Activity Report Paris 2019

Edition: 2020-03-23

List of Inria's Research Teams

Project-Team ALMANACH

Automatic Language Modelling and Analysis & Computational Humanities

RESEARCH CENTER
Paris

THEME
Language, Speech and Audio

Table of contents

Project-Team ALMANACH

Creation of the Project-Team: 2019 July 01

Keywords:

Computer Science and Digital Science:

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
A3.4.6. - Neural networks
A3.4.8. - Deep learning
A9.1. - Knowledge
A9.2. - Machine learning
A9.4. - Natural language processing
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B1.2.2. - Cognitive science
B1.2.3. - Computational neurosciences
B9.1.1. - E-learning, MOOC
B9.5.6. - Data science
B9.6.5. - Sociology
B9.6.6. - Archeology, History
B9.6.8. - Linguistics
B9.6.10. - Digital humanities
B9.7. - Knowledge dissemination
B9.7.1. - Open access
B9.7.2. - Open data
B9.8. - Reproducibility

1. Team, Visitors, External Collaborators

Research Scientists

Benoît Sagot [Team leader, Inria, Researcher, HDR]
Pierre Boullier [Inria, Emeritus, until Apr 2019]
Laurent Romary [Inria, Senior Researcher, HDR]
Djamé Seddah [Inria, Researcher, “détachement” at Inria from Sorbonne Univ.]
Éric Villemonte de La Clergerie [Inria, Researcher]

Faculty Member

Kim Gerdes [Univ Sorbonne Nouvelle, Associate Professor, from Mar 2019 (“délégation”), HDR]

Technical Staff

Achraf Azhar [Inria, Engineer, until Apr 2019]
Alix Chagué [Inria, Engineer, from Feb 2019]

Farah Essaidi [Inria, Engineer, from Oct 2019]
Clémentine Fourrier [Inria, Engineer, from Mar 2019 until Sep 2019]
Ganesh Jawahar [Inria, Engineer, until Sep 2019]
Tanti Kristanti [Inria, Engineer]
Alba Marina Malaga Sabogal [Inria, Engineer, until Aug 2019]
Charles Riondet [Inria, Engineer, until May 2019]
Dorian Seillier [Inria, Engineer, until May 2019]
Lionel Tadonfouet [Inria, Engineer]

PhD Students

Jack Bowers [Austrian Academy of Sciences, PhD Student]
Marine Courtin [Sorbonne Université, PhD Student, from Mar 2019]
Yixuan Li [Sorbonne Université, PhD Student, from Mar 2019]
Clémentine Fourrier [Inria, PhD Student, from Oct 2019]
Loïc Grobol [École Normale Supérieure de Paris, PhD Student]
Mohamed Khemakhem [Inria, PhD Student]
Louis Martin [Facebook, PhD Student]
Benjamin Muller [Inria, PhD Student]
Pedro Ortiz Suárez [Inria, PhD Student]
Mathilde Regnault [École Normale Supérieure de Paris, PhD Student]
Jose Rosales Nuñez [CNRS, PhD Student]

Post-Doctoral Fellows

Yoann Dupont [Univ d'Orléans, Post-Doctoral Fellow, from Jun 2019]
Murielle Fabre [Inria, Post-Doctoral Fellow]
Gaël Guibon [Univ Denis Diderot, Post-Doctoral Fellow, from Jun 2019]
Ilia Markov [Inria, Post-Doctoral Fellow, until May 2019]

Administrative Assistant

Meriem Guemair [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

The ALMAnaCH project-team [0] brings together specialists of a pluri-disciplinary research domain at the interface between computer science, linguistics, statistics, and the humanities, namely that of **natural language processing**, **computational linguistics** and **digital and computational humanities and social sciences**.

Computational linguistics is an interdisciplinary field dealing with the computational modelling of natural language. Research in this field is driven both by the theoretical goal of understanding human language and by practical applications in **Natural Language Processing** (hereafter NLP) such as linguistic analysis (syntactic and semantic parsing, for instance), machine translation, information extraction and retrieval and human-computer dialogue. Computational linguistics and NLP, which date back at least to the early 1950s, are among the key sub-fields of **Artificial Intelligence**.

Digital Humanities and social sciences (hereafter DH) is an interdisciplinary field that uses computer science as a source of techniques and technologies, in particular NLP, for exploring research questions in social sciences and humanities. **Computational Humanities** and computational social sciences aim at improving the state of the art in both computer sciences (e.g. NLP) and social sciences and humanities, by involving computer science as a research field.

[0] ALMAnaCH was created as an Inria team ("équipe") on the 1st January, 2017 and as a project-team on the 1st July 2019.

ALMAnaCH is a follow-up to the ALPAGE project-team, which came to an end in December 2016. ALPAGE was created in 2007 in collaboration with Paris-Diderot University and had the status of an UMR-I since 2009. This joint team involved computational linguists from Inria as well as computational linguists from Paris-Diderot University with a strong background in linguistics, and proved successful. However, the context has changed since then, with the recent emergence of digital humanities and, more importantly, of computational humanities. This presents both an opportunity and a challenge for Inria computational linguists, as it provides them with new types of data (on which their tools, resources and algorithms can be used, thereby leading to new results in human sciences), as well as with new and challenging research problems, which, if solved, provide new ways of studying human sciences.

The scientific positioning of ALMAnaCH therefore extends that of ALPAGE. We remain committed to developing state-of-the-art NLP software and resources that can be used by academics and in the industry, including recent approaches based on deep learning. At the same time we continue our work on language modelling in order to provide a better understanding of languages, an objective that is reinforced and addressed in the broader context of computational humanities. Finally, we remain dedicated to having an impact on the industrial world and more generally on society, via multiple types of collaboration with companies and other institutions (startup creation, industrial contracts, expertise, etc.).

One of the main challenges in computational linguistics is **to model and to cope with language variation**. Language varies with respect to domain and genre (news wires, scientific literature, poetry, oral transcripts...), sociolinguistic factors (age, background, education; variation attested for instance on social media), geographical factors (dialects) and other dimensions (disabilities, for instance). But language also constantly evolves at all time scales. Addressing this variability is still an open issue for NLP. Commonly used approaches, which often rely on supervised and semi-supervised machine learning methods, require very large amounts of annotated data. They still suffer from the high level of variability found for instance in **user-generated content**, **non-contemporary texts**, as well as in **domain-specific documents** (e.g. financial, legal).

ALMAnaCH tackles the challenge of language variation in two complementary directions, supported by a third, transverse research axis on language resources. These three research axes do not reflect an internal organisation of eparate teams. They are meant to structure our scientific agenda, and most members of the project-team are involved in two or all of them.

ALMAnaCH's research axes, themselves structured in sub-axis, are the following:

1. Automatic Context-augmented Linguistic Analysis
    1. Processing of natural language at all levels: morphology, syntax, semantics
    2. Integrating context in NLP systems
    3. Information and knowledge extraction
2. Computational Modelling of Linguistic Variation
    1. Theoretical and empirical synchronic linguistics
    2. Sociolinguistic variation
    3. Diachronic variation
    4. Accessibility-related variation
3. Modelling and development of Language Resources
    1. Construction, management and automatic annotation of text corpora
    2. Development of lexical resources
    3. Development of annotated corpora

3. Research Program

3.1. Research strands

As described above, ALMAnaCH's scientific programme is organised around three research axes. The first two aim to tackle the challenge of language variation in two complementary directions. They are supported by a third, transverse research axis on language resources. Our four-year objectives are described in much greater detail in the project-team proposal, whose very recent final validation in June 2019 resulted in the upgrade of ALMAnaCH to the "project-team" status in July 2019. They can be summarised as follows:

3.1.1. Research axis 1

Our first objective is to **stay at a state-of-the-art level in key NLP tasks** such as shallow processing, part-of-speech tagging and (syntactic) parsing, which are core expertise domains of ALMAnaCH members. This will also require us to improve the **generation of semantic representations (semantic parsing)**, and to begin to explore tasks such as machine translation, which now relies on neural architectures also used for some of the above-mentioned tasks. Given the generalisation of neural models in NLP, we will also be involved in better understanding how such models work and what they learn, something that is directly related to the investigation of language variation (Research axis 2). We will also work on the **integration of both linguistic and non-linguistic contextual information** to improve automatic linguistic analysis. This is an emerging and promising line of research in NLP. We will have to identify, model and take advantage of each type of contextual information available. Addressing these issues will enable the development of new lines of research related to conversational content. Applications include improved information and knowledge extraction algorithms. We will especially focus on challenging datasets such as domain-specific texts (e.g. financial, legal) as well as historical documents, in the larger context of the development of digital humanities. We currently also explore the even more challenging new direction of a cognitively inspired NLP, in order to tackle the possibility to enrich the architecture of state-of-the-art algorithms, such as RNNGs, based on human neuroimaging-driven data.

3.1.2. Research axis 2

Language variation must be better understood and modelled in all its forms. In this regard, we will put a strong emphasis on **four types** of language variation and their mutual interaction: **sociolinguistic variation** in synchrony (including non-canonical spelling and syntax in user-generated content), **complexity-based variation** in relation to language-related disabilities, and **diachronic variation** (computational exploration of language change and language history, with a focus on Old to all forms of Modern French, as well as Indo-European languages in general). In addition, the noise introduced by Optical Character Recognition and Handwritten Text Recognition systems, especially in the context of historical documents, bears some similarities to that of non-canonical input in user-generated content (e.g. erroneous characters). This noise constitutes a more transverse kind of variation stemming from the way language is graphically encoded, which we call **language-encoding variation**. Other types of language variation will also become important research topics for ALMAnaCH in the future. This includes dialectal variation (e.g. work on Arabic varieties, something on which we have already started working, producing the first annotated data set on Maghrebi Arabizi, the Arabic variants used on social media by people from North-African countries, written using a non-fixed Latin-script transcription) as well as the study and exploitation of paraphrases in a broader context than the above-mentioned complexity-based variation.

Both research axes above rely on the availability of language resources (corpora, lexicons), which is the focus of our third, transverse research axis.

3.1.3. Research axis 3

Language resource development (raw and annotated corpora, lexical resources) is not just a necessary preliminary step to create both evaluation datasets for NLP systems and training datasets for NLP systems based on machine learning. When dealing with datasets of interest to researchers from the humanities (e.g. large archives), it is also a goal *per se* and a preliminary step before making such datasets available and exploitable online. It involves a number of scientific challenges, among which (i) tackling issues related to the digitalisation of non-electronic datasets, (ii) tackling issues related to the fact that many DH-related datasets are domain-specific and/or not written in contemporary languages; (iii) the development of semi-automatic and automatic algorithms to speed up the work (e.g. automatic extraction of lexical information, low-resource learning for the development of pre-annotation algorithms, transfer methods to leverage existing tools and/or resources for other languages, etc.) and (iv) the development of formal models to represent linguistic information in the best possible way, thus requiring expertise at least in NLP and in typological and formal linguistics. Such endeavours are domains of expertise of the ALMAnaCH team, and a large part of our research activities will be dedicated to language resource development. In this regard, we aim to retain our leading role in the representation and management of lexical resource and treebank development and also to develop a complete processing line for the transcription, analysis and processing of complex documents of interest to the humanities, in particular archival documents. This research axis 3 will benefit the whole team and beyond, and will benefit from and feed the work of the other research axes.

3.2. Automatic Context-augmented Linguistic Analysis

This first research strand is centred around NLP technologies and some of their applications in Artificial Intelligence (AI). Core NLP tasks such as part-of-speech tagging, syntactic and semantic parsing is improved by integrating new approaches, such as (deep) neural networks, whenever relevant, while preserving and taking advantage of our expertise on symbolic and statistical system: hybridisation not only couples symbolic and statistical approaches, but neural approaches as well. AI applications are twofold, notwithstanding the impact of language variation (see the next strand): (i) information and knowledge extraction, whatever the type of input text (from financial documents to ancient, historical texts and from Twitter data to Wikipedia) and (ii) chatbots and natural language generation. In many cases, our work on these AI applications is carried out in collaboration with industrial partners. The specificities and issues caused by language variation (a text in Old French, a contemporary financial document and tweets with a non-canonical spelling cannot be processed in the same way) are addressed in the next research strand.

3.2.1. Processing of natural language at all levels: morphology, syntax, semantics

Our expertise in NLP is the outcome of more than 10 years in developing new models of analysis and accurate techniques for the full processing of any kind of language input since the early days of the Atoll project-team and the rise of linguistically informed data-driven models as put forward within the Alpage project-team.

Traditionally, a full natural language process (NLP) chain is organised as a pipeline where each stage of analysis represents a traditional linguistic field (in a *structuralism* view) from morphological analysis to purely semantic representations. The problem is that this architecture is vulnerable to error propagation and very domain sensitive: each of these stage must be compatible at the lexical and structure levels they provide. We arguably built the best performing NLP chain for French [74], [112] and one of the best for robust multilingual parsing as shown by our results in various shared tasks over the years [108], [105], [111], [82]. So we pursue our efforts on each of our components we developed: tokenisers (e.g. SxPipe), part-of-speech taggers (e.g. MElt), constituency parsers and dependency parsers (e.g. FRMG, DyALog-SR) as well as our recent neural semantic graph parsers [105].

In particular, we continue to explore the hybridisation of symbolic and statistical approaches, and extend it to neural approaches, as initiated in the context of our participation to the CoNLL 2017 multilingual parsing shared task [0] and to Extrinsic Parsing Evaluation Shared Task [0].

[0] We ranked 3 for UPOS tagging and 6 for dependency parsing out of 33 participants.

[0] Semantic graph parsing, evaluated on biomedical data, speech and opinion. We ranked 1 in a joint effort with the Stanford NLP team

Fundamentally, we want to build tools that are less sensitive to variation, more easily configurable, and self-adapting. Our short-term goal is to explore techniques such as multi-task learning (cf. already [110]) to propose a joint model of tokenisation, normalisation, morphological analysis and syntactic analysis. We also explore adversarial learning, considering the drastic variation we face in parsing user-generated content and processing historical texts, both seen as noisy input that needs to be handled at training and decoding time.

3.2.2. Integrating context in NLP systems

While those points are fundamental, therefore necessary, if we want to build the next generation of NLP tools, we need to *push the envelop* even further by tackling the biggest current challenge in NLP: handling the context within which a speech act is taking place.

There is indeed a strong tendency in NLP to assume that each sentence is independent from its siblings sentences as well as its context of enunciation, with the obvious objective to simplify models and reduce the complexity of predictions. While this practice is already questionable when processing full-length edited documents, it becomes clearly problematic when dealing with short sentences that are noisy, full of ellipses and external references, as commonly found in User-Generated Content (UGC).

A more expressive and context-aware structural representation of a linguistic production is required to accurately model UGC. Let us consider for instance the case for Syntax-based Machine Translation of social media content, as is carried out by the ALMAnaCH-led ANR project Parsiti (PI: DS). A Facebook post may be part of a discussion thread, which may include links to external content. Such information is required for a complete representation of the post's context, and in turn its accurate machine translation. Even for the presumably simpler task of POS tagging of dialogue sequences, the addition of context-based features (namely information about the speaker and dialogue moves) was beneficial [85]. In the case of UGC, working across sentence boundaries was explored for instance, with limited success, by [73] for document-wise parsing and by [96] for POS tagging.

Taking the context into account requires new inference methods able to share information between sentences as well as new learning methods capable of finding out which information is to be made available, and where. Integrating contextual information at all steps of an NLP pipeline is among the main research questions addressed in this research strand. In the short term, we focus on morphological and syntactic disambiguation within close-world scenarios, as found in video games and domain-specific UGC. In the long term, we investigate the integration of linguistically motivated semantic information into joint learning models.

From a more general perspective, contexts may take many forms and require imagination to discern them, get useful data sets, and find ways to exploit them. A context may be a question associated with an answer, a rating associated with a comment (as provided by many web services), a thread of discussions (e-mails, social media, digital assistants, chatbots—on which see below–), but also meta data about some situation (such as discussions between gamers in relation with the state of the game) or multiple points of views (pictures and captions, movies and subtitles). Even if the relationship between a language production and its context is imprecise and indirect, it is still a valuable source of information, notwithstanding the need for less supervised machine learning techniques (cf. the use of LSTM neural networks by Google to automatically suggest replies to emails).

3.2.3. Information and knowledge extraction

The use of local contexts as discussed above is a new and promising approach. However, a more traditional notion of global context or world knowledge remains an open question and still raises difficult issues. Indeed, many aspects of language such as ambiguities and ellipsis can only be handled using world knowledge. Linked Open Data (LODs) such as DBpedia, WordNet, BabelNet, or Framebase provide such knowledge and we plan to exploit them.

However, each specialised domain (economy, law, medicine...) exhibits its own set of concepts with associated terms. This is also true of communities (e.g. on social media), and it is even possible to find communities discussing the same topics (e.g. immigration) with very distinct vocabularies. Global LODs weakly related to language may be too general and not sufficient for a specific language variant. Following and extending previous work in ALPAGE, we put an emphasis on information acquisition from corpora, including error mining techniques in parsed corpora (to detect specific usages of a word that are missing in existing resources), terminology extraction, and word clustering.

Word clustering is of specific importance. It relies on the distributional hypothesis initially formulated by Harris, which states that words occurring in similar contexts tend to be semantically close. The latest developments of these ideas (with word2vec or GloVe) have led to the embedding of words (through vectors) in low-dimensional semantic spaces. In particular, words that are typical of several communities (see above) can be embedded in a same semantic space in order to establish mappings between them. It is also possible in such spaces to study static configurations and vector shifts with respect to variables such as time, using topological theories (such as pretopology), for instance to explore shifts in meaning over time (cf. the ANR project Profiterole concerning ancient French texts) or between communities (cf. the ANR project SoSweet). It is also worth mentioning on-going work (in computational semantics) whose goal is to combine word embeddings to embed expressions, sentences, paragraphs or even documents into semantic spaces, e.g. to explore the similarity of documents at various time periods.

Besides general knowledge about a domain, it is important to detect and keep trace of more specific pieces of information when processing a document and maintaining a context, especially about (recurring) Named Entities (persons, organisations, locations...) —something that is the focus of future work in collaboration with Patrice Lopez on named entity detection in scientific texts. Through the co-supervision of a PhD funded by the LabEx EFL (see below), we are also involved in pronominal coreference resolution (finding the referent of pronouns). Finally, we plan to continue working on deeper syntactic representations (as initiated with the Deep Sequoia Treebank), thus paving the way towards deeper semantic representations. Such information is instrumental when looking for more precise and complete information about who does what, to whom, when and where in a document. These lines of research are motivated by the need to extract useful contextual information, but it is also worth noting their strong potential in industrial applications.

3.3. Computational Modelling of Linguistic Variation

NLP and DH tools and resources are very often developed for contemporary, edited, non-specialised texts, often based on journalistic corpora. However, such corpora are not representative of the variety of existing textual data. As a result, the performance of most NLP systems decreases, sometimes dramatically, when faced with non-contemporary, non-edited or specialised texts. Despite the existence of domain-adaptation techniques and of robust tools, for instance for social media text processing, dealing with linguistic variation is still a crucial challenge for NLP and DH.

Linguistic variation is not a monolithic phenomenon. Firstly, it can result from different types of processes, such as variation over time (diachronic variation) and variation correlated with sociological variables (sociolinguistic variation, especially on social networks). Secondly, it can affect all components of language, from spelling (languages without a normative spelling, spelling errors of all kinds and origins) to morphology/syntax (especially in diachrony, in texts from specialised domains, in social media texts) and semantics/pragmatics (again in diachrony, for instance). Finally, it can constitute a property of the data to be analysed or a feature of the data to be generated (for instance when trying to simplify texts for increasing their accessibility for disabled and/or non-native readers).

Nevertheless, despite this variability in variation, the underlying mechanisms are partly comparable. This motivates our general vision that many generic techniques could be developed and adapted to handle different types of variation. In this regard, three aspects must be kept in mind: spelling variation (human errors, OCR/HTR errors, lack of spelling conventions for some languages...), lack or scarcity of parallel data aligning "variation-affected" texts and their "standard/edited" counterpart, and the sequential nature of the problem at hand. We will therefore explore, for instance, how unsupervised or weakly-supervised techniques could be developed and feed dedicated sequence-to-sequence models. Such architectures could help develop "normalisation" tools adapted, for example, to social media texts, texts written in ancient/dialectal varieties of well-resourced languages (e.g. Old French texts), and OCR/HTR system outputs.

Nevertheless, the different types of language variation will require specific models, resources and tools. All these directions of research constitute the core of our second research strand described in this section.

3.3.1. Theoretical and empirical synchronic linguistics

Permanent members involved: all

We aim to explore computational models to deal with language variation. It is important to get more insights about language in general and about the way humans apprehend it. We will do so in at least two directions, associating computational linguistics with formal and descriptive linguistics on the one hand (especially at the morphological level) and with cognitive linguistics on the other hand (especially at the syntactic level).

Recent advances in morphology rely on quantitative and computational approaches and, sometimes, on collaboration with descriptive linguists—see for instance the special issue of the *Morphology* journal on "computational methods for descriptive and theoretical morphology", edited and introduced by [70]. In this regard, ALMAnaCH members have taken part in the design of quantitative approaches to defining and measuring morphological complexity and to assess the internal structure of morphological systems (inflection classes, predictability of inflected forms...). Such studies provide valuable insights on these prominent questions in theoretical morphology. They also improve the linguistic relevance and the development speed of NLP-oriented lexicons, as also demonstrated by ALMAnaCH members. We shall therefore pursue these investigations, and orientate them towards their use in diachronic models (see section 3.3.3).

Regarding cognitive linguistics, we have the perfect opportunity with the starting ANR-NSF project "Neuro-Computational Models of Natural Language" (NCM-NL) to go in this direction, by examining potential correlations between medical imagery applied on patients listening to a reading of "Le Petit Prince" and computation models applied on the novel. A secondary prospective benefit from the project will be information about processing evolution (by the patients) along the novel, possibly due to the use of contextual information by humans.

3.3.2. Sociolinguistic variation

Because language is central in our social interactions, it is legitimate to ask how the rise of digital content and its tight integration in our daily life has become a factor acting on language. This is even more actual as the recent rise of novel digital services opens new areas of expression, which support new linguistic behaviours. In particular, social media such as Twitter provide channels of communication through which speakers/writers use their language in ways that differ from standard written and oral forms. The result is the emergence of new language varieties.

A very similar situation exists with regard to historical texts, especially documentary texts or graffiti but even literary texts, that do not follow standardised orthography, morphology or syntax.

However, NLP tools are designed for standard forms of language and exhibit a drastic loss of accuracy when applied to social media varieties or non-standardised historical sources. To define appropriate tools, descriptions of these varieties are needed. However, to validate such descriptions, tools are also needed. We address this chicken-and-egg problem in an interdisciplinary fashion, by working both on linguistic descriptions and on the development of NLP tools. Recently, socio-demographic variables have been shown to bear a strong impact on NLP processing tools (see for instance [80] and references therein). This is why, in a first step, jointly with researchers involved in the ANR project SoSweet (ENS Lyon and Inria project-team Dante), we will study how these variables can be factored out by our models and, in a second step, how they can be accurately predicted from sources lacking these kinds of featured descriptions.

3.3.3. Diachronic variation

Language change is a type of variation pertaining to the diachronic axis. Yet any language change, whatever its nature (phonetic, syntactic...), results from a particular case of synchronic variation (competing phonetic realisations, competing syntactic constructions...). The articulation of diachronic and synchronic variation is influenced to a large extent by both language-internal factors (i.e. generalisation of context-specific facts) and/or external factors (determined by social class, register, domain, and other types of variation).

Very few computational models of language change have been developed. Simple deterministic finite-state-based phonetic evolution models have been used in different contexts. The PIElexicon project [90] uses such models to automatically generate forms attested in (classical) Indo-European languages but is based on an idiosyncratic and unacceptable reconstruction of the Proto-Indo-European language. Probabilistic finite-state models have also been used for automatic cognate detection and proto-form reconstruction, for example by [71] and [81]. Such models rely on a good understanding of the phonetic evolution of the languages at hand.

In ALMAnaCH, our goal is to work on modelling phonetic, morphological and lexical diachronic evolution, with an emphasis on computational etymological research and on the computational modelling of the evolution of morphological systems (morphological grammar and morphological lexicon). These efforts will be in direct interaction with sub-strand 3b (development of lexical resources). We want to go beyond the above-mentioned purely phonetic models of language and lexicon evolution, as they fail to take into account a number of crucial dimensions, among which: (1) spelling, spelling variation and the relationship between spelling and phonetics; (2) synchronic variation (geographical, genre-related, etc.); (3) morphology, especially through intra-paradigmatic and inter-paradigmatic analogical levelling phenomena, (4) lexical creation, including via affixal derivation, back-formation processes and borrowings.

We apply our models to two main tasks. The first task, as developed for example in the context of the ANR project Profiterole, consists in predicting non-attested or non-documented words at a certain date based on attestations of older or newer stages of the same word (e.g., predicting a non-documented Middle French word based on its Vulgar Latin and Old French predecessors and its Modern French successor). Morphological models and lexical diachronic evolution models will provide independent ways to perform the same predictions, thus reinforcing our hypotheses or pointing to new challenges.

The second application task is computational etymology and proto-language reconstruction. Our lexical diachronic evolution models will be paired with semantic resources (wordnets, word embeddings, and other corpus-based statistical information). This will allow us to formally validate or suggest etymological or cognate relations between lexical entries from different languages of a same language family, provided they are all inherited. Such an approach could also be adapted to include the automatic detection of borrowings from one language to another (e.g. for studying the non-inherited layers in the Ancient Greek lexicon). In the longer term, we will investigate the feasibility of the automatic (unsupervised) acquisition of phonetic change models, especially when provided with lexical data for numerous languages from the same language family.

These lines of research will rely on etymological data sets and standards for representing etymological information (see Section 3.4.2).

Diachronic evolution also applies to syntax, and in the context of the ANR project Profiterole, we are beginning to explore more or less automatic ways of detecting these evolutions and suggest modifications, relying on fine-grained syntactic descriptions (as provided by meta-grammars), unsupervised sentence clustering (generalising previous works on error mining, cf. [6]), and constraint relaxation (in meta-grammar classes). The underlying idea is that a new syntactic construction evolves from a more ancient one by small, iterative modifications, for instance by changing word order, adding or deleting functional words, etc.

3.3.4. Accessibility-related variation

Language variation does not always pertain to the textual input of NLP tools. It can also be characterised by their intended output. This is the perspective from which we investigate the issue of text simplification (for a recent survey, see for instance [109]). Text simplification is an important task for improving the accessibility to information, for instance for people suffering from disabilities and for non-native speakers learning a given language [91]. To this end, guidelines have been developed to help writing documents that are easier to read and understand, such as the FALC ("Facile À Lire et à Comprendre") guidelines for French. [0]

[0] Please click here for an archived version of these guidelines (at the time this footnote is begin written, the original link does not seem to work any more).

Fully automated text simplification is not suitable for producing high-quality simplified texts. Besides, the involvement of disabled people in the production of simplified texts plays an important social role. Therefore, following previous works [79], [103], our goal will be to develop tools for the computer-aided simplification of textual documents, especially administrative documents. Many of the FALC guidelines can only be linguistically expressed using complex, syntactic constraints, and the amount of available "parallel" data (aligned raw and simplified documents) is limited. We will therefore investigate hybrid techniques involving rule-based, statistical and neural approaches based on parsing results (for an example of previous parsing-based work, see [68]). Lexical simplification, another aspect of text simplification [86], [92], will also be pursued. In this regard, we have already started a collaboration with Facebook's AI Research in Paris, the UNAPEI (the largest French federation of associations defending and supporting people with intellectual disabilities and their families), and the French Secretariat of State in charge of Disabled Persons.

Accessibility can also be related to the various presentation forms of a document. This is the context in which we have initiated the OPALINE project, funded by the *Programme d'Investissement d'Avenir - Fonds pour la Société Numérique*. The objective is for us to further develop the GROBID text-extraction suite [0] in order to be able to re-publish existing books or dictionaries, available in PDF, in a format that is accessible by visually impaired persons.

3.4. Modelling and Development of Language Resources

Language resources (raw and annotated corpora, lexical resources, etc.) are required in order to apply any machine learning technique (statistical, neural, hybrid) to an NLP problem, as well as to evaluate the output of an NLP system.

In data-driven, machine-learning-based approaches, language resources are the place where linguistic information is stored, be it implicitly (as in raw corpora) or explicitly (as in annotated corpora and in most lexical resources). Whenever linguistic information is provided explicitly, it complies to guidelines that formally define which linguistic information should be encoded, and how. Designing linguistically meaningful and computationally exploitable ways to encode linguistic information within language resources constitutes the first main scientific challenge in language resource development. It requires a strong expertise on both the linguistic issues underlying the type of resource under development (e.g. on syntax when developing a treebank) and the NLP algorithms that will make use of such information.

The other main challenge regarding language resource development is a consequence of the fact that it is a costly, often tedious task. ALMAnaCH members have a long track record of language resource development, including by hiring, training and supervising dedicated annotators. But a manual annotation can be speeded up by automatic techniques. ALMAnaCH members have also work on such techniques, and published work on approaches such as automatic lexical information extraction, annotation transfer from a language to closely related languages, and more generally on the use of pre-annotation tools for treebank development and on the impact of such tools on annotation speed and quality. These techniques are often also relevant for Research strand 1. For example, adapting parsers from one language to the other or developing parsers that work on more than one language (e.g. a non-lexicalised parser trained on the concatenation of treebanks from different languages in the same language family) can both improve parsing results on low-resource languages and speed up treebank development for such languages.

3.4.1. Construction, management and automatic annotation of Text Corpora

Corpus creation and management (including automatic annotation) is often a time-consuming and technically challenging task. In many cases, it also raises scientific issues related for instance with linguistic questions (what is the elementary unit in a text?) as well as computer-science challenges (for instance when OCR or HTR are involved). It is therefore necessary to design a work-flow that makes it possible to deal with data collections, even if they are initially available as photos, scans, wikipedia dumps, etc.

[0] https://github.com/kermitt2/grobid

These challenges are particularly relevant when dealing with ancient languages or scripts where fonts, OCR techniques, language models may be not extant or of inferior quality, as a result, among others, of the variety of writing systems and the lack of textual data. We will therefore work on improving print OCR for some of these languages, especially by moving towards joint OCR and language models. Of course, contemporary texts can be often gathered in very large volumes, as we already do within the ANR project SoSweet, resulting in different, specific issues.

ALMAnaCH pays a specific attention to the re-usability [0] of all resources produced and maintained within its various projects and research activities. To this end, we will ensure maximum compatibility with available international standards for representing textual sources and their annotations. More precisely we will take the TEI (*Text Encoding Initiative*) guidelines as well the standards produced by ISO committee TC 37/SC 4 as essential points of reference.

From our ongoing projects in the field of Digital Humanities and emerging initiatives in this field, we observe a real need for complete but easy work-flows for exploiting corpora, starting from a set of raw documents and reaching the level where one can browse the main concepts and entities, explore their relationship, extract specific pieces of information, always with the ability to return to (fragments of) the original documents. The pieces of information extracted from the corpora also need to be represented as knowledge databases (for instance as RDF "linked data"), published and linked with other existing databases (for instance for people and locations).

The process may be seen as progressively enriching the documents with new layers of annotations produced by various NLP modules and possibly validated by users, preferably in a collaborative way. It relies on the use of clearly identified representation formats for the annotations, as advocated within ISO TC 37/SC 4 standards and the TEI guidelines, but also on the existence of well-designed collaborative interfaces for browsing, querying, visualisation, and validation. ALMAnaCH has been or is working on several of the NLP bricks needed for setting such a work-flow, and has a solid expertise in the issues related to standardisation (of documents and annotations). However, putting all these elements in a unified work-flow that is simple to deploy and configure remains to be done. In particular, work-flow and interface should maybe not be dissociated, in the sense that the work-flow should be easily piloted and configured from the interface. An option will be to identify pertinent emerging platforms in DH (such as Transkribus) and to propose collaborations to ensure that NLP modules can be easily integrated.

It should be noted that such work-flows have actually a large potential besides DH, for instance for exploiting internal documentation (for a company) or exploring existing relationships between entities.

3.4.2. Development of Lexical Resources

ALPAGE, the Inria predecessor of ALMAnaCH, has put a strong emphasis in the development of morphological, syntactic and wordnet-like semantic lexical resources for French as well as other languages (see for instance [5], [1]). Such resources play a crucial role in all NLP tools, as has been proven among other tasks for POS tagging [101], [97], [111] and parsing, and some of the lexical resource development will be targeted towards the improvement of NLP tools. They will also play a central role for studying diachrony in the lexicon, for example for Ancient to Contemporary French in the context of the Profiterole project. They will also be one of the primary sources of linguistic information for augmenting language models used in OCR systems for ancient scripts, and will allow us to develop automatic annotation tools (e.g. POS taggers) for low-resourced languages (see already [113]), especially ancient languages. Finally, semantic lexicons such as wordnets will play a crucial role in assessing lexical similarity and automating etymological research.

Therefore, an important effort towards the development of new morphological lexicons will be initiated, with a focus on ancient languages of interest. Following previous work by ALMAnaCH members, we will try and leverage all existing resources whenever possible such as electronic dictionaries, OCRised dictionaries, both modern and ancient [100], [83], [102], while using and developing (semi)automatic lexical information extraction techniques based on existing corpora [98], [104]. A new line of research will be to integrate

[0] From a larger point of view we intend to comply with the so-called FAIR principles (http://force11.org/group/fairgroup/fairprinciples).

the diachronic axis by linking lexicons that are in diachronic relation with one another thanks to phonetic and morphological change laws (e.g. XIIth century French with XVth century French and contemporary French). Another novelty will be the integration of etymological information in these lexical resources, which requires the formalisation, the standardisation, and the extraction of etymological information from OCRised dictionaries or other electronic resources, as well as the automatic generation of candidate etymologies. These directions of research are already investigated in ALMAnaCH [83], [102].

An underlying effort for this research will be to further the development of the GROBID-dictionaries software, which provides cascading CRF (Conditional Random Fields) models for the segmentation and analysis of existing print dictionaries. The first results we have obtained have allowed us to set up specific collaborations to improve our performances in the domains of a) recent general purpose dictionaries such as the Petit Larousse (Nénufar project, funded by the DGLFLF in collaboration with the University of Montpellier), b) etymological dictionaries (in collaboration with the Berlin Brandenburg Academy of sciences) and c) patrimonial dictionaries such as the Dictionnaire Universel de Basnage (an ANR project, including a PhD thesis at ALMAnaCH, has recently started on this topic in collaboration with the University of Grenoble-Alpes and the University Sorbonne Nouvelle in Paris).

In the same way as we signalled the importance of standards for the representation of interoperable corpora and their annotations, we will keep making the best use of the existing standardisation background for the representation of our various lexical resources. There again, the TEI guidelines play a central role, and we have recently participated in the "TEI Lex 0" initiative to provide a reference subset for the "Dictionary" chapter of the guidelines. We are also responsible, as project leader, of the edition of the new part 4 of the ISO standard 24613 (LMF, Lexical Markup Framework) [94] dedicated to the definition of the TEI serialisation of the LMF model (defined in ISO 24613 part 1 'Core model', 2 'Machine Readable Dictionaries' and 3 'Etymology'). We consider that contributing to standards allows us to stabilise our knowledge and transfer our competence.

3.4.3. Development of Annotated Corpora

Along with the creation of lexical resources, ALMAnaCH is also involved in the creation of corpora either fully manually annotated (gold standard) or automatically annotated with state-of-the-art pipeline processing chains (silver standard). Annotations will either be only morphosyntactic or will cover more complex linguistic levels (constituency and/or dependency syntax, deep syntax, maybe semantics). Former members of the ALPAGE project have a renowned experience in those aspects (see for instance [107], [93], [106], [88]) and will participate to the creation of valuable resources originating from the historical domain genre.

Under the auspices of the ANR Parsiti project, led by ALMAnaCH (PI: DS), we aim to explore the interaction of extra-linguistic context and speech acts. Exploiting extra-linguistics context highlights the benefits of expanding the scope of current NLP tools beyond unit boundaries. Such information can be of spatial and temporal nature, for instance. They have been shown to improve Entity Linking over social media streams [76]. In our case, we decided to focus on a closed world scenario in order to study context and speech acts interaction. To do so, we are developing a multimodal data set made of live sessions of a first person shooter video game (Alien vs. Predator) where we transcribed all human players interactions and face expressions streamlined with a log of all in-game events linked to the video recording of the game session, as well as the recording of the human players themselves. The in-games events are ontologically organised and enable the modelling of the extra-linguistics context with different levels of granularity. Recorded over many games sessions, we already transcribed over 2 hours of speech that will serve as a basis for exploratory work, needed for the prototyping of our context-enhanced NLP tools. In the next step of this line of work, we will focus on enriching this data set with linguistic annotations, with an emphasis on co-references resolutions and predicate argument structures. The midterm goal is to use that data set to validate a various range of approaches when facing multimodal data in a close-world environment.

4. Application Domains

4.1. Application domains for ALMAnaCH

ALMAnaCH's research areas cover Natural Language Processing (nowadays identified as a sub-domain of Artificial Intelligence) and Digital Humanities. Application domains are therefore numerous, as witnessed by ALMAnaCH's multiple academic and industrial collaborations, for which see the relevant sections. Examples of application domains for NLP include:

- Information extraction, information retrieval, text mining (ex.: opinion surveys)
- Text generation, text simplification, automatic summarisation
- Spelling correction (writing aid, post-OCR, normalisation of noisy/non-canonical texts)
- Machine translation, computer-aided translation
- Chatbots, conversational agents, question answering systems
- Medical applications (early diagnosis, language-based medical monitoring...)
- Applications in linguistics (modelling languages and their evolution, sociolinguistic studies...)
- Digital humanities (exploitation of text documents, for instance in historical research)

5. Highlights of the Year

5.1. Highlights of the Year

The main highlight for ALMAnaCH in 2019 is the publication of CamemBERT, a French neural language model trained on the French section of OSCAR, our very large multilingual web-based raw corpus. The publication of this first Transformer-based language model for French, which allowed us to improve the state of the art in several classical NLP tasks, met a large success both in the academic and industrial worlds. It is the topic of an article in the major French daily newspaper Le Monde and of a broadcast on the national radio France Culture.

6. New Software and Platforms

6.1. Enqi

- Author: Benoît Sagot
- Contact: Benoît Sagot

6.2. SYNTAX

KEYWORD: Parsing

FUNCTIONAL DESCRIPTION: Syntax system includes various deterministic and non-deterministic CFG parser generators. It includes in particular an efficient implementation of the Earley algorithm, with many original optimizations, that is used in several of Alpage's NLP tools, including the pre-processing chain Sx Pipe and the LFG deep parser SxLfg . This implementation of the Earley algorithm has been recently extended to handle probabilistic CFG (PCFG), by taking into account probabilities both during parsing (beam) and after parsing (n-best computation).

- Participants: Benoît Sagot and Pierre Boullier
- Contact: Pierre Boullier
- URL: http://syntax.gforge.inria.fr/

6.3. FRMG

KEYWORDS: Parsing - French

FUNCTIONAL DESCRIPTION: FRMG is a large-coverage linguistic meta-grammar of French. It can be compiled (using MGCOMP) into a Tree Adjoining Grammar, which, in turn, can be compiled (using DyALog) into a parser for French.

- Participant: Eric de La Clergerie
- Contact: Éric De La Clergerie
- URL: http://mgkit.gforge.inria.fr/

6.4. MElt

Maximum-Entropy lexicon-aware tagger

KEYWORD: Part-of-speech tagger

FUNCTIONAL DESCRIPTION: MElt is a freely available (LGPL) state-of-the-art sequence labeller that is meant to be trained on both an annotated corpus and an external lexicon. It was developed by Pascal Denis and Benoît Sagot within the Alpage team, a joint Inria and Université Paris-Diderot team in Paris, France. MElt allows for using multiclass Maximum-Entropy Markov models (MEMMs) or multiclass perceptrons (multitrons) as underlying statistical devices. Its output is in the Brown format (one sentence per line, each sentence being a space-separated sequence of annotated words in the word/tag format).

MElt has been trained on various annotated corpora, using Alexina lexicons as source of lexical information. As a result, models for French, English, Spanish and Italian are included in the MElt package.

MElt also includes a normalization wrapper aimed at helping processing noisy text, such as user-generated data retrieved on the web. This wrapper is only available for French and English. It was used for parsing web data for both English and French, respectively during the SANCL shared task (Google Web Bank) and for developing the French Social Media Bank (Facebook, twitter and blog data).

- Contact: Benoît Sagot
- URL: https://team.inria.fr/almanach/melt/

6.5. dyalog-sr

KEYWORDS: Parsing - Deep learning - Natural language processing

FUNCTIONAL DESCRIPTION: DyALog-SR is a transition-based dependency parser, built on top of DyALog system. Parsing relies on dynamic programming techniques to handle beams. Supervised learning exploit a perceptron and aggressive early updates. DyALog-SR can handle word lattice and produce dependency graphs (instead of basic trees). It was tested during several shared tasks (SPMRL'2013 and SEMEVAL'2014). It achieves very good accuracy on French TreeBank, alone or by coupling with FRMG parser. In 2017, DyALog-SR has been extended into DyALog-SRNN by adding deep neuronal layers implemented with the Dynet library. The new version has participated to the evaluation campaigns CONLL UD 2017 (on more than 50 languages) and EPE 2017.

- Contact: Éric De La Clergerie

6.6. FSMB

French Social Media Bank

KEYWORDS: Treebank - User-generated content

FUNCTIONAL DESCRIPTION: The French Social Media Bank is a treebank of French sentences coming from various social media sources (Twitter(c), Facebook(c)) and web forums (JeuxVidéos.com(c), Doctissimo.fr(c)). It contains different kind of linguistic annotations: - part-of-speech tags - surface syntactic representations (phrase-based representations) as well as normalized form whenever necessary.

- Contact: Djamé Seddah

6.7. DyALog

KEYWORD: Logic programming

FUNCTIONAL DESCRIPTION: DyALog provides an environment to compile and execute grammars and logic programs. It is essentially based on the notion of tabulation, i.e. of sharing computations by tabulating traces of them. DyALog is mainly used to build parsers for Natural Language Processing (NLP). It may nevertheless be used as a replacement for traditional PROLOG systems in the context of highly ambiguous applications where sub-computations can be shared.

- Participant: Eric de La Clergerie
- Contact: Eric de La Clergerie
- URL: http://dyalog.gforge.inria.fr/

6.8. SxPipe

KEYWORD: Surface text processing

SCIENTIFIC DESCRIPTION: Developed for French and for other languages, Sx Pipe includes, among others, various named entities recognition modules in raw text, a sentence segmenter and tokenizer, a spelling corrector and compound words recognizer, and an original context-free patterns recognizer, used by several specialized grammars (numbers, impersonal constructions, quotations...). It can now be augmented with modules developed during the former ANR EDyLex project for analysing unknown words, this involves in particular (i) new tools for the automatic pre-classification of unknown words (acronyms, loan words...) (ii) new morphological analysis tools, most notably automatic tools for constructional morphology (both derivational and compositional), following the results of dedicated corpus-based studies. New local grammars for detecting new types of entities and improvement of existing ones, developed in the context of the PACTE project, will soon be integrated within the standard configuration.

FUNCTIONAL DESCRIPTION: SxPipe is a modular and customizable processing chain dedicated to applying to raw corpora a cascade of surface processing steps (tokenisation, wordform detection, non-deterministic spelling correction. . .). It is used as a preliminary step before ALMAnaCH's parsers (e.g., FRMG) and for surface processing (named entities recognition, text normalization, unknown word extraction and processing...).

- Participants: Benoît Sagot, Djamé Seddah and Eric de La Clergerie
- Contact: Benoît Sagot
- URL: http://lingwb.gforge.inria.fr/

6.9. Mgwiki

KEYWORDS: Parsing - French

FUNCTIONAL DESCRIPTION: Mgwiki is a linguistic wiki that may used to discuss linguistic phenomena with the possibility to add annotated illustrative sentences. The work is essentially devoted to the construction of an instance for documenting and discussing FRMG, with the annotations of the sentences automatically provided by parsing them with FRMG. This instance also offers the possibility to parse small corpora with FRMG and an interface of visualization of the results. Large parsed corpora (like French Wikipedia or Wikisource) are also available. The parsed corpora can also be queried through the use of the DPath language.

- Participant: Eric de La Clergerie
- Contact: Eric de La Clergerie
- URL: http://alpage.inria.fr/frmgwiki/

6.10. WOLF

WOrdnet Libre du Français (Free French Wordnet)

KEYWORDS: WordNet - French - Semantic network - Lexical resource

FUNCTIONAL DESCRIPTION: The WOLF (Wordnet Libre du Français, Free French Wordnet) is a free semantic lexical resource (wordnet) for French.

The WOLF has been built from the Princeton WordNet (PWN) and various multilingual resources.

- Contact: Benoît Sagot
- URL: http://alpage.inria.fr/~sagot/wolf-en.html

6.11. vera

KEYWORD: Text mining

FUNCTIONAL DESCRIPTION: Automatic analysis of answers to open-ended questions based on NLP and statistical analysis and visualisation techniques (vera is currently restricted to employee surveys).

- Participants: Benoît Sagot and Dimitri Tcherniak
- Partner: Verbatim Analysis
- Contact: Benoît Sagot

6.12. Alexina

Atelier pour les LEXiques INformatiques et leur Acquisition

KEYWORD: Lexical resource

FUNCTIONAL DESCRIPTION: Alexina is ALMAnaCH's framework for the acquisition and modeling of morphological and syntactic lexical information. The first and most advanced lexical resource developed in this framework is the Lefff, a morphological and syntactic lexicon for French.

- Participant: Benoît Sagot
- Contact: Benoît Sagot
- URL: http://gforge.inria.fr/projects/alexina/

6.13. FQB

French QuestionBank

KEYWORD: Treebank

FUNCTIONAL DESCRIPTION: The French QuestionBanks is a corpus of around 2000 questions coming from various domains (TREC data set, French governmental organisation, NGOs, etc..) it contains different kind of annotations - morpho-syntactic ones (POS, lemmas) - surface syntaxe (phrase based and dependency structures) with long-distance dependency annotations.

The TREC part is aligned with the English QuestionBank (Judge et al, 2006).

- Contact: Djamé Seddah

6.14. Sequoia corpus

KEYWORD: Treebank

FUNCTIONAL DESCRIPTION: The Sequoia corpus contains French sentences, annotated with various linguistic information: - parts-of-speech - surface syntactic representations (both constituency trees and dependency trees) - deep syntactic representations (which are deep syntactic dependency graphs)

- Contact: Djamé Seddah

7. New Results

7.1. New results on text simplification

Participants: Benoît Sagot, Éric Villemonte de La Clergerie, Louis Martin.

Text simplification (TS) aims at making a text easier to read and understand by simplifying grammar and structure while keeping the underlying meaning and information identical. It is therefore an instance of language variation, based on language complexity. It can benefit numerous audiences, such as people with disabilities, language learners or even everyone, for instance when dealing with intrinsically complex texts such as legal documents.

We have initiated in 2017 a collaboration with the Facebook Artificial Intelligence Research (FAIR) lab in Paris and with the UNAPEI, the federation of French associations helping people with mental disabilities and their families. The objective of this collaboration is to develop tools for helping the simplification of texts aimed at mentally disabled people. More precisely, the is to develop a computer-assisted text simplification platform (as opposed to an automatic TS system). In this context, a CIFRE PhD thesis has started in collaboration with the FAIR on the TS task. We have first dedicated important efforts to the problem of the evaluation of TS systems, which remains an open challenge. As the task has common points with machine translation (MT), TS is often evaluated using MT metrics such as BLEU. However, such metrics require high quality reference data, which is rarely available for TS. TS has the advantage over MT of being a monolingual task, which allows for direct comparisons to be made between the simplified text and its original version. We compared multiple approaches to reference-less quality estimation of sentence-level TS systems, based on the dataset used for the QATS 2016 shared task. We distinguished three different dimensions: grammaticality, meaning preservation and simplicity. We have shown that n-gram-based MT metrics such as BLEU and METEOR correlate the most with human judgement of grammaticality and meaning preservation, whereas simplicity is best evaluated by basic length-based metrics [87]. Our implementations of several metrics have been made this year easily accessible and described in a demo paper in collaboration with the University of Sheffield [16].

In 2019, we have also investigated an important issue inherent to the TS task. Although it is often considered an all-purpose generic task where the same simplification is suitable for all, multiple audiences can benefit from simplified text in different ways. We have therefore introduced a discrete parametrisation mechanism that provides explicit control on TS systems based on Seq2Seq neural models. As a result, users can condition the simplifications returned by a model on parameters such as length and lexical complexity. We also show that carefully chosen values of these parameters allow out-of-the-box Seq2Seq neural models to outperform their standard counterparts on simplification benchmarks. Our best parametrised model improves over the previous state of the art performance [61].

Finally, we are involved in the development of a new text simplification corpus. In order to simplify a sentence, human editors perform multiple rewriting transformations: splitting it into several shorter sentences, paraphrasing (i.e. replacing complex words or phrases by simpler synonyms), reordering components, and/or deleting information deemed unnecessary. Despite the vast range of possible text alterations, current models for automatic sentence simplification are evaluated using datasets that are focused on single transformations, such as paraphrasing or splitting. This makes it impossible to understand the ability of simplification models in more abstractive and realistic settings. This is what motivated the development of ASSET, a new dataset for assessing sentence simplification in English, in collaboration with the University of Sheffield (United Kingdom). ASSET is a crowdsourced multi-reference corpus where each simplification was produced by executing several rewriting transformations. Through quantitative and qualitative experiments, we have shown that simplifications in ASSET are better at capturing characteristics of simplicity when compared to other standard evaluation datasets for the task. Furthermore, we have motivated the need for developing better methods for automatic evaluation using ASSET, since we show that current popular metrics may not be suitable for assessment when multiple simplification transformations were performed.

7.2. NLP and computational neurolinguistics

Participants: Éric Villemonte de La Clergerie, Murielle Fabre.

In the context of the CRCNS international network, the ANR-NSF NCM-ML project (dubbed "*Petit Prince* project") aims to discover and explore correlations between features (or predictors) provided by NLP tools such as parsers, and brain imagery (fMRI) data resulting from listening of the novel Le Petit Prince. Following the availability of an increasing amount of fMRI datasets in French and English, the project has investigated the correlations between fMRI observations and an increasing number of parser-based features based on several parsers representing a number of architecture types (LSTM, RNN, Dyalog-SR [statistical], FRMG [hybrid symbolic/statistical]) [20].

While pursuing the purely computation goal of developing a method of variable beam size inference for Recurrent Neural Network Grammar (RNNG) the project investigated how different beam search methods can show different goodness of fit with fMRI signal recorded during naturalistic story listening [58]. This approach is part of a new trend that is now emerging under the name of cognitively inspired NLP, where the effort to leverage from what we know of human cognition to increase machine processing of language data. Drawing inspiration from sequential Monte-Carlo methods such as particle filtering, we illustrated the relevance of our new method for speeding up the computations of direct generative parsing for RNNG, and revealing the potential cognitive interpretation of the underlying representations built by the search method and its beam activity through the analysis of neuro-imaging signal.

A second focus of the project is on compositionality, memory retrieval and syntactic composition during language comprehension. By using quantifications of these hypothesised processes as obtained from computational linguistics we seek to highlight their neural substrates and better understand or model human cognitition.

While linguistic expressions have been binarised as compositional and non-compositional given the lack of compositional linguistic analysis, the so-called Multi-word Expressions (MWEs) demonstrate finer-grained degrees of conventionalisation and predictability in psycho-linguisitcs, which can be quantified through computational Association Measures, like Point-wise Mutual Information and Dice's Coefficient [57]. An fMRI analysis was conducted to investigate to what extent these computational measures and the underlying cognitive processes they reflect are observable during on-line naturalistic sentence processing. Our results show that predictability, as quantified through Dice's Coefficent, is a better predictor of neural activation for processing MWEs and the more cognitively plausible computational metric. Computational results (1348) were obtained on MWE identification in French based on new method searching for frequent dependency-patterns [13]. These identifications in the Little Prince are contrasted with the ones published for English [69] and will yield an fMRI analysis comparing the two languages and the possible typological differences that the two languages may reflect in terms of morphological strategies to achieve lexical conventionalisation.

7.3. Large-scale raw corpus development

Participants: Benoît Sagot, Éric Villemonte de La Clergerie, Laurent Romary, Pedro Ortiz Suárez, Murielle Fabre, Louis Martin, Benjamin Muller, Yoann Dupont.

In order to be in phase (and comparable) with the US partners of the "Petit-Prince" ANR project, Murielle Fabre assembled two French corpora:

- a small corpus for domain adaptation to children's books: it will permit the fine tuning of the different parsers to a great amount of dialogues and Q&A present in *Le Petit Prince*.
- a large corpus of Contemporary French oral transcriptions and texts to calculate lexical association measures (AM) like PMI (Point-wise Mutual information) or Dice scores on the MWEs found in *Le Petit Prince*. This corpus of approx. 600 millions words, called CaBerNET, represents a balanced counterpart to the American COCA corpus. [0]

[0] https://corpus.byu.edu/coca/

We have also developed a general, highly parallel, multi-threaded pipeline to clean and classify Common Crawl by language. Common Crawl is a huge (over 20TB), heterogeneous multilingual corpus comprised of documents crawled from the internet, not sorted per language. We designed our pipeline, called goclassy, so that it runs efficiently on medium to low resource infrastructures where I/O speeds are the main constraint. We have created and we distribute a 6.3TB version of Common Crawl, called OSCAR, which is filtered, classified by language, shuffled at line level in order to avoid copyright issues, and ready to be used for NLP applications [29]. OSCAR corpora served as input data to train a variety of neural language models, including the French BERT model CamemBERT (see relevant module for more information). Bridging corpus development, NLP and computational neurolinguistics on of our next step is to train BERT model with the above cited French balanced corpus CaBerNet to create CaBERTnet and extract form it parsing metrics that will be correlated with brain activity as measured by French fMRI recording while listening *Le Petit Prince* in French.

7.4. Neural language modelling

Participants: Benoît Sagot, Djamé Seddah, Éric Villemonte de La Clergerie, Laurent Romary, Louis Martin, Benjamin Muller, Pedro Ortiz Suárez, Yoann Dupont, Ganesh Jawahar.

Pretrained language models are now ubiquitous in Natural Language Processing. Despite their success, most available models have either been trained on English data or on the concatenation of data in multiple languages. This makes practical use of such models—in all languages except English—very limited. In 2019, one of the most visible achievements of the ALMAnaCH team was the training and release of CamemBERT, a BERT-like [75] (rather, RoBERTa-like) neural language model for French trained on the French section of our large-scale web-based OSCAR corpus, together with CamemBERT variants [60]. Our goal was to investigate the feasibility of training monolingual Transformer-based language models for other languages, taking French as an example and evaluating our language models on part-of-speech tagging, dependency parsing, named entity recognition and natural language inference tasks. We have shown that the use of web-crawled data such as found in OSCAR to train such language models is preferable to the use of Wikipedia data, because of the homogeneity of Wikipedia data. More surprisingly, we have also shown that a relatively small web crawled dataset (4GB randomly extracted from the French section of OSCAR) leads to results that are as good as those obtained using larger datasets (130+GB, i.e. the whole French section of OSCAR). CamemBERT allowed us to reach or improve the state of the art in all four downstream tasks.

Beyond training neural language models, we have reinforced the exploration of an active question, that of their interpretability. With the emergence of contextual vector representations of words, such as the ELMo [89] and BERT language models and word embeddings, the interpretability of neural models becomes a key research topic. It is a way to understand what such neural networks actually learn in an unsupervised way from (huge amounts of) textual data, and in which circumstances they manage to do so. The work carried out in the team this year to identify where morphological vs. syntactic vs. semantic information is stored in a BERT language model [26] was part of a more general trend (see for example [78]). And our work on training ELMo models for five mid-resourced languages has shown that such LSTM-based models, when trained on large scale although non edited dataset such as our web-based corpora OSCAR, can lead to outperforming state-of-the-art performance on an number of downstream tasks such as part-of-speech tagging and parsing. Finally, we have carried out comparative evaluations of the performance of CamemBERT and of ELMo models trained on the same French section of OSCAR on a number of downstream task, with an emphasis on named-entity recognition—a work that led us to publish a new version of the named-entity-annotated version of the French TreeBank [67] that we published in 2012 [99].

We have also investigated how word embeddings can capture the evolution of word usage and meaning over time, at a fine-grained scale. As part of the ANR SoSweet and the PHC Maimonide projects (in collaboration with Bar Ilan University for the latter), ALMAnaCH has invested a lot of efforts since 2018 into studying language variation within user-generated content (UGC), taking into account two main interrelated dimensions: how language variation is related to socio-demographic and dynamic network variables, and how UGC language evolves over time. Taking advantage of the SoSweet corpus (600 millions tweet) and of the Bar Ilan Hebrew Tweets (180M tweets) both collected over the last 5 years, we have been addressing the problem of studying semantic changes via the use of dynamic word embeddings, that is embeddings evolving over time. We devised a novel attention model, based on Bernouilli word embeddings, that are conditioned on contextual extra-linguistic features such as network, spatial and socio-economic variables, which can be inferred from Twitter users metadata, as well as topic-based features. We posit that these social features provide an inductive bias that is susceptible to helping our model to overcome the narrow time-span regime problem. Our extensive experiments reveal that, as a result of being less biased towards frequency cues, our proposed model was able to capture subtle semantic shifts and therefore benefits from the inclusion of a reduced set of contextual features. Our model thus fit the data better than current state-of-the-art dynamic word embedding models and therefore is a promising tool to study diachronic semantic changes over small time periods. We published these ideas and results in [41].

A deep understanding of what is learned, and, beyond that, of how it is learned by neural language models, both synchronic and diachronic, will be a crucial step towards the improvement of such architectures (e.g. targeting low-resource languages or scenarios) and the design and deployment of new generations of neural networks for NLP. Particularly important is to assess the role of the training corpus size and heterogeneity, as well as the impact of the properties of the language at hand (e.g. morphological richness, token-type ratio, etc.). This line of research will also have an impact on our understanding of language variation and on our ability to improve the robustness of neural-network-based NLP tools to such variation.

7.5. Processing non-standard language: user-generated content and code-mixed language

Participants: Djamé Seddah, Benoît Sagot, Éric Villemonte de La Clergerie, Benjamin Muller, Ganesh Jawahar, Abhishek Srivastava, Jose Rosales Nuñez, Hafida Le Cloirec, Farah Essaidi, Matthieu Futeral.

In 2019, we have resumed our long-lasting efforts towards increasing the robustness of our language analysis tools to the variation found in user-generated content (UGC). We have done this in two directions, in the context of the SoSweet and Parsiti projects.

Firstly, we have investigated how our state-of-the-art hybrid (symbolic and statistical) parsing architecture for French, based on SxPipe, FRMG and the Lefff, behaves on French UGC data, namely on 20 millions tweets from the SoSweet corpus. A first observation was that the current level of pre-parsing normalization was not sufficient to ensure a good parsing coverage with FRMG (around 67%, to be compared with around 93% on journalistic texts such as the French TreeBank), also leading to high parsing times because of correction strategies. However, we applied our error mining strategy [6] to identify a first set of easy errors. Clustering and word embedding were also tried for lemmas relying on the dependency parse trees, again leading to semi-successful results due to the poor quality of the pre-parsing phases.

Secondly, we have investigated the normalisation task, whose goal is to transform possibly noisy UGC into less noisy inputs that are more adapted to our standard neural analysis models (e.g. taggers and parsers). More precisely, we have investigated how useful a language model such as BERT [75], trained on standard data, can be in handling non-canonical text. We study the ability of BERT to perform lexical normalisation in a realistic, and therefore low-resource, English UGC scenario [28]. By framing lexical normalisation as a token prediction task, by enhancing its architecture and by carefully fine-tuning it, we have shown that BERT can be a competitive lexical normalisation model without the need of any UGC resources aside from 3,000 training sentences. To the best of our knowledge, it is the first work done in adapting and analysing the ability of this model to handle noisy UGC data.

Thirdly, we have compared the performances achieved by Phrase-Based Statistical Machine Translation systems (PBSMT) and attention-based Neural Machine Translation systems (NMT) when translating UGC from French to English [44]. We have shown that, contrary to what could have been expected, PBSMT outperforms NMT when translating non-canonical inputs. Our error analysis uncovers the specificities of UGC that are problematic for sequential NMT architectures and suggests new avenue for improving NMT models.

Finally, building natural language processing systems for highly variable and low resource languages is a hard challenge. The recent success of large-scale multilingual pretrained neural language models (including our CamemBERT language model for French) provides us with new modeling tools to tackle it. We have studied the ability of the multilingual version of BERT to model an unseen dialect, namely the Latin-script user-generated North African Arabic dialect called Arabizi. We have shown in different scenarios that multilingual language models are able to transfer to such an unseen dialect, specifically in two extreme cases: across script (Arabic to Latin) and from Maltese, a related language written in the Arabic script, unseen during pretraining. Preliminary results have already been published [66].

7.6. Long-range diachronic variation

Participants: Benoît Sagot, Laurent Romary, Éric Villemonte de La Clergerie, Clémentine Fourrier, Gaël Guibon, Mathilde Regnault, Kim Gerdes.

ALMAnaCH members have resumed their work on longer-range diachronic variation, in two distinct directions:

- Firstly, we have been working on resources and tools for Old French, using contemporary French as a starting point for which resources and tools are available. This work is carried out within the ANR project "Profiterole", whose goal is to automatically annotate a large corpus of medieval French (9th-15th centuries) in dependency syntax and to provide a methodology for dealing with heterogeneous data as found in such a corpus. Indeed, Old French does not only involve diachronic variation when contrasted with contemporary French. It also involve large internal variation, notably because of diachronic (within Old French), dialectal, geographic, stylistic and genre-based variation. We have carried out experiments on morphosyntactic tagging by trying to determine which parameters and which training sets are the best ones to use when annotating a new text. We explored two approaches for parsing. On the one hand, an ongoing thesis aims at adapting the FRMG metagrammar to medieval French, notably by changing the constraints on certain syntactic phenomena and relaxing the order of words [31], [30]. This work relies on the new morphological and syntactic lexicon for Old French, OFrLex, developed at ALMAnaCH [34]. On the other hand, we conducted parsing experiments with neural models (DyALog's SRNN models).
- Secondly, we have started experiments to investigate whether and under which conditions neural networks can be used for learning sound correspondences between two related languages, i.e. for predicting cognates of source language words in a related target language. In order to obtain suitably large homogeneously phonetised data, we extracted bilingual lexicons and cognate sets from available resources, including our EtymDB etymological database, of which a new, extended version was created in 2019. This data was then used to train and evaluate several neural architectures (seq2seq, Siamese). Preliminary results are promising, but further investigation is required.

These two research directions will find a common ground now that we have begun to investigate, in the context of the Profiterole ANR project, how we can model the diachronic evolution of the lexicon from Old French to contemporary French. Moreover, our work on Basnage's 1701 *Dictionnaire Universel*, in the context of the BASNUM ANR project might draw some inspiration from the Profiterole project. But since 1700's French is much closer from contemporary French than Old French, another source of inspiration for BASNAGE might come from our work on sociolinguistic variation in contemporary French and more generally on our work on User-Generated Content (UCG).

7.7. Syntax and treebanking

Participants: Djamé Seddah, Benoît Sagot, Kim Gerdes, Benjamin Muller, Pedro Ortiz Suárez, Marine Courtin.

In 2019 we have introduced the first treebank for a romanized user-generated content of Algerian, a North-African Arabic dialect called Arabizi. It contains 1500 sentences, fully annotated in morpho-syntax and universal dependencies, and is freely available. We complement it with 50k unlabeled sentences that were collected using intensive data-mining techniques from Common Crawl and web-crawled data. Preliminary results show its usefulness for POS tagging and dependency parsing.

We have also developed the first syntactic treebank for spoken Naija, an English pidgincreole, which is rapidly spreading across Nigeria. The syntactic annotation is developed in the Surface-Syntactic Universal Dependency annotation scheme (SUD) [77] and automatically converted into Universal Dependencies (UD). A crucial step in the syntactic analysis of a spoken language consists in manually adding a markup onto the transcription, indicating the segmentation into major syntactic units and their internal structure. We have shown that this so-called "macrosyntactic" markup improves parsing results. We have also studied some iconic syntactic phenomena that clearly distinguish Naija from English. This work is published in [36].

We have carried out two pilot studies in empirical syntax based on UD treebanks. In a first study [38], we investigate the relationship between dependency distance and frequency based on the analysis of an English dependency treebank. The preliminary result shows that there is a non-linear relation between dependency distance and frequency. This relation between them can be further formalised as a power law function which can be used to predict the distribution of dependency distance in a treebank. In a second study [40], we discussed an empirical refoundation of selected Greenbergian word order univer-sals based on a data analysis of the Universal Dependencies project. The nature of the data we worked on allows us to extract rich details for testing well-known typological universals and constitutes therefore a valuable basis for validating Greenberg's universals. Our results show that we can refine some Greenbergian universals in a more empirical and accurate way by means of a data-driven typological analysis.

Finally, we have introduced a new schema to annotate Chinese Treebanks on the character level. The original UD and SUD projects provide token-level resources with rich morphosyntactic language details. However, without any commonly accepted word definition for Chinese, the dependency parsing always faces the dilemma of word segmentation. Therefore we have presented a character-level annotation schema integrated into the existing Universal Dependencies schema as an extension [39]. The different SUD projects were also presented at the Journées scientifiques "Linguistique informatique, formelle et de terrain" (LIFT 2019), Nov 28-29, 2019 at the University of Orléans.

7.8. Analysing and enriching legacy dictionaries

Participants: Laurent Romary, Benoît Sagot, Mohamed Khemakhem, Pedro Ortiz Suárez, Achraf Azhar.

2019 has been a year of deployment and large-scale experiment of the work initiated in 2016 on the analysis and enrichment of legacy dictionaries and implemented in the GROBID-dictionary framework [84]. GROBID-dictionary is an extension of the generic GROBID Suite [95] and implements an architecture of cascading CRF models with the purpose to parse and categorize components of a pdf documents, whether born-digital or resulting from an OCR. It is developped as part of the doctoral work of Mohamed Khemakhem. GROBID dictionaries produces an output that is conformant to the Text Encoding Initiative guideline and thus easy to distribute and further process in an open science context. We have had the opportunity the show the performances and robustness of the architecture on a variety of dictionaries and contexts resulting both from internal and external collaborations:

- In the context of the language documentation project of Jack Bowers dealing with Mixtepec-Mixtec (ISO 639-3: mix, [72], we have been successful in completely parsing a new edition of an historical lexical resource of Colonial Mixtec 'Voces del Dzaha Dzahui' published by the Dominican fray Francisco Alvarado in the year 1593, published by Jansen and Perez Jiménez (2009). The result is now integrated into the reference lexical description maintained by Jack). See [18];
- Within the Nénufar project, a collaboration with the Praxiling laboratory in Montpellier, we have been contributing to the analyses and encoding of several editions of the Petit Larousse Illustré, a central legacy publication for the French language. [17], [27];
- For the ANR funded project BASNUM, we are deeply involved in understanding how a complex, semi-structured dictionary, for which we do not necessarily have a high quality digitized primary source, can be properly segmented in lexical entries and subfields from which we expect being able to extract fine-grained linguistic content (e.g. named entities for literary sources). In [42], we have shown for instant how the GROBID-dictionary framework could be robust to variations in scanning and thus OCR quality;

- In the same context of the BASNUM project, we have also started to explore the possibility of deploying deep learning components. As shown in [43], the main challenges is the lack of available annotated data in order to train machine learning models, decreased accuracy when using modern pre-trained models due to the differences between present-day and 18th century French, and even unreliable or low quality OCRisation;
- These various experiments have been accompanied by an intense training and hand-on activity in the context in particular of the Lexical Data Master Class and collaboration within the ELEXIS project, which has opted for using the system for building a dictionary matrix from legacy dictionaries [0].Further alignments with the ongoing standardisation activities around TEI Lex0 and ISO 24613 (LMF) has been carried out to ensure a proper standards compliance of the generated output.
- Finally, and as a nice example of the kind of DH collaborations that our researches can lead to, we should mention here the targeted experiments that we carried out on extending the GROBID-dictionary framework to deal with objects which, although analogous with dictionary entries from a distance, appear to have a highly specific structure. This is the case Manuscript Sales Catalogues, which are highly important for authenticating documents and studying the reception of authors. Their regular publication throughout Europe since the beginning of the 19th c. has raised the interest around scaling up the means for automatically structuring their contents. [33] presents the results of advanced tests of the system's capacity to handle a large corpus with MSC of different dealers, and therefore multiple layouts.

7.9. Coreference resolution

Participants: Loïc Grobol, Éric Villemonte de La Clergerie.

In 2019 we have resumed our work on coreference resolution for French with the release in [25] of the first end-to-end automatic coreference resolution system for spoken French by adapting state-of-the art neural network system to the case of noisy non-standard inputs.

This first release uses no external knowledge beyond pretrained non-contextual word embeddings, making it suitable for applications to languages with less pre-existing resources. We also investigated the integration of further knowledge, both in the form of contextual embedding techniques such as CamemBERT and syntactic parsers developped at ALMAnaCH (works to be published in 2020).

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Ongoing contracts:

Verbatim Analysis Verbatim Analysis is an Inria start-up co-created in 2009 by Benoît Sagot. It uses some of ALMAnaCH's free NLP software (SxPipe) as well as a data mining solution co-developed by Benoît Sagot, VERA, for processing employee surveys with a focus on answers to open-ended questions.

opensquare was co-created in December 2016 by Benoît Sagot with 2 senior specialists of HR (human resources) consulting. It is dedicated to designing, carrying out and analysing employee surveys as well as HR consulting based on these results. It uses a new employee survey analysis tool, *enqi*, which is still under development. This tool being co-owned by opensquare and Inria, both parties have signed a Software Licence Agreement in exchange for a yearly fee paid by opensquare to ALMAnaCH based on its turnover. Benoît Sagot currently contributes to opensquare, under the "Concours scientifique" scheme.

[0] https://grobid.elex.is

Facebook A collaboration on text simplification ("français Facile À Lire et à Comprendre", FALC) is ongoing with Facebook's Parisian FAIR laboratory. It involves a co-supervised (CIFRE) PhD thesis in collaboration with UNAPEI, the largest French federation of associations defending and supporting people with special needs and their families. This collaboration, is part of a larger initiative called Cap'FALC involving (at least) these three partners as well as the relevant ministries. Funding received as a consequence of the CIFRE PhD thesis: 60,000 euros

Bluenove A contract with this company has been signed in 2018, which initiated a collaboration for the integration of NLP tools within Bluenove's platform Assembl, dedicated to online employee and citizen debating forums. It involved 12 months of fixed-term contracts (a post-doc, who worked at ALMAnaCH in 2018-2019). Funding received: 77,137 euros

Active collaborations without a contract:

Science Miner ALMAnaCH (following ALPAGE) has collaborated since 2014 with this company founded by Patrice Lopez, a specialist in machine learning techniques and initiator of the Grobid and NERD (now entity-fishing) suites. Patrice Lopez provides scientific support for the corresponding software components in the context of the Parthenos, EHRI and Iperion projects, as well as in the context of the Inria anHALytics initiative, aiming to provide a scholarly dashboard on scientific papers available from the HAL national publication repository.

Software Heritage , whose goal is to collect and preserve software in source code form. ALMAnaCH's collaboration with Software Heritage, on large-scale programming language identification, also involves Qwant, who provided some funding to Software Heritage.

Fortia Financial Solutions ALMAnaCH members led a proposal for the creation of an ANR LabCom with this French FinTech company on the analysis of (raw, PDF) financial documents from investment funds. The proposal was rejected, but future collaboration is still planned.

Hyperlex A collaboration was initiated in 2018 on NLP and information extraction from raw legal documents (mostly PDF format), involving especially Éric de La Clergerie, who is now a part-time employee of the company.

Ongoing discussions that should/could be formalised in the form of a contract in 2020:

Winespace : information extraction from wine descriptions to develop a wine recommendation system

Newsbridge : automatic extraction of short summaries of filmed events (e.g. sport events) based on social media coverage analysis

INPI : Patent classification

Cour de cassation (in the context of the LabIA): retrieval of relevant jurisprudence

DGCCRF (in the context of the LabIA): automatic identification of illicit clauses in B-to-C commercial contracts

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

- **ANR SoSweet** (2015-2019, PI J.-P. Magué, resp. ALMAnaCH: DS; Other partners: ICAR [ENS Lyon, CRNS], Dante [Inria]). Topic: studying sociolinguistic variability on Twitter, comparing linguistic and graph-based views on tweets
- **ANR ParSiTi** (2016-2021, PI Djamé Seddah, Other partners: LIMSI, LIPN). Topic: context-aware parsing and machine translation of user-generated content

- **ANR PARSE-ME** (2015-2020, PI. Matthieu Constant, resp. Marie Candito [ALPAGE, then LLF], ALMAnaCH members are associated with Paris-Diderot's LLF for this project). Topic: multi-word expressions in parsing
- **ANR Profiterole** (2016-2020, PI Sophie Prévost [LATTICE], resp. Benoit Crabbé [ALPAGE, then LLF], ALMAnaCH members are associated with Paris-Diderot's LLF for this project). Topic: modelling and analysis of Medieval French
- **ANR TIME-US** (2016-2019, PI Manuela Martini [LARHRA], ALMAnaCH members are associated with Paris-Diderot's CEDREF for this project). Topic: Digital study of remuneration and time budget textile trades in XVIIIth and XIXth century France
- **ANR BASNUM** (2018-2021, PI Geoffrey Williams [Université Grenoble Alpes], resp. ALMAnaCH: LR). Topic: Digitalisation and computational linguistic study of Basnage de Beauval's *Dictionnaire universel* published in 1701.

9.1.2. Competitivity Clusters and Thematic Institutes

- **PRAIRIE institute** (2019-2024, Dir.: Isabelle Ryl). Benoît Sagot was granted a Chair in this newly created research institute dedicated to Artificial Intelligence.
- **GDR LiFT** (2019-): LiFT is a CNRS-funded national coordination structure (GDR) involving many French teams involved in computational, formal and descriptive linguistics, in order to facilitate the emergence of fruitful collaborations. ALMAnaCH is involved in the GDR.
- **LabEx EFL** (2010-2019, PI Christian Puech [HTL, Paris 3], Sorbonne Paris Cité). Topic: empirical foundations of linguistics, including computational linguistics and natural language processing. ALPAGE was one of the partner teams of this LabEx, which gathers a dozen of teams within and around Paris whose research interests include one aspects of linguistics or more. BS serves as deputy head (and former head) of one of the scientific strands of the LabEx, namely strand 6 dedicated to language resources. BS and DS are in charge of a number of scientific "operations" within strands 6, 5 ("computational semantic analysis") and 2 ("experimental grammar"). BS, EVdLC and DS are now individual members of the LabEx EFL since 1st January 2017, and BS still serves as the deputy head of strand 6. Main collaborations are on language resource development (strands 5 and 6), syntactic and semantic parsing (strand 5, especially with LIPN [CNRS and U.Paris 13]) and computational morphology (strands 2 and 6, especially with CRLAO [CNRS and Inalco]).

9.1.3. Other National Initiatives

- **LECTAUREP project** (2017-2019): A preliminary study has been launched in collaboration with the National Archives in France, in the context of the framework agreement between Inria and the Ministry of Culture, to explore the possibility of extracting various components from digitised 19th Century notary registers.
- **Nénufar (DGLFLF - Délégation générale à la langue française et aux langues de France)**: The projects is intended to digitize and exploit the early editions (beginning of the 20th Century) of the Petit Larousse dictionary. ALMAnaCH is involve to contribute to the automatic extraction of the dictionary content by means of GROBID-Dictionaries and define a TEI compliant interchange format for all results.
- **PIA Opaline** (2017-2020): The objective of the project is to provide a better access to published French literature and reference material for visually impaired persons. Financed by the Programme d'Investissement d'Avenir, it will integrate technologies related to document analysis and re-publishing, textual content enrichment and dedicated presentational interfaces. Inria participate to deploy the GROBID tool suite for the automatic structuring of content from books available as plain PDF files.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

- **H2020 Parthenos** (2015-2019, PI Franco Niccolucci [University of Florence]; LR is a work package coordinator) Topic: strengthening the cohesion of research in the broad sector of Linguistic Studies, Humanities, Cultural Heritage, History, Archaeology and related fields through a thematic cluster of European Research Infrastructures, integrating initiatives, e-infrastructures and other world-class infrastructures, and building bridges between different, although tightly interrelated, fields.
- **H2020 EHRI** "European Holocaust Research Infrastructure" (2015-2019, PI Conny Kristel [NIOD-KNAW, NL]; LR is task leader) Topic: transform archival research on the Holocaust, by providing methods and tools to integrate and provide access to a wide variety of archival content.
- **H2020 Iperion CH** (2015-2019, PI Luca Pezzati [CNR, IT], LR is task leader) Topic: coordinating infrastructural activities in the cultural heritage domain.
- **H2020 HIRMEOS**: HIRMEOS objective is to improve five important publishing platforms for the open access monographs in the humanities and enhance their technical capacities and services and rendering technologies, while making their content interoperable. Inria is responsable for improving integrating the entity-fishing component deplyed as an infrastructural service for the five platforms.
- **H2020 DESIR**: The DESIR project aims at contributing to the sustainability of the DARIAH infrastructure along all its dimensions: dissemination, growth, technology, robustness, trust and education. Inria is responsable for providing of a portfolio of text analytics services based on GROBID and entity-fishing.

9.2.2. Collaborations in European Programs, Except FP7 & H2020

- **ERIC DARIAH "Digital Research Infrastructure for the Arts and Humanities"** (set up as a consortium of states, 2014-2034; LR served president of the board of director until August 2018) Topic: coordinating Digital Humanities infrastructure activities in Europe (17 partners, 5 associated partners).
- **COST enCollect** (2017-2020, PI Lionel Nicolas [European Academy of Bozen/Bolzano]) Topic: combining language learning and crowdsourcing for developing language teaching materials and more generic language resources for NLP

9.2.3. Collaborations with Major European Organizations

Collaborations with institutions not cited above (for the SPMRL initiative, see below):

- Berlin-Brandenburgische Akademie der Wissenschaften [Berlin-Brandenburg Academy of Sciences and Humanities], Berlin, Germany (Alexander Geyken) [lexicology]
- Österreichische Akademie der Wissenschaften [Austrian Academy of Sciences], Vienna, Austria (Karlheinz Moerth) [lexicology]
- Bar Ilan University (Yoav Goldberg, Hila Gonen) [non-canonical text processing]
- Dublin City University, Ireland (Teresa Lynn) [low-resource languages, user-generated content]
- University of Sheffield, United Kingdom (Lucia Specia, Carolina Scarton, Fernando Alva-Manchego) [text simplification]
- Univerza v Ljubljani [University of Ljubljana], Ljubljana, Slovenia (Darja Fišer) [wordnet development]

9.3. International Initiatives

9.3.1. Participation in Other International Programs

ANR-NSF project MCM-NL "Petit Prince" (2016-2020, PI John Hale [Cornell University, USA], resp. for Inria Paris/ALMAnaCH: Éric de La Clergerie) Topic: exploring correlations between data from neuro-imagery (fMRI, EEG) and data from NLP tools (mostly parsers). The data will come from "Le Petit Prince" read in French and English, and parsed with different parsers. Other partners: Cornell Univ., Univ. Michigan, Paris Saclay/Neurospin, Univ. Paris 8. Grant for ALMAnaCH: 108,500 euros

PHC Maïmonide (2018-2019, PI Djamé Seddah, co-PI Yoav Goldberg [Bar Ilan University]). Topics: Building NLP resources for analysing reactions to major events in Hebrew and French social media. Amount of the grant for the French side: 59,000 euros (89,000 euros for the whole project).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- BS organised with Annie Rialland and Catherine Schnedecker the 2019 "Journée de la Société de Linguistique de Paris", dedicated this year to "Corpus, analyses quantitatives et modèles linguistiques" (January 2019)
- DS was the co-chair of the SyntaxFest (Paris, 26-30 August 2019), a one-week event colocating a number of previously independent conferences and workshops on topics ranging from syntax to parsing and treebank development (including the UD workshop), as well as the co-chair of one of these conferences (Treebank and Linguistic Theories, TLT).
- KG was the co-chair of the International Conference on Dependency Linguistics, Depling 2019, another event of the SyntaxFest.

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program/Scientific/Reviewing Committees

- BS: Member of the Program, Scientific or Reviewing Committee of the following conferences and workshops: EMNLP-IJCNLP 2019, ACL 2019, NAACL 2019
- LR: Member of the Program, Scientific or Reviewing Committee of the following conferences and workshops: CMLC 2019, eLex 2019, ISA-15, TOTh 2019, LDK 2019, DATeCH 2019, ElPub 2019
- KG: Member of the Program, Scientific or Reviewing Committee of the following conferences and workshops: ACL 2019, SyntaxFest 2019 including Depling and Workshop on Universal Dependencies, Workshop on Multilingual Surface Realization, 2019, CoNLL 2019, EMNLP-IJCNLP 2019, International Conference on Natural Language Generation (INLG) 2019, LAW XIII 2019 (The 13th Linguistic Annotation Workshop), Rencontre des Jeunes Chercheurs en Parole

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- LR: Member of the scientific committe of the French speaking DH journal *Humanités numériques*
- KG: Member of the scientific committe of the *Journal of Linguistics*

10.1.3.2. Reviewer - Reviewing Activities

- BS: Reviewer for *Traitement Automatique des Langues*, *wék^{W}os*
- AC: Reviewer for *Digital Humanities 2020 - Intersections/Carrefours*

10.1.4. Invited Talks

- DS, invited talk, "Faire face au syndrome du Jabberwocky: Analyses morpho-syntaxiques en environnement hostile", at the Societé Linguistique de Paris (SLP) (Janvier 2019)
- BS, "Représentation et exploitation des informations lexicales", CENTAL, Louvain-la-Neuve, Belgium (April 2019)
- BS, "Morphological complexities", joint invited talk (with Géraldine Walther) at ACL 2019's SIGMORPHON workshop, Florence, Italy (July 2019) [15]

- LR, “The TEI as a modeling infrastructure: TEI beyond the TEI realms”. Ringvorlesung Digital Humanities, Paderborn, Germany (July 2019).
- LR, “Where is the place of scientific lexicography in science?”, Lorentz workshop: The Future of Academic Lexicography, Leiden, The Netherlands (Nov. 2019)
- LR, “Digital transition of SSH”, LERU Information and Open Access (INFO) Policy Group meeting, Strasbourg, France (Dec. 2019)
- EVdlC, “Acquisition de connaissances à partir de corpus analysés syntaxiquement” Seminar at Nokia Bell Labs, France (June 2019)
- EVdlC, “Évolutions et pertinence des plongements lexicaux”, workshop on “Machine learning, données textuelles et recherche en sciences humaines et sociales”, ENS Lyon, France (Nov. 2019)
- Benjamin Muller, “Transfer Learning on an unseen North African Arabic Dialect”, Bar Ilan University, Israel (24 Nov 2019)
- Mohamed Khemakhem was invited to the “ELEXIS Observers” event to give a talk [0] about structuring historical dictionaries and the use of GROBID-Dictionaries, Austrian Academy of Sciences, Austria (February 2019)
- KG, “Instant treebanks: Tools and methods to quickly build syntactically annotated corpora from scratch”, Workshop “Annotation of non-standard corpora (2019)”, University of Bamberg, Germany in September.

10.1.5. Training

- Mohamed Khemakhem chaired and tutored the GROBID-Dictionaries workshop series:
  - ELEXIS workshop - Berlin (March 2019)
  - Atelier de formation CollEx-Persée - Paris (September 2019)
- Following the aforementioned ‘Symposium on Naija’, a Master Class of one week was organised in June 2019 at the ARCIS institute of the University of Ibadan, Nigeria, on “Crowd-sourcing Web Corpora of Nigerian Languages” taught by KG and Slavomír Čéplö (from the Austrian Academy of Sciences).

10.1.6. Leadership within the Scientific Community

An important aspect of ALMAnaCH’s work relates to standardisation initiatives, especially within the ISO TC37 committee on language and terminology. Laurent Romary is the President of this committee and the convenor of its working group TC37/SC4 on lexical resources, within which he and a number of other ALMAnaCH members (Jack Bowers, Mohamed Khemakhem, Benoît Sagot, Éric de La Clergerie) have responsibilities (as project leaders or co-leaders). Most of them are related to the revision as a multi-part standard of ISO 24613 (Lexical Markup Framework, [14]), the first part of which was published in June 2019.

ALMAnaCH members have also played a key role in developing the Standardisation Survival Kit, an online tool hosted by Huma-num which focuses on giving researchers access to standards in a meaningful way by using research scenarios which cover all the domains of the Humanities, from literature to heritage science, including history, social sciences, linguistics, etc. We have published one publications on this topic in 2019 [64], following numerous ones over the last few years.

Other examples of ALMAnaCH’s leadership within the scientific community are the following:

- LR: Member of the ELEXIS Interoperability and Sustainability Committee (ISC) — ELEXIS is the European Lexicographic Infrastructure (https://elex.is)
- EVdLC: Chairman of the ACL special interest group SIGPARSE (ended in 2019)
- BS: Member, Deputy Treasurer and Member of the Board of the Société de Linguistique de Paris
- DS: Board member of the French NLP society (Atala, 2017-2020), Vice-President of the Atala and program chair of the “journée d’études”.
- DS: Member of the ACL’s BIG (Broad Interest Group) Diversity group.
- Mohamed Khemakhem: Member of the DARIAH Working Group “Bibliographical Data”

[0] http://videolectures.net/elexisobserver2019_khemakhem_dictionaries/

10.1.7. Scientific Expertise

- LR is an advisor for scientific information to Inria's deputy CEO for Science.

10.1.8. Research Administration

- BS is a member of Inria Paris's Scientific Committee ("Comité des Projets") and of its Board ("Bureau du Comité des Projets"), and a member of the International Relations Working Group of Inria's Scientific and Technological Orientation Council (COST-GTRI)
- BS is the Deputy Head (and former Head) of the research strand on Language Resources of the LabEx EFL (Empirical Foundations of Linguistics), and is therefore a deputy member of the Governing Board of the LabEx
- LR is the President of the scientific committee of ABES (Agence Bibliographique de l'Enseignement Supérieur), and a Member of the Text Encoding Initiative board

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Master: Benoît Sagot (with Emmanuel Dupoux), "Speech and Language Processing", 20h, M2, Master "Mathématiques, Vision, Apprentissage", ENS Paris-Saclay, France
- Master: Loïc Grobol, "Introduction à la fouille de textes", 39h, M1, Université Sorbonne Nouvelle, France
- Master: Loïc Grobol, "Langages de script", 39h, M2, INaLCO, France
- Eq. Master: Clémentine Fourrier, "Module TDLog - Techniques de Developpement Logiciel" (computer science opt-in course in Python), 18h, M1 and M2, École des Ponts ParisTech, France
- Eq. Bachelor second year: Gaël Guibon, "Introduction to Python and Programming" (for business school students and english speakers) 25h, International BBA 2, ESSEC Business School, France
- Master: Alix Chagué, "Humanités Numériques", 18h, M1, Institut d'études culturelles et internationales (IECI), Université Versailles-Saint-Quentin-en-Yvelines, France.
- Master: Alix Chagué, "Introduction à LaTeX", 5h, M1, Institut d'études culturelles et internationales (IECI), Université Versailles-Saint-Quentin-en-Yvelines, France.
- Master: Alix Chagué, "Introduction à Linux", 5h, M1, Institut d'études culturelles et internationales (IECI), Université Versailles-Saint-Quentin-en-Yvelines, France.
- Licence: Pedro Ortiz, "Mathématiques discrètes", 40h, L2, Licence de Sciences et Technologies, Faculté de Sciences et Ingénierie, Sorbonne Université, Paris, France.
- Licence: Mathilde Regnault, "Informatique et industries de la langue", 20h, L2, Licence Sciences du langage, Université Sorbonne nouvelle - Paris 3, Paris, France.

10.2.2. Supervision

PhD in progress: Mohamed Khemakhem, "Structuration automatique de dictionnaires à partir de modèles lexicaux standardisés", September 2016, Paris Diderot, supervised by Laurent Romary

PhD in progress: Loïc Grobol, "Coreference resolution for spoken French", "Université Sorbonne Nouvelle", started in Oct. 2016, supervised by Frédéric Landragin (main superviser), Isabelle Tellier† (main superviser), Éric de La Clergerie and Marco Dinarelli

PhD in progress: Jack Bowers, "Technology, description and theory in language documentation: creating a comprehensive body of multi-media resources for Mixtepec-Mixtec using standards, ontology and Cognitive Linguistics", started in Oct. 2016, EPHE, supervised by Laurent Romary

PhD in progress: Axel Herold, "Automatic identification and modeling of etymological information from retro-digitized dictionaries", October 2016, EPHE, Laurent Romary

PhD in progress: Mathilde Regnault,“Annotation et analyse de corpus hétérogènes”, “Université Sorbonne Nouvelle”, started in Oct. 2017, supervised by Sophie Prévost (main superviser), Isabelle Tellier†, and Éric de la Clergerie

PhD in progress: Louis Martin, “Text Simplification”, June 2018, Facebook & Sorbonne Université (“CIFRE” PhD), supervised by Benoît Sagot and Éric de La Clergerie

PhD in progress: Pedro Ortiz, “Automatic Enrichment of Ancient Dictionaries”, October 2018, Sorbonne Université, supervised by Laurent Romary and Benoît Sagot

PhD in progress: Benjamin Muller, “Multi-task learning for text normalisation, parsing and machine translation”, October 2018, Sorbonne Université, supervised by Benoît Sagot and Djamé Seddah

PhD in progress: José Carlos Rosales, supervised by Guillaume Wisnewski (Limsi) and Djamé Seddah, October 2018

PhD in progress: Clémentine Fourrier, “Neural approaches to the modelling of phonetic evolution”, October 2019, EPHE (Inria fellowship), Benoît Sagot

10.2.3. Juries

- BS: reviewer (“rapporteur”) of the PhD committee for Nourredine Alliane at Université Paris 8 on 17th May (Title: “Évaluation des représentations vectorielles de mots”; Supervisor: Gilles Bernard)
- BS: reviewer of the PhD committee for Tamara Álvarez López at Université de Vigo, Spain, on 29th November (Title: “Sentiment analysis in social media contents using natural language processing techniques”; Supervisors: Enrique Costa Montenegro and Milagros Fernández Gavilanes)
- BS: member of the PhD committee for Hazem Al Saied at Université de Lorraine on 20th December (Title: “Analyse automatique par transitions pour l’identification des expressions polylexicales”; Supervisors: Matthieu Constant and Marie Candito)
- LR: Reviewer of the PhD committee for Elina Leblanc, Université de Grenoble Alpes (title: “Bibliothèques numériques enrichies et participatives : utilisateurs, services, interfaces”, Supervisors: Elena Pierazzo and Hervé Blanchon)
- EVdlC: Reviewer of the PhD committee for Diana Nicoleta Popa, Université de Grenoble Alpes (title: “Vers des représentations contextualisées de mots”, Supervisor: Éric Gaussier)
- KG: reviewer of the PhD committee for Mohammed Galal at the University of Ain Shams, Egypt, on January 8 (Title: “Les constructions exceptives du français et de l’arabe : syntaxe et interface sémantique-syntaxe”; Supervisor: Sylvain Kahane in codirection with Dina El-Kassas)
- KG: President of the PhD committee for Marie-Amélie Botalla at the Sorbonne Nouvelle on May 14 (Title: “Modélisation de la production des énoncés averbaux : le cas des compléments différés”, Supervisors: Jeanne-Marie Debaisieux and Sylvain Kahane)

10.3. Popularization

10.3.1. Articles and contents

- BS was interviewed by David Larousserie, from the French daily newspaper Le Monde, about the release of CamemBERT, in October and December 2019. This resulted in an article in the newspaper, accessible here (subscription required).
- LR was interviewed the radio France Culture about the release of CamemBERT, in October 2019. The interview was aired on the radio.
- AC has taken in charge the content management of the Time Us project’s research blog since April 2019, which aims at presenting sources, good practices and results used and elaborated by all the project members.

10.3.2. Interventions

- Welcoming of schoolchildren at Inria Paris (half a day with ALMAnaCH members within an one-week-long stay; December 2019)
- Clémentine Fourrier co-organised the Young Women Mathematician Days (Inria Paris edition): 2 days of talks, mathematical animations and speed meetings with women scientists or engineers at the Inria Paris center, for high school girls
- Clémentine Fourrier was a speaker at the ESPCI highschool students summer school, to introduce the different jobs of computer science engineer to high school girls, in research and the industry, and animated an introductory software development workshop in Python over an afternoon
- Alix Chagué was the speaker for 6th workshop organised by the Alumni Association of the École nationale des chartes's Masters (ADEMEC), to introduce the software Transkribus and guidelines for text extraction (March 2019).

10.3.3. Internal action

- BS presented he January 2019 edition of the "demi-heure de science" at Inria Paris, whose target audience is researchers from all Inria Paris teams
- BS presented he June 2019 edition of Inria Paris' "my team in 180 seconds" event, whose target audience is both researchers and non-scientific Inria personnel
- CF presented softwares and good practices (Docker, GitLab and version control, Diverting Gitlab to create a small website) at the Inria developper meetup, whose target audience is both researchers and engineers.

11. Bibliography

Major publications by the team in recent years

[1] D. FIŠER, B. SAGOT.*Constructing a poor man's wordnet in a resource-rich world*, in "Language Resources and Evaluation", 2015, vol. 49, n^O 3, p. 601-635 [*DOI :* 10.1007/s10579-015-9295-6], https://hal.inria.fr/hal-01174492

[2] P. LOPEZ, L. ROMARY.*HUMB: Automatic Key Term Extraction from Scientific Articles in GROBID*, in "SemEval 2010 Workshop", Uppsala, Sweden, ACL SigLex event, July 2010, p. 248-251, https://hal.inria.fr/inria-00493437

[3] C. RIBEYRE, É. VILLEMONTE DE LA CLERGERIE, D. SEDDAH.*Because Syntax does Matter: Improving Predicate-Argument Structures Parsing Using Syntactic Features*, in "Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies", Denver, USA, United States, Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, June 2015, https://hal.archives-ouvertes.fr/hal-01174533

[4] L. ROMARY.*TEI and LMF crosswalks*, in "JLCL - Journal for Language Technology and Computational Linguistics", 2015, vol. 30, n^O 1, https://hal.inria.fr/hal-00762664

[5] B. SAGOT.*The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French*, in "7th international conference on Language Resources and Evaluation (LREC 2010)", Valletta, Malta, May 2010, https://hal.inria.fr/inria-00521242

[6] B. SAGOT, É. VILLEMONTE DE LA CLERGERIE.*Error mining in parsing results*, in "The 21st International Conference of the Association for Computational Linguistics (ACL 2006)", Sydney, Australia, July 2006, p. 329-336, https://hal.inria.fr/hal-02270412

[7] D. SEDDAH, B. SAGOT, M. CANDITO, V. MOUILLERON, V. COMBET.*The French Social Media Bank: a Treebank of Noisy User Generated Content*, in "COLING 2012 - 24th International Conference on Computational Linguistics", Mumbai, Inde, Kay, Martin and Boitet, Christian, December 2012, http://hal.inria.fr/hal-00780895

[8] R. TSARFATY, D. SEDDAH, Y. GOLDBERG, S. KÜBLER, Y. VERSLEY, M. CANDITO, J. FOSTER, I. REHBEIN, L. TOUNSI.*Statistical Parsing of Morphologically Rich Languages (SPMRL) What, How and Whither*, in "Proceedings of the NAACL HLT 2010 First Workshop on Statistical Parsing of Morphologically-Rich Languages", États-Unis Los Angeles, Association for Computational Linguistics, 2010, p. 1–12

[9] R. TSARFATY, D. SEDDAH, S. KÜBLER, J. NIVRE.*Parsing Morphologically Rich Languages: Introduction to the Special Issue*, in "Computational Linguistics", March 2013, vol. 39, n^o^ 1, 8 [*DOI :* 10.1162/COLI_A_00133], https://hal.inria.fr/hal-00780897

[10] É. VILLEMONTE DE LA CLERGERIE.*Improving a symbolic parser through partially supervised learning*, in "The 13th International Conference on Parsing Technologies (IWPT)", Naria, Japan, November 2013, https://hal.inria.fr/hal-00879358

Publications of the year

Articles in International Peer-Reviewed Journal

[11] D. REINEKE, L. ROMARY.*Bridging the gap between SKOS and TBX*, in "edition - Die Fachzeitschrift für Terminologie", November 2019, vol. 19, n^o^ 2, https://hal.inria.fr/hal-02398820

[12] L. ROMARY, C. RIONDET.*Towards multiscale archival digital data*, in "Umanistica digitale", 2019 [*DOI :* 10.6092/ISSN.2532-8816/9045], https://hal.inria.fr/hal-01586389

Invited Conferences

[13] M. FABRE, Y. DUPONT, É. VILLEMONTE DE LA CLERGERIE.*Syntactic Parsing versus MWEs: What can fMRI signal tell us*, in "PARSEME-FR 2019 consortium meeting", Blois, France, PARSEME-FR 2019, June 2019, https://hal.inria.fr/hal-02272288

[14] L. ROMARY, M. KHEMAKHEM, F. KHAN, J. BOWERS, N. CALZOLARI, M. GEORGE, M. PET, P. BAŃSKI.*LMF Reloaded*, in "AsiaLex 2019: Past, Present and Future", Istanbul, Turkey, June 2019, https://arxiv.org/abs/1906.02136 , https://hal.inria.fr/hal-02118319

[15] G. WALTHER, B. SAGOT.*Morphological complexities*, in "16th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology", Florence, Italy, August 2019, https://hal.inria.fr/hal-02266999

International Conferences with Proceedings

[16] F. ALVA-MANCHEGO, L. MARTIN, C. SCARTON, L. SPECIA.*EASSE: Easier Automatic Sentence Simplification Evaluation*, in "EMNLP-IJCNLP 2019 - Conference on Empirical Methods in Natural Language Pro-

cessing and 9th International Joint Conference on Natural Language Processing (demo session)", Hong Kong, China, November 2019, p. 49-54, https://hal.inria.fr/hal-02272950

[17] H. BOHBOT, F. FRONTINI, F. KHAN, M. KHEMAKHEM, L. ROMARY.*Nénufar: Modelling a Diachronic Collection of Dictionary Editions as a Computational Lexical Resource*, in "ELEX 2019: smart lexicography", Sintra, Portugal, October 2019, https://hal.inria.fr/hal-02272978

[18] J. BOWERS, M. KHEMAKHEM, L. ROMARY.*TEI Encoding of a Classical Mixtec Dictionary Using GROBID-Dictionaries*, in "ELEX 2019: Smart Lexicography", Sintra, Portugal, October 2019, https://hal.inria.fr/hal-02264033

[19] J. BOWERS, L. ROMARY.*TEI and the Mixtepec-Mixtec corpus: data integration, annotation and normalization of heterogeneous data for an under-resourced language*, in "6th International Conference on Language Documentation and Conservation (ICLDC)", Honolulu, United States, February 2019, https://hal.inria.fr/hal-02075475

[20] B. CRABBÉ, M. FABRE, C. PALLIER.*Variable beam search for generative neural parsing and its relevance for the analysis of neuro-imaging signal*, in "EMNLP-IJCNLP 2019 - Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing", Hong-Kong, China, November 2019, https://hal.inria.fr/hal-02272303

[21] M. DINARELLI, L. GROBOL.*Hybrid Neural Networks for Sequence Modelling : The Best of Three Worlds*, in "TALN-RECITAL 2019 - 26ème Conférence sur le Traitement Automatique des Langues Naturelles", Toulouse, France, Conférence sur le Traitement Automatique des Langues Naturelles (TALN-RECITAL), ATALA, July 2019, https://hal.archives-ouvertes.fr/hal-02157160

[22] M. DINARELLI, L. GROBOL.*Seq2Biseq: Bidirectional Output-wise Recurrent Neural Networks for Sequence Modelling*, in "CICLing 2019 - 20th International Conference on Computational Linguistics and Intelligent Text Processing", La Rochelle, France, April 2019, https://hal.inria.fr/hal-02085093

[23] L. FOPPIANO, L. ROMARY, M. ISHII, M. TANIFUJI.*Automatic Identification and Normalisation of Physical Measurements in Scientific Literature*, in "DocEng '19 - ACM Symposium on Document Engineering 2019", Berlin, Germany, ACM Press, September 2019, p. 1-4 [*DOI :* 10.1145/3342558.3345411], https://hal.inria.fr/hal-02294424

[24] S. GABAY, L. RONDEAU DU NOYER, M. KHEMAKHEM.*Selling autograph manuscripts in 19th c. Paris: digitising the Revue des Autographes*, in "IX Convegno AIUCD", Milan, Italy, AIUCD, January 2020, https://hal.archives-ouvertes.fr/hal-02388407

[25] L. GROBOL.*Neural Coreference Resolution with Limited Lexical Context and Explicit Mention Detection for Oral French*, in "Second Workshop on Computational Models of Reference, Anaphora and Coreference (CRAC19)", Minneapolis, United States, June 2019, https://hal.inria.fr/hal-02151569

[26] G. JAWAHAR, B. SAGOT, D. SEDDAH.*What does BERT learn about the structure of language?*, in "ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics", Florence, Italy, July 2019, https://hal.inria.fr/hal-02131630

[27] A. F. KHAN, H. BOHBOT, F. FRONTINI, M. KHEMAKHEM, L. ROMARY.*Historical Dictionaries as Digital Editions and Connected Graphs: the Example of Le Petit Larousse Illustré*, in "Digital Humanities 2019", Utrech, Netherlands, July 2019, https://hal.inria.fr/hal-02111199

[28] B. MULLER, B. SAGOT, D. SEDDAH.*Enhancing BERT for Lexical Normalization*, in "The 5th Workshop on Noisy User-generated Text (W-NUT)", Hong Kong, China, November 2019, https://hal.inria.fr/hal-02294316

[29] P. J. ORTIZ SUÁREZ, B. SAGOT, L. ROMARY.*Asynchronous Pipeline for Processing Huge Corpora on Medium to Low Resource Infrastructures*, in "7th Workshop on the Challenges in the Management of Large Corpora (CMLC-7)", Cardiff, United Kingdom, P. BAŃSKI, A. BARBARESI, H. BIBER, E. BREITENEDER, S. CLEMATIDE, M. KUPIETZ, H. LÜNGEN, C. ILIADI (editors), Leibniz-Institut für Deutsche Sprache, July 2019, https://hal.inria.fr/hal-02148693

[30] M. REGNAULT, S. PRÉVOST, É. VILLEMONTE DE LA CLERGERIE.*Challenges of language change and variation: towards an extended treebank of Medieval French*, in "TLT 2019 - 18th International Workshop on Treebanks and Linguistic Theories", Paris, France, August 2019, https://hal.inria.fr/hal-02272560

[31] M. REGNAULT.*Adapting a Metagrammar for Contemporary French to Medieval French*, in "TALN-RECITAL 2019 - 26e édition de la conférence TALN (Traitement Automatique des Langues Naturelles) et 21e édition de la conférence jeunes chercheur·euse·s RECITAL", Toulouse, France, July 2019, https://hal.inria.fr/hal-02147686

[32] L. ROMARY.*The place of lexicography in (computer) science*, in "The Future of Academic Lexicography: Linguistic Knowledge Codification in the Era of Big Data and AI", Leiden, Netherlands, Frieda Steurs and Dirk Geeraerts and Niels Schiller and Marian Klamer and Iztok Kosem, November 2019, https://hal.inria.fr/hal-02358218

[33] L. RONDEAU DU NOYER, S. GABAY, M. KHEMAKHEM, L. ROMARY.*Scaling up Automatic Structuring of Manuscript Sales Catalogues*, in "TEI 2019: What is text, really? TEI and beyond", Graz, Austria, September 2019, https://hal.inria.fr/hal-02272962

[34] B. SAGOT.*Development of a morphological and syntactic lexicon of Old French*, in "26ème Conférence sur le Traitement Automatique des Langues Naturelles (TALN)", Toulouse, France, July 2019, https://hal.inria.fr/hal-02148701

Conferences without Proceedings

[35] S. BASSETT, L. WESSELS, S. KRAUWER, B. MAEGAARD, H. HOLLANDER, F. ADMIRAAL, L. ROMARY, F. UITERWAAL.*Connecting the Humanities through Research Infrastructures*, in "4th Digital Humanities in the Nordic Countries (DHN 2019)", Copenhagen, Denmark, March 2019, https://hal.inria.fr/hal-02047512

[36] B. CARON, M. COURTIN, K. GERDES, S. KAHANE.*A Surface-Syntactic UD Treebank for Naija*, in "TLT 2019, Treebanks and Linguistic Theories, Syntaxfest", Paris, France, August 2019, https://hal.archives-ouvertes.fr/hal-02270530

[37] A. CHAGUÉ, V. LE FOURNER, M. MARTINI, É. VILLEMONTE DE LA CLERGERIE.*Deux siècles de sources disparates sur l'industrie textile en France : comment automatiser les traitements d'un corpus non-uniforme ?*, in "Colloque DHNord 2019 "Corpus et archives numériques"", Lille, France, MESHS Lille Nord de France, October 2019, https://hal.inria.fr/hal-02448921

[38] X. CHEN, K. GERDES.*The relation between dependency distance and frequency*, in "Quasy 2019, Quantitative Syntax 2019, Syntaxfest", Paris, France, August 2019, https://hal.archives-ouvertes.fr/hal-02270528

[39] C. DONG, Y. LI, K. GERDES.*Character-level Annotation for Chinese Surface-Syntactic Universal Dependencies*, in "Depling 2019 - International Conference on Dependency Linguistics", Paris, France, August 2019, https://hal.archives-ouvertes.fr/hal-02270535

[40] K. GERDES, S. KAHANE, X. CHEN.*Rediscovering Greenberg's Word Order Universals in UD*, in "UDW, Universal Dependencies Workshop 2019, Syntaxfest", Paris, France, August 2019, https://hal.archives-ouvertes.fr/hal-02270531

[41] G. JAWAHAR, D. SEDDAH.*Contextualized Diachronic Word Representations*, in "1st International Workshop on Computational Approaches to Historical Language Change 2019 (colocated with ACL 2019)", Florence, Italy, August 2019, https://hal.archives-ouvertes.fr/hal-02194763

[42] M. KHEMAKHEM, I. GALLERON, G. WILLIAMS, L. ROMARY, P. J. ORTIZ SUÁREZ.*How OCR Performance can Impact on the Automatic Extraction of Dictionary Content Structures*, in "19th annual Conference and Members' Meeting of the Text Encoding Initiative Consortium (TEI) - What is text, really? TEI and beyond", Graz, Austria, September 2019, https://hal.archives-ouvertes.fr/hal-02263276

[43] P. J. ORTIZ SUÁREZ, L. ROMARY, B. SAGOT.*Preparing the Dictionnaire Universel for Automatic Enrichment*, in "10th International Conference on Historical Lexicography and Lexicology (ICHLL)", Leeuwarden, Netherlands, June 2019, https://hal.inria.fr/hal-02131598

[44] J. C. ROSALES NUNEZ, D. SEDDAH, G. WISNIEWSKI.*A Comparison between NMT and PBSMT Performance for Translating Noisy User-Generated Content*, in "The 22nd Nordic Conference on Computational Linguistics (NoDaLiDa'19)", Turku, Finland, September 2019, https://hal.archives-ouvertes.fr/hal-02270524

Scientific Books (or Scientific Book chapters)

[45] *Proceedings of the Fifth International Conference on Dependency Linguistics (Depling, SyntaxFest 2019)*, August 2019, https://hal.inria.fr/hal-02450315

[46] PARTHENOS (editor). *Share - Publish - Store - Preserve. Methodologies, Tools and Challenges for 3D Use in Social Sciences and Humanities*, PARTHENOS and consortium 3D-SHS and LIA MAP-ISTI, Marseille, France, May 2019, https://hal.archives-ouvertes.fr/hal-02155055

[47] J. EDMOND, F. FISCHER, L. ROMARY, T. TASOVAC.*9. Springing the Floor for a Different Kind of Dance : Building DARIAH as a Twenty-First-Century Research Infrastructure for the Arts and Humanities*, in "Digital Technology and the Practices of Humanities Research", Open Book Publishers, February 2020, p. 207-234 [*DOI :* 10.11647/OBP.0192.09], https://hal.inria.fr/hal-02464622

[48] J. EDMOND, L. ROMARY.*3. Academic Publishing*, in "Digital Technology and the Practices of Humanities Research", Open Book Publishers, February 2020, p. 49-80 [*DOI :* 10.11647/OBP.0192.03], https://hal.inria.fr/hal-02464616

[49] K. GERDES, S. KAHANE, R. BAWDEN, J. BELIAO, É. VILLEMONTE DE LA CLERGERIE, I. WANG.*Annotation tools for syntax*, in "Rhapsodie: A Prosodic and Syntactic Treebank for Spoken French", June 2019, https://hal.inria.fr/hal-02450311

[50] S. Kahane, K. Gerdes, R. Bawden.*The microsyntactic annotation*, in "Rhapsodie: A Prosodic and Syntactic Treebank for Spoken French", June 2019, https://hal.inria.fr/hal-02450018

[51] S. Kahane, P. Pietrandrea, K. Gerdes.*The annotation of list structures*, in "Rhapsodie: A Prosodic and Syntactic Treebank for Spoken French", June 2019, https://hal.inria.fr/hal-02450034

[52] L. Romary, J. Edmond.*A Tangential View on Impact for the Arts and Humanities through the Lens of the DARIAH-ERIC*, in "Stay Tuned To The Future - Impact of the Research Infrastructures for Social Sciences and Humanities", B. Maegaard, R. Pozzo (editors), Leo S. Olschki Editore, 2019, https://hal.inria.fr/hal-02094713

Other Publications

[53] A. Bertino, L. Foppiano, L. Romary, P. Mounier.*Leveraging Concepts in Open Access Publications*, March 2019, working paper or preprint, https://hal.inria.fr/hal-01981922

[54] J. Bowers.*Language Documentation and Standards in Digital Humanities: TEI and the documentation of Mixtepec-Mixtec*, February 2019, working paper or preprint, https://hal.inria.fr/hal-02004005

[55] J. Bowers.*Pathways and patterns of metaphor & metonymy in Mixtepec-Mixtec body-part terms*, February 2020, working paper or preprint, https://hal.inria.fr/hal-02075731

[56] Y. Dupont.*Un corpus libre, évolutif et versionné en entités nommées du français*, July 2019, TALN 2019 - Traitement Automatique des Langues Naturelles, Poster, https://hal.archives-ouvertes.fr/hal-02448590

[57] M. Fabre, S. Bhattasali, C. Pallier, J. Hale.*Modeling Conventionalization and Predictability in Multi-Word Expressions at Brain-level*, September 2019, CRCNS 2019, Poster, https://hal.inria.fr/hal-02272435

[58] M. Fabre, B. Crabbe, C. Pallier.*Variable beam search for generative neural parsing and its fit with neuro-imaging signal*, September 2019, CRCNS 2019, Poster, https://hal.inria.fr/hal-02272475

[59] K. Gerdes, B. Guillaume, S. Kahane, G. Perrier.*Pourquoi se tourner vers le SUD : L'importance de choisir un schéma d'annotation en dépendance surface-syntaxique*, November 2019, LIFT 2019 - Journées scientifiques "Linguistique informatique, formelle & de terrain", https://hal.inria.fr/hal-02449922

[60] L. Martin, B. Muller, P. J. Ortiz Suárez, Y. Dupont, L. Romary, É. Villemonte de La Clergerie, D. Seddah, B. Sagot.*CamemBERT: a Tasty French Language Model*, October 2019, https://arxiv.org/abs/1911.03894 - Web site: https://camembert-model.fr, https://hal.inria.fr/hal-02445946

[61] L. Martin, B. Sagot, É. Villemonte de La Clergerie, A. Bordes.*Controllable Sentence Simplification*, October 2019, https://arxiv.org/abs/1910.02677 - Code and models: https://github.com/facebookresearch/access [*DOI :* 10.02677], https://hal.inria.fr/hal-02445874

[62] A. Málaga Sabogal, S. Troubetzkoy.*Unique ergodicity for infinite area Translation Surfaces*, August 2019, https://arxiv.org/abs/1908.04019 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02265283

[63] C. ROCHEREAU, B. SAGOT, E. DUPOUX.*Modeling German Verb Argument Structures: LSTMs vs. Humans*, December 2019, https://arxiv.org/abs/1912.00239 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02417640

[64] L. ROMARY, D. BIABIANY, K. ILLMAYER, M. PUREN, C. RIONDET, D. SEILLIER, L. TADJOU.*SSK by example - Make your Arts and Humanities research go standard*, May 2019, DARIAH Annual Event, Poster, https://hal.inria.fr/hal-02151788

[65] L. ROMARY.*The TEI as a modeling infrastructure: TEI beyond the TEI realms*, July 2019, Ringvorlesung Digital Humanities, https://hal.inria.fr/hal-02265036

[66] A. SRIVASTAVA, B. MULLER, D. SEDDAH.*Unsupervised Learning for Handling Code-Mixed Data: A Case Study on POS Tagging of North-African Arabizi Dialect*, October 2019, EurNLP - First annual EurNLP, Poster, https://hal.archives-ouvertes.fr/hal-02270527

References in notes

[67] A. ABEILLÉ, L. CLÉMENT, F. TOUSSENEL.*10*, in "Building a Treebank for French", Kluwer, Dordrecht, 2003, p. 165-187

[68] M. J. ARANZABE, A. D. DE ILARRAZA, I. GONZALEZ-DIOS.*Transforming complex sentences using dependency trees for automatic text simplification in Basque*, in "Procesamiento del lenguaje natural", 2013, vol. 50, p. 61–68

[69] S. BHATTASALI, M. FABRE, W.-M. LUH, H. AL SAIED, M. CONSTANT, C. PALLIER, J. R. BRENNAN, R. N. SPRENG, J. HALE.*Localising Memory Retrieval and Syntactic Composition: An fMRI Study of Naturalistic Language Comprehension*, in "Language, Cognition and Neuroscience", 2018, vol. 34, n^o 4, p. 1-20 [*DOI :* 10.1080/23273798.2018.1518533], https://hal.archives-ouvertes.fr/hal-01930201

[70] O. BONAMI, B. SAGOT.*Computational methods for descriptive and theoretical morphology: a brief introduction*, in "Morphology", 2017, vol. 27, n^o 4, p. 1-7 [*DOI :* 10.1017/CBO9781139248860], https://hal.inria.fr/hal-01628253

[71] A. BOUCHARD-CÔTÉ, D. HALL, T. GRIFFITHS, D. KLEIN.*Automated Reconstruction of Ancient Languages using Probabilistic Models of Sound Change*, in "Proceedings of the National Academy of Sciences", 2013, n^o 110, p. 4224–4229

[72] J. BOWERS, L. ROMARY.*Bridging the Gaps between Digital Humanities, Lexicography, and Linguistics: A TEI Dictionary for the Documentation of Mixtepec-Mixtec*, in "Dictionaries: Journal of the Dictionary Society of North America", 2018, vol. 39, n^o 2, p. 79-106, https://hal.inria.fr/hal-01968871

[73] J. C. K. CHEUNG, G. PENN.*Utilizing Extra-sentential Context for Parsing*, in "Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing", Cambridge, Massachusetts, EMNLP '10, 2010, p. 23–33

[74] M. CONSTANT, M. CANDITO, D. SEDDAH.*The LIGM-Alpage Architecture for the SPMRL 2013 Shared Task: Multiword Expression Analysis and Dependency Parsing*, in "Fourth Workshop on Statistical Parsing of Morphologically Rich Languages", Seattle, United States, October 2013, p. 46-52, https://hal.archives-ouvertes.fr/hal-00932372

[75] J. DEVLIN, M. CHANG, K. LEE, K. TOUTANOVA.*BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*, in "Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)", 2019, p. 4171–4186, https://www.aclweb.org/anthology/N19-1423/

[76] Y. FANG, M. CHANG.*Entity Linking on Microblogs with Spatial and Temporal Signals*, in "TACL", 2014, vol. 2, p. 259–272, https://tacl2013.cs.columbia.edu/ojs/index.php/tacl/article/view/323

[77] K. GERDES, B. GUILLAUME, S. KAHANE, G. PERRIER.*SUD or Surface-Syntactic Universal Dependencies: An annotation scheme near-isomorphic to UD*, in "Universal Dependencies Workshop 2018", Brussels, Belgium, November 2018, https://hal.inria.fr/hal-01930614

[78] J. HEWITT, C. D. MANNING.*A Structural Probe for Finding Syntax in Word Representations*, in "Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies", Association for Computational Linguistics, 2019, https://nlp.stanford.edu/pubs/hewitt2019structural.pdf

[79] J. E. HOARD, R. WOJCIK, K. HOLZHAUSER.*An automated grammar and style checker for writers of Simplified English*, in "Computers and Writing: State of the Art", 1992, p. 278–296

[80] D. HOVY, T. FORNACIARI.*Increasing In-Class Similarity by Retrofitting Embeddings with Demographic Information*, in "Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing", Association for Computational Linguistics, 2018, p. 671–677, http://aclweb.org/anthology/D18-1070

[81] D. HRUSCHKA, S. BRANFORD, E. SMITH, J. WILKINS, A. MEADE, M. PAGEL, T. BHATTACHARYA.*Detecting Regular Sound Changes in Linguistics as Events of Concerted Evolution*, in "Current Biology", 2015, vol. 1, n^{o} 25, p. 1–9

[82] G. JAWAHAR, B. MULLER, A. FETHI, L. MARTIN, É. VILLEMONTE DE LA CLERGERIE, B. SAGOT, D. SEDDAH.*ELMoLex: Connecting ELMo and Lexicon features for Dependency Parsing*, in "CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies", Brussels, Belgium, October 2018 [*DOI :* 10.18653/V1/K18-2023], https://hal.inria.fr/hal-01959045

[83] M. KHEMAKHEM, L. FOPPIANO, L. ROMARY.*Automatic Extraction of TEI Structures in Digitized Lexical Resources using Conditional Random Fields*, in "electronic lexicography, eLex 2017", Leiden, Netherlands, September 2017, https://hal.archives-ouvertes.fr/hal-01508868

[84] M. KHEMAKHEM, L. FOPPIANO, L. ROMARY.*Automatic Extraction of TEI Structures in Digitized Lexical Resources using Conditional Random Fields*, in "electronic lexicography, eLex 2017", Leiden, Netherlands, September 2017, https://hal.archives-ouvertes.fr/hal-01508868

[85] S. KÜBLER, M. SCHEUTZ, E. BAUCOM, R. ISRAEL.*Adding Context Information to Part Of Speech Tagging for Dialogues*, in "NEALT Proceedings Series", M. DICKINSON, K. MUURISEP, M. PASSAROTTI (editors), 2010, vol. 9, p. 115-126

[86] A.-L. LIGOZAT, C. GROUIN, A. GARCIA-FERNANDEZ, D. BERNHARD.*Approches à base de fréquences pour la simplification lexicale*, in "TALN-RÉCITAL 2013", 2013, 493

[87] L. MARTIN, S. HUMEAU, P.-E. MAZARÉ, A. BORDES, É. VILLEMONTE DE LA CLERGERIE, B. SAGOT.*Reference-less Quality Estimation of Text Simplification Systems*, in "1st Workshop on Automatic Text Adaptation (ATA)", Tilburg, Netherlands, November 2018, https://hal.inria.fr/hal-01959054

[88] H. MARTÍNEZ ALONSO, D. SEDDAH, B. SAGOT.*From Noisy Questions to Minecraft Texts: Annotation Challenges in Extreme Syntax Scenarios*, in "2nd Workshop on Noisy User-generated Text (W-NUT) at CoLing 2016", Osaka, Japan, December 2016, https://hal.inria.fr/hal-01584054

[89] M. E. PETERS, M. NEUMANN, M. IYYER, M. GARDNER, C. CLARK, K. LEE, L. ZETTLEMOYER.*Deep Contextualized Word Representations*, in "Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2018, New Orleans, Louisiana, USA, June 1-6, 2018, Volume 1 (Long Papers)", 2018, p. 2227–2237, https://www.aclweb.org/anthology/N18-1202/

[90] J. PYSSALO.*System PIE: the Primary Phoneme Inventory and Sound Law System for Proto-Indo-European*, University of Helsinki, 2013

[91] L. RELLO, R. BAEZA-YATES, S. BOTT, H. SAGGION.*Simplify or help?: text simplification strategies for people with dyslexia*, in "Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility", ACM, 2013, 15

[92] L. RELLO, R. BAEZA-YATES, L. DEMPERE-MARCO, H. SAGGION.*Frequent words improve readability and short words improve understandability for people with dyslexia*, in "IFIP Conference on Human-Computer Interaction", Springer, 2013, p. 203–219

[93] C. RIBEYRE, M. CANDITO, D. SEDDAH.*Semi-Automatic Deep Syntactic Annotations of the French Treebank*, in "The 13th International Workshop on Treebanks and Linguistic Theories (TLT13)", Tübingen, Germany, Proceedings of TLT 13, Tübingen Universität, December 2014, https://hal.inria.fr/hal-01089198

[94] L. ROMARY, M. KHEMAKHEM, F. KHAN, J. BOWERS, N. CALZOLARI, M. GEORGE, M. PET, P. BAŃSKI.*LMF Reloaded*, in "AsiaLex 2019: Past, Present and Future", Istanbul, Turkey, June 2019, https://hal.inria.fr/hal-02118319

[95] L. ROMARY, P. LOPEZ.*GROBID - Information Extraction from Scientific Publications*, in "ERCIM News", January 2015, vol. 100, https://hal.inria.fr/hal-01673305

[96] A. M. RUSH, R. REICHART, M. COLLINS, A. GLOBERSON.*Improved Parsing and POS Tagging Using Inter-sentence Consistency Constraints*, in "Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning", Jeju Island, Korea, EMNLP-CoNLL '12, 2012, p. 1434–1444

[97] B. SAGOT, H. MARTÍNEZ ALONSO.*Improving neural tagging with lexical information*, in "15th International Conference on Parsing Technologies", Pisa, Italy, September 2017, p. 25-31, https://hal.inria.fr/hal-01592055

[98] B. SAGOT, D. NOUVEL, V. MOUILLERON, M. BARANES.*Extension dynamique de lexiques morphologiques pour le français à partir d'un flux textuel*, in "TALN - Traitement Automatique du Langage Naturel", Les sables d'Olonne, France, June 2013, p. 407-420, https://hal.inria.fr/hal-00832078

[99] B. SAGOT, M. RICHARD, R. STERN.*Annotation référentielle du Corpus Arboré de Paris 7 en entités nommées*, in "Traitement Automatique des Langues Naturelles (TALN)", Grenoble, France, G. ANTONIADIS, H. BLANCHON, G. SÉRASSET (editors), Actes de la conférence conjointe JEP-TALN-RECITAL 2012, June 2012, vol. 2 - TALN, https://hal.inria.fr/hal-00703108

[100] B. SAGOT.*DeLex, a freely-avaible, large-scale and linguistically grounded morphological lexicon for German*, in "Language Resources and Evaluation Conference", Reykjavik, Iceland, European Language Resources Association, May 2014, https://hal.inria.fr/hal-01022288

[101] B. SAGOT.*External Lexical Information for Multilingual Part-of-Speech Tagging*, Inria Paris, June 2016, n^{o} RR-8924, https://hal.inria.fr/hal-01330301

[102] B. SAGOT.*Extracting an Etymological Database from Wiktionary*, in "Electronic Lexicography in the 21st century (eLex 2017)", Leiden, Netherlands, September 2017, p. 716-728, https://hal.inria.fr/hal-01592061

[103] C. SCARTON, M. DE OLIVEIRA, A. CANDIDO JR, C. GASPERIN, S. M. ALUÍSIO.*SIMPLIFICA: a tool for authoring simplified texts in Brazilian Portuguese guided by readability assessments*, in "Proceedings of the NAACL HLT 2010 Demonstration Session", Association for Computational Linguistics, 2010, p. 41–44

[104] Y. SCHERRER, B. SAGOT.*A language-independent and fully unsupervised approach to lexicon induction and part-of-speech tagging for closely related languages*, in "Language Resources and Evaluation Conference", Reykjavik, Iceland, European Language Resources Association, May 2014, https://hal.inria.fr/hal-01022298

[105] S. SCHUSTER, É. VILLEMONTE DE LA CLERGERIE, M. CANDITO, B. SAGOT, C. D. MANNING, D. SEDDAH.*Paris and Stanford at EPE 2017: Downstream Evaluation of Graph-based Dependency Representations*, in " EPE 2017 - The First Shared Task on Extrinsic Parser Evaluation", Pisa, Italy, Proceedings of the 2017 Shared Task on Extrinsic Parser Evaluation, September 2017, p. 47-59, https://hal.inria.fr/hal-01592051

[106] D. SEDDAH, M. CANDITO.*Hard Time Parsing Questions: Building a QuestionBank for French*, in "Tenth International Conference on Language Resources and Evaluation (LREC 2016)", Portorož, Slovenia, Proceedings of the 10th edition of the Language Resources and Evaluation Conference (LREC 2016), May 2016, https://hal.archives-ouvertes.fr/hal-01457184

[107] D. SEDDAH, B. SAGOT, M. CANDITO, V. MOUILLERON, V. COMBET.*The French Social Media Bank: a Treebank of Noisy User Generated Content*, in "COLING 2012 - 24th International Conference on Computational Linguistics", Mumbai, India, Kay, Martin and Boitet, Christian, December 2012, https://hal.inria.fr/hal-00780895

[108] D. SEDDAH, B. SAGOT, M. CANDITO.*The Alpage Architecture at the SANCL 2012 Shared Task: Robust Pre-Processing and Lexical Bridging for User-Generated Content Parsing*, in "SANCL 2012 - First Workshop on Syntactic Analysis of Non-Canonical Language , an NAACL-HLT'12 workshop", Montréal, Canada, June 2012, https://hal.inria.fr/hal-00703124

[109] M. SHARDLOW.*A survey of automated text simplification*, in "International Journal of Advanced Computer Science and Applications", 2014, vol. 4, n^{o} 1, p. 58–70

[110] A. SØGAARD, Y. GOLDBERG.*Deep multi-task learning with low level tasks supervised at lower layers*, in "Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)", Berlin, Germany, 2016, p. 231–235

[111] É. VILLEMONTE DE LA CLERGERIE, B. SAGOT, D. SEDDAH.*The ParisNLP entry at the ConLL UD Shared Task 2017: A Tale of a #ParsingTragedy*, in "Conference on Computational Natural Language Learning", Vancouver, Canada, Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, August 2017, p. 243-252 [*DOI :* 10.18653/v1/K17-3026], https://hal.inria.fr/hal-01584168

[112] É. VILLEMONTE DE LA CLERGERIE.*Jouer avec des analyseurs syntaxiques*, in "TALN 2014", Marseilles, France, ATALA, July 2014, https://hal.inria.fr/hal-01005477

[113] G. WALTHER, B. SAGOT.*Speeding up corpus development for linguistic research: language documentation and acquisition in Romansh Tuatschin*, in "Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature", Vancouver, Canada, Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, August 2017, p. 89-94 [*DOI :* 10.18653/v1/W17-2212], https://hal.inria.fr/hal-01570614

Project-Team ALPINES

Algorithms and parallel tools for integrated numerical simulations

IN COLLABORATION WITH: Laboratoire Jacques-Louis Lions (LJLL)

IN PARTNERSHIP WITH:
CNRS

Sorbonne Université (UPMC)

RESEARCH CENTER
Paris

THEME
Distributed and High Performance Computing

Table of contents

Project-Team ALPINES

Creation of the Team: 2013 January 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.4. - Multiscale modeling
A6.1.5. - Multiphysics modeling
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.5. - Numerical Linear Algebra
A6.2.7. - High performance computing
A6.3. - Computation-data interaction
A6.3.1. - Inverse problems
A7.1. - Algorithms

Other Research Topics and Application Domains:

B3.3.1. - Earth and subsoil
B9.5.2. - Mathematics
B9.5.3. - Physics

1. Team, Visitors, External Collaborators

Research Scientists

Laura Grigori [Team leader, Inria, Senior Researcher, HDR]
Sever Hirstoaga [Inria, Researcher, HDR]
Frédéric Nataf [CNRS, Senior Researcher, HDR]

Faculty Members

Xavier Claeys [Sorbonne Université, Associate Professor, HDR]
Frédéric Hecht [Sorbonne Université, Professor, HDR]

Technical Staff

Axel Fourmont [Inria, Engineer, from Mar 2019]
Jan Papez [Inria, Engineer, from Mar 2019 until Jun 2019]
Pierre-Henri Tournier [CNRS, Engineer]

PhD Students

Matthias Beaupere [Inria, PhD Student, from Sep 2019]
Igor Chollet [Sorbonne Université, PhD Student]
Thibault Cimic [Inria, PhD Student]
Zakariae Jorti [IFPEN, PhD Student, until Apr 2019]
Houssam Houssein [Sorbonne Université, PhD Student]
Pierre Marchand [Inria, PhD Student, until Sep 2019]
Van Thanh Nguyen [Inria, PhD Student]
Olivier Tissot [Inria, PhD Student, until Jan 2019]
Jeremy Weisman [Inria, PhD Student, from Oct 2019]

Post-Doctoral Fellows

Oleg Balabanov [Inria, Post-Doctoral Fellow, from Nov 2019]
Suraj Kumar [Inria, Post-Doctoral Fellow, from Nov 2019]

Jan Papez [Inria, Post-Doctoral Fellow, until Feb 2019]

Administrative Assistants

Laurence Bourcier [Inria, Administrative Assistant]
Maryse Desnous [Inria, Administrative Assistant, until Apr 2019]
Julien Guieu [Inria, Administrative Assistant, from Apr 2019]

2. Overall Objectives

2.1. Introduction

The focus of our research is on the development of novel parallel numerical algorithms and tools appropriate for state-of-the-art mathematical models used in complex scientific applications, and in particular numerical simulations. The proposed research program is by nature multi-disciplinary, interweaving aspects of applied mathematics, computer science, as well as those of several specific applications, as porous media flows, elasticity, wave propagation in multi-scale media and inverse problems.

Our first objective is to develop numerical methods and tools for complex scientific and industrial applications, that will enhance their scalable execution on the emergent heterogeneous hierarchical models of massively parallel machines. Our second objective is to integrate the novel numerical algorithms into a middle-layer that will hide as much as possible the complexity of massively parallel machines from the users of these machines.

3. Research Program

3.1. Overview

The research described here is directly relevant to several steps of the numerical simulation chain. Given a numerical simulation that was expressed as a set of differential equations, our research focuses on mesh generation methods for parallel computation, novel numerical algorithms for linear algebra, as well as algorithms and tools for their efficient and scalable implementation on high performance computers. The validation and the exploitation of the results is performed with collaborators from applications and is based on the usage of existing tools. In summary, the topics studied in our group are the following:

- Numerical methods and algorithms
  - Mesh generation for parallel computation
  - Solvers for numerical linear algebra
    - Domain decomposition methods
    - Preconditioning for iterative methods
  - Computational kernels for numerical linear algebra
  - Tensor computations
- Validation on numerical simulations and other numerical applications

3.2. Domain specific language - parallel FreeFem++

In the engineering, researchers, and teachers communities, there is a strong demand for simulation frameworks that are simple to install and use, efficient, sustainable, and that solve efficiently and accurately complex problems for which there are no dedicated tools or codes available. In our group we develop FreeFem++ (see https://www.freefem.org/), a user dedicated language for solving PDEs. The goal of FreeFem++ is not to be a substitute for complex numerical codes, but rather to provide an efficient and relatively generic tool for:

- getting a quick answer to a specific problem,
- prototyping the resolution of a new complex problem.

The current users of FreeFem++ are mathematicians, engineers, university professors, and students. In general for these users the installation of public libraries as MPI, MUMPS, Ipopt, Blas, lapack, OpenGL, fftw, scotch, PETSc, SLEPc is a very difficult problem. For this reason, the authors of FreeFem++ have created a user friendly language, and over years have enriched its capabilities and provided tools for compiling FreeFem++ such that the users do not need to have special knowledge of computer science. This leads to an important work on porting the software on different emerging architectures.

Today, the main components of parallel FreeFem++ are:

1. definition of a coarse grid,
2. splitting of the coarse grid,
3. mesh generation of all subdomains of the coarse grid, and construction of parallel data structures for vectors and sparse matrices from the mesh of the subdomain,
4. call to a linear solver,
5. analysis of the result.

All these components are parallel, except for point (5) which is not in the focus of our research. However for the moment, the parallel mesh generation algorithm is very simple and not sufficient, for example it addresses only polygonal geometries. Having a better parallel mesh generation algorithm is one of the goals of our project. In addition, in the current version of FreeFem++, the parallelism is not hidden from the user, it is done through direct calls to MPI. Our goal is also to hide all the MPI calls in the specific language part of FreeFem++. In addition to these in-house domain decomposition methods, FreeFem++ is also linked to PETSc solvers which enables an easy use of third parties parallel multigrid methods.

3.3. Solvers for numerical linear algebra

Iterative methods are widely used in industrial applications, and preconditioning is the most important research subject here. Our research considers domain decomposition methods and iterative methods and its goal is to develop solvers that are suitable for parallelism and that exploit the fact that the matrices are arising from the discretization of a system of PDEs on unstructured grids.

One of the main challenges that we address is the lack of robustness and scalability of existing methods as incomplete LU factorizations or Schwarz-based approaches, for which the number of iterations increases significantly with the problem size or with the number of processors. This is often due to the presence of several low frequency modes that hinder the convergence of the iterative method. To address this problem, we study different approaches for dealing with the low frequency modes as coarse space correction in domain decomposition or deflation techniques.

We also focus on developing boundary integral equation methods that would be adapted to the simulation of wave propagation in complex physical situations, and that would lend themselves to the use of parallel architectures. The final objective is to bring the state of the art on boundary integral equations closer to contemporary industrial needs. From this perspective, we investigate domain decomposition strategies in conjunction with boundary element method as well as acceleration techniques (H-matrices, FMM and the like) that would appear relevant in multi-material and/or multi-domain configurations. Our work on this topic also includes numerical implementation on large scale problems, which appears as a challenge due to the peculiarities of boundary integral equations.

3.4. Computational kernels for numerical linear and multilinear algebra

The design of new numerical methods that are robust and that have well proven convergence properties is one of the challenges addressed in Alpines. Another important challenge is the design of parallel algorithms for the novel numerical methods and the underlying building blocks from numerical linear algebra. The goal is to enable their efficient execution on a diverse set of node architectures and their scaling to emerging high-performance clusters with an increasing number of nodes.

Increased communication cost is one of the main challenges in high performance computing that we address in our research by investigating algorithms that minimize communication, as communication avoiding algorithms. We propose to integrate the minimization of communication into the algorithmic design of numerical linear algebra problems. This is different from previous approaches where the communication problem was addressed as a scheduling or as a tuning problem. The communication avoiding algorithmic design is an approach originally developed in our group since 2007 (initially in collaboration with researchers from UC Berkeley and CU Denver). While at mid term we focus on reducing communication in numerical linear algebra, at long term we aim at considering the communication problem one level higher, during the parallel mesh generation tool described earlier.

Our research also focuses on solving problems of large size that feature high dimensions as in molecular simulations. The data in this case is represented by objects called tensors, or multilinear arrays. The goal is to design novel tensor techniques to allow their effective compression, i.e. their representation by simpler objects in small dimensions, while controlling the loss of information. The algorithms are aiming to being highly parallel to allow to deal with the large number of dimensions and large data sets, while preserving the required information for obtaining the solution of the problem.

4. Application Domains

4.1. Compositional multiphase Darcy flow in heterogeneous porous media

We study the simulation of compositional multiphase flow in porous media with different types of applications, and we focus in particular on reservoir/bassin modeling, and geological CO2 underground storage. All these simulations are linearized using Newton approach, and at each time step and each Newton step, a linear system needs to be solved, which is the most expensive part of the simulation. This application leads to some of the difficult problems to be solved by iterative methods. This is because the linear systems arising in multiphase porous media flow simulations cumulate many difficulties. These systems are non-symmetric, involve several unknowns of different nature per grid cell, display strong or very strong heterogeneities and anisotropies, and change during the simulation. Many researchers focus on these simulations, and many innovative techniques for solving linear systems have been introduced while studying these simulations, as for example the nested factorization [Appleyard and Cheshire, 1983, SPE Symposium on Reservoir Simulation].

4.2. Inverse problems

We focus on methods related to the blend of time reversal techniques and absorbing boundary conditions (ABC) used in a non standard way. Since the seminal paper by [M. Fink et al., Imaging through inhomogeneous media using time reversal mirrors. Ultrasonic Imaging, 13(2):199, 1991.], time reversal is a subject of very active research. The principle is to back-propagate signals to the sources that emitted them. The initial experiment was to refocus, very precisely, a recorded signal after passing through a barrier consisting of randomly distributed metal rods. In [de Rosny and Fink. Overcoming the diffraction limit in wave physics using a time-reversal mirror and a novel acoustic sink. Phys. Rev. Lett., 89 (12), 2002], the source that created the signal is time reversed in order to have a perfect time reversal experiment. In [39], we improve this result from a numerical point of view by showing that it can be done numerically without knowing the source. This is done at the expense of not being able to recover the signal in the vicinity of the source. In [40], time dependent wave splitting is performed using ABC and time reversal techniques. We now work on extending these methods to non uniform media.

All our numerical simulations are performed in FreeFem++ which is very flexible. As a byproduct, it enables us to have an end user point of view with respect to FreeFem++ which is very useful for improving it.

4.3. Numerical methods for wave propagation in multi-scale media

We are interested in the development of fast numerical methods for the simulation of electromagnetic waves in multi-scale situations where the geometry of the medium of propagation may be described through caracteristic lengths that are, in some places, much smaller than the average wavelength. In this context, we propose to develop numerical algorithms that rely on simplified models obtained by means of asymptotic analysis applied to the problem under consideration.

Here we focus on situations involving boundary layers and *localized* singular perturbation problems where wave propagation takes place in media whose geometry or material caracteristics are submitted to a small scale perturbation localized around a point, or a surface, or a line, but not distributed over a volumic sub-region of the propagation medium. Although a huge literature is already available for the study of localized singular perturbations and boundary layer pheneomena, very few works have proposed efficient numerical methods that rely on asymptotic modeling. This is due to their functional framework that naturally involves singular functions, which are difficult to handle numerically. The aim of this part of our reasearch is to develop and analyze numerical methods for singular perturbation methods that are prone to high order numerical approximation, and robust with respect to the small parameter characterizing the singular perturbation.

4.4. Data analysis in astrophysics

We focus on computationally intensive numerical algorithms arising in the data analysis of current and forthcoming Cosmic Microwave Background (CMB) experiments in astrophysics. This application is studied in collaboration with researchers from University Paris Diderot, and the objective is to make available the algorithms to the astrophysics community, so that they can be used in large experiments.

In CMB data analysis, astrophysicists produce and analyze multi-frequency 2D images of the universe when it was 5% of its current age. The new generation of the CMB experiments observes the sky with thousands of detectors over many years, producing overwhelmingly large and complex data sets, which nearly double every year therefore following Moore's Law. Planck (http://planck.esa.int/) is a keystone satellite mission which has been developed under auspices of the European Space Agency (ESA). Planck has been surveying the sky since 2010, produces terabytes of data and requires 100 Petaflops per image analysis of the universe. It is predicted that future experiments will collect on the order of petabyte of data around 2025. This shows that data analysis in this area, as many other applications, will keep pushing the limit of available supercomputing power for the years to come.

4.5. Molecular simulations

Molecular simulation is one of the most dynamic areas of scientific computing. Its field of application is very broad, ranging from theoretical chemistry and drug design to materials science and nanotechnology. It provides many challenging problems to mathematicians and computer scientists.

In the context of the ERC Synergy Grant EMC2 we address several important limitations of state of the art molecular simulation. In particular, the simulation of very large molecular systems, or smaller systems in which electrons interact strongly with each other, remains out of reach today. In an interdisciplinary collaboration between chemists, mathematicians and computer scientists, we focus on developing a new generation of reliable molecular simulation algorithms and software.

5. Highlights of the Year

5.1. Highlights of the Year

- HPDDM software is now available through PETSc PCHPDDM routine, for details see the routine. Both Geneo and its multilevel extension presented in [32] are available.
- ERC Synergy EMC2 project kick off meeting took place in September 2019, at Sorbonne University.

6. New Software and Platforms

6.1. FreeFem++

FeeFrem++

SCIENTIFIC DESCRIPTION: FreeFem++ is a partial differential equation solver. It has its own language. freefem scripts can solve multiphysics non linear systems in 2D and 3D.

Problems involving PDE (2d, 3d) from several branches of physics such as fluid-structure interactions require interpolations of data on several meshes and their manipulation within one program. FreeFem++ includes a fast 2dˆ-tree-based interpolation algorithm and a language for the manipulation of data on multiple meshes (as a follow up of bamg (now a part of FreeFem++).

FreeFem++ is written in C++ and the FreeFem++ language is a C++ idiom. It runs on Macs, Windows, Unix machines. FreeFem++ replaces the older freefem and freefem+.

FUNCTIONAL DESCRIPTION: FreeFem++ is a PDE (partial differential equation) solver based on a flexible language that allows a large number of problems to be expressed (elasticity, fluids, etc) with different finite element approximations on different meshes.

- Partner: UPMC
- Contact: Frederic Hecht
- URL: http://www.freefem.org/ff++/

6.2. HPDDM

SCIENTIFIC DESCRIPTION: HPDDM is an efficient implementation of various domain decomposition methods (DDM) such as one- and two-level Restricted Additive Schwarz methods, the Finite Element Tearing and Interconnecting (FETI) method, and the Balancing Domain Decomposition (BDD) method. This code has been proven to be efficient for solving various elliptic problems such as scalar diffusion equations, the system of linear elasticity, but also frequency domain problems like the Helmholtz equation. A comparison with modern multigrid methods can be found in the thesis of Pierre Jolivet.

FUNCTIONAL DESCRIPTION: HPDDM is an efficient implementation of various domain decomposition methods (DDM) such as one- and two-level Restricted Additive Schwarz methods, the Finite Element Tearing and Interconnecting (FETI) method, and the Balancing Domain Decomposition (BDD) method.

- Participants: Frédéric Nataf and Pierre Jolivet
- Contact: Pierre Jolivet
- URL: https://github.com/hpddm

6.3. LORASC

LORASC preconditioner

KEYWORD: Preconditioner

- Participants: Laura Grigori and Rémi Lacroix
- Contact: Laura Grigori

6.4. Platforms

6.4.1. preAlps

KEYWORD: Preconditioned enlarged Krylov subspace methos

FUNCTIONAL DESCRIPTION: Contains enlarged Conjugate Gradient Krylov subspace method and Lorasc preconditioner.

- Partners: Inria
- Contact: Simplice Donfack, Laura Grigori, Olivier Tissot
- URL: https://github.com/NLAFET/preAlps

6.4.2. BemTool

KEYWORD: Boundary Element Method
FUNCTIONAL DESCRIPTION: BemTool is a C++ header-only library implementing the boundary element method for the discretisation of the Laplace, Helmholtz and Maxwell equations, in 2D and 3D. Its main purpose is the assembly of classic boundary element matrices, which can be compressed and inverted through its interface with HTOOL.

- Partners: UPMC - ANR NonlocalDD
- Contact: Xavier Claeys
- URL: https://github.com/xclaeys/BemTool

6.4.3. HTool

KEYWORD: Hierarchical Matrices
FUNCTIONAL DESCRIPTION: HTOOL is a C++ header-only library implementing compression techniques (e.g. Adaptive Cross Approximation) using hierarchical matrices. The library uses MPI and OpenMP for parallelism, and is interfaced with HPDDM for the solution of linear systems.

- Partners: CNRS - UPMC - ANR NonlocalDD
- Contact: Pierre Marchand
- URL: https://github.com/PierreMarchand20/htool

7. New Results

7.1. Adaptive Domain Decomposition Method for Saddle Point Problem

In [37], we introduce an adaptive domain decomposition (DD) method for solving saddle point problems defined as a block two by two matrix. The algorithm does not require any knowledge of the constrained space. We assume that all sub matrices are sparse and that the diagonal blocks are the sum of positive semi definite matrices. The latter assumption enables the design of adaptive coarse space for DD methods.

7.2. A Class of Efficient Locally Constructed Preconditioners Based on Coarse Spaces

In [24] we present a class of robust and fully algebraic two-level preconditioners for SPD matrices. We introduce the notion of algebraic local SPSD splitting of an SPD matrix and we give a characterization of this splitting. It helps construct *algebraically and locally* a class of efficient coarse subspaces which bound the spectral condition number of the preconditioned system by a number defined a priori. Some PDEs-dependant preconditioners correspond to a special case of the splitting. The examples of the algebraic coarse subspaces in this paper are not practical due to expensive construction. We propose an heuristic approximation that is not costly. Numerical experiments illustrate the efficiency of the proposed method.

7.3. A Multilevel Schwarz Preconditioner Based on a Hierarchy of Robust Coarse Spaces

In [32] we present a multilevel preconditioner for SPD matrices. Robust two-level additive Schwarz preconditioners guarantee a fast convergence of the Krylov method. To maintain the robustness each subdomain contributes a small number of vectors to construct a basis for the second level (the coarse space). As long as the dimension of the coarse space is reasonable i.e., direct solvers can be used efficiently, the two-level method scales well. However, the bottleneck arises when factoring the coarse space matrix becomes costly. Using an iterative Krylov method on the second level might be the right choice. Nevertheless, the condition number of the coarse space matrix is typically larger than the one of the first level. One of the difficulties of using two-level methods to solve the coarse problem is that the matrix does not arise from a PDE anymore. We introduce in this paper a practical method of applying a multilevel additive Schwarz preconditioner efficiently. This multilevel preconditioner is implemented in HPDDM and the code for reproducing the results from the paper is available here.

7.4. Inverse scattering problems without knowing the source term

The solution of inverse scattering problems always presupposed knowledge of the incident wave-field and require repeated computations of the forward problem, for which knowing the source term is crucial. In [26], we present a three-step strategy to solve inverse scattering problems when the time signature of the source is unknown. The proposed strategy combines three recent techniques: (i) wave splitting to retrieve the incident and the scattered wavefields, (ii) time-reversed absorbing conditions (TRAC) for redatuming the data inside the computational domain, (iii) adaptive eigenspace inversion (AEI) to solve the inverse problem. Numerical results illustrate step-by-step the feasibility of the proposed strategy.

7.5. Envelope following methods

One difficulty when solving problems in plasma physics is the behaviour at several scales in time and space of the solutions of equations. For example, central equations in this domain of application are highly oscillatory in time. The multiscale aspect makes the models difficult to tackle when we aim at avoiding a high computational cost. A solution to this problem is to solve the models by designing adapted numerical methods with a low computational cost and which are able to deal efficiently with rapid and slow scales in time. In this direction, we worked on envelope following methods, which have been efficiently applied in the community of oscillators in RF circuits. The method has (at least) two variants: in a first place, it is based on the concept of using extra variables to represent the changing rapid period and the cumulative effect of changing periods and then, use of Newton iterations allows to find these unknowns. In a second place, we adopt a similar strategy except that the rapid period is not an extra variable but a direct outcome of the numerical integration by the use of the Poincaré map. We implemented and tested both approaches for equations of interest in plasma physics and we observed that these methods didn't perform accurate results.

7.6. Domain decomposition preconditioning for high frequency wave propagation problems

The work about domain decomposition preconditioning for Maxwell equations has been published in [21]. It studies two-level preconditioners where the coarse space is based on the discretisation of the PDE on a coarse mesh. The PDE is discretised using finite-element methods of fixed, arbitrary order. The theoretical part of this work is the Maxwell analogue of a previous work for Helmholtz equation, and shows that for Maxwell problems with absorption, if the absorption is large enough and if the subdomain and coarse mesh diameters are chosen appropriately, then classical two-level overlapping Additive Schwarz Domain Decomposition preconditioning performs optimally – in the sense that GMRES converges in a wavenumber-independent number of iterations. The theory is also illustrated by various numerical experiments.

Ongoing studies are being conducted on recursive one-level optimized Schwarz methods for the high frequency Helmholtz and Maxwell equations. The method consists in solving the subdmain problems in a one-level optimized Schwarz preconditioner only approximately, using inner GMRES iterations preconditioned again by a one-level method, with smaller subdomains. The asymptotic behaviour and parallel scalability of the method are being investigated. Exhaustive numerical experiments are being conducted to compare the efficiency of this method with two-level preconditioners, including cavity problems and benchmarks in seismic imaging.

7.7. The boundary element method in FreeFEM

The BemTool and HTOOL libraries developed by the team, implementing respectively the Boundary Element Method and Hierarchical Matrices, have been interfaced with FreeFEM to allow FreeFEM users to use the Boundary Element Method (BEM) in their FreeFEM scripts. New additions to the Domain Specific Language (DSL) of FreeFEM allows the user to define and manipulate curved (1D) and surface (2D) meshes, as well as define and solve BEM variational problems in a high-level manner, similarly to FEM problems. The parallelization of the HTOOL library allows the user to assemble and solve their BEM problems in parallel in a transparent way.
Ongoing work consists in finalizing the BEM DSL to propose complete and documented features to the FreeFEM user in the next release, as well as investigating FEM-BEM coupling.

7.8. New Optimised Schwarz Method for dealing with cross-points

We consider a scalar wave propagation in harmonic regime modelled by Helmholtz equation with heterogeneous coefficients. Using the Multi-Trace Formalism (MTF), we propose a new variant of the Optimized Schwarz Method (OSM) that can accomodate the presence of cross-points in the subdomain partition. This leads to the derivation of a strongly coercive formulation of our Helmholtz problem posed on the union of all interfaces. The corresponding operator takes the form "identity + contraction".

7.9. Two-level preconditioning for h-version boundary element approximation of hypersingular operator with GenEO

We consider symmetric positive definite operators stemming from boundary integral equation (BIE), and we analysed a two-level preconditioner where the coarse space is built using local generalized eigenproblems in the overlap. We will refer to this coarse space as the GenEO coarse space. We obtained bounds on the condition number of the preconditionned system. In this work package, we also performed large scale numerical experiments for testing the scalability of our approach. We relied on parallel implementation of our algorithm.

7.10. Adaptive resolution of linear systems based on a posteriori error estimators

In [18] we discuss a new adaptive approach for iterative solution of sparse linear systems arising from partial differential equations (PDEs) with self-adjoint operators. The idea is to use the a posteriori estimated local distribution of the algebraic error in order to steer and guide the solve process in such way that the algebraic error is reduced more efficiently in the consecutive iterations. We first explain the motivation behind the proposed procedure and show that it can be equivalently formulated as constructing a special combination of preconditioner and initial guess for the original system. We present several numerical experiments in order to identify when the adaptive procedure can be of practical use.

7.11. Adaptive hierarchical subtensor partitioning for tensor compression

In [33] a numerical method is proposed to compress a tensor by constructing a piece-wise tensor approximation. This is defined by partitioning a tensor into sub-tensors and by computing a low-rank tensor approximation (in a given format) in each sub-tensor. Neither the partition nor the ranks are fixed a priori, but, instead, are obtained in order to fulfill a prescribed accuracy and optimize, to some extent, the storage. The different steps of the method are detailed and some numerical experiments are proposed to assess its performances.

7.12. Frictionless contact problem for hyper-elastic materials with interior point optimizer

In [35] we present a method to solve the mechanical problems undergoing finite deformations and the unilateral contact problems without friction for hyperelastic materials. We apply it to an industrial application: contact between a mechanical gasket and an obstacle. The main idea is to formulate the contact problem into an optimization one, in or- der to use the Interior Point OPTimizer (IPOPT) to solve it. Finally, the FreeFEM software is used to compute and solve the contact problem. Our method is validated against several benchmarks and used on an industrial application example.

7.13. A posteriori error estimates for Darcy's problem coupled with the heat equation

In [25] we derive a posteriori error estimates, in two and three dimensions, for the heat equation coupled with Darcy's law by a nonlinear viscosity depending on the temperature. We introduce two variational formulations and discretize them by finite element methods. We prove optimal a posteriori errors with two types of computable error indicators. The first one is linked to the linearization and the second one to the discretization. Then we prove upper and lower error bounds under regularity assumptions on the solutions. Finally, numerical computations are performed to show the effectiveness of the error indicators.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Contract with IFPen, February 2016 - April 2019, that funds the Phd thesis of Zakariae Jorti on adaptive preconditioners using a posteriori error estimators. Supervisor L. Grigori.
- Contract with IFPen, October 2016 - October 2019, that funds the Phd thesis of Julien Coulet on the virtual element method (VEM). Supervisor F. Nataf and V. Girault.

9. Partnerships and Cooperations

9.1. Regional Initiatives

GIS, Géosciences franciliennes: scientific collaboration network between ten public institutions from the Paris (Ile-de-France) region, focused on natural resources and environment. The project-team Alpines is a member.

9.2. National Initiatives

9.2.1. ANR

9.2.1.1. B3DCMB

ANR Decembre 2017 - Novembre 2021 This project is in the area of data analysis of cosmological data sets as collected by contemporary and forthcoming observatories. This is one of the most dynamic areas of modern cosmology. Our special target are data sets of Cosmic Microwave Background (CMB) anisotropies, measurements of which have been one of the most fruitful of cosmological probes. CMB photons are remnants of the very early evolution of the Universe and carry information about its physical state at the time when the Universe was much younger, hotter and denser, and simpler to model mathematically. The CMB has been, and continue to be, a unique source of information for modern cosmology and fundamental physics. The main objective of this project is to empower the CMB data analysis with novel high performance tools and algorithms superior to those available today and which are capable of overcoming the existing performance gap. Partners: AstroParticules et Cosmologie Paris 7 (PI R. Stompor), ENSAE Paris Saclay.

9.2.1.2. ANR Cine-Para

October 2015 - September 2019, Laura Grigori is Principal Coordinator for Inria Paris. Funding for Inria Paris is 145 Keuros. The funding for Inria is to combine Krylov subspace methods with parallel in time methods. Partners: University Pierre and Marie Curie, J. L. Lions Laboratory (PI Y. Maday), CEA, Paris Dauphine University, Paris 13 University.

9.2.1.3. Non-local DD

ANR appel à projet générique October 2015 - September 2020

This project in scientific computing aims at developing new domain decomposition methods for massively parallel simulation of electromagnetic waves in harmonic regime. The specificity of the approach that we propose lies in the use of integral operators not only for solutions local to each subdomain, but for coupling subdomains as well. The novelty of this project consists, on the one hand, in exploiting multi-trace formalism for domain decomposition and, on the other hand, considering optimized Schwarz methods relying on Robin type transmission conditions involving quasi-local integral operators.

9.2.1.4. Soilμ-3D

ANR appel à projet générique October 2015 - april 2019

In spite of decades of work on the modeling of greenhouse gas emission such as CO2 and N2O and on the feedback effects of temperature and water content on soil carbon and nitrogen transformations, there is no agreement on how these processes should be described, and models are widely conflicting in their predictions. Models need improvements to obtain more accurate and robust predictions, especially in the context of climate change, which will affect soil moisture regime.

The goal of this new project is now to go further using the models developed in MEPSOM to upscale heterogeneities identified at the scale of microbial habitats and to produce macroscopic factors for biogeochemical models running at the field scale.

To achieve this aim, it will be necessary to work at different scales: the micro-scale of pores (μm) where the microbial habitats are localized, the meso-scale of cores at which laboratory measurements on CO2 and N2O fluxes can be performed, and the macro-scale of the soil profile at which outputs are expected to predict greenhouse gas emission. The aims of the project are to (i) develop new descriptors of the micro-scale 3D soil architecture that explain the fluxes measured at the macro-scale, (ii) Improve the performance of our 3D pore scale models to simulate both micro-and meso- scales at the same time. Upscaling methods like “homogeneization” would help to simulate centimeter samples which cannot be achieved now. The reduction of the computational time used to solve the diffusion equations and increase the number of computational units, (iii) develop new macro-functions describing the soil micro-heterogeneity and integrate these features into the field scale models.

9.2.1.5. Muffin

ANR appel à projet générique 2019.

S. Hirstoaga and P.-H. Tournier are members of the project MUFFIN, whose objective is to explore and optimize original computational scenarios for multi-scale and high dimensional transport codes, with priority applications in plasma physics. Several approximation methods are planed to be developed. It is at the frontier of computing and numerical analysis and intends to reduce the computational burden in the context of intensive calculation. Principal Investigator: B. Després (Sorbonne University).

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. NLAFET (197)

Title: Parallel Numerical Linear Algebra for Future Extreme-Scale Systems

Programm: H2020

Duration: November 2015 - April 2019

Coordinator: UMEÅ Universitet

Partners:

Science and Technology Facilities Council (United Kingdom)

Computer Science Department, UmeåUniversitet (Sweden)

Mathematics Department, The University of Manchester (United Kingdom)

Inria, Alpines group

Inria contact: Laura Grigori

The NLAFET proposal is a direct response to the demands for new mathematical and algorithmic approaches for applications on extreme scale systems, as identified in the FETHPC work programme and call. This project will enable a radical improvement in the performance and scalability of a wide range of real-world applications relying on linear algebra software, by developing novel architecture-aware algorithms and software libraries, and the supporting runtime capabilities to achieve scalable performance and resilience on heterogeneous architectures. The focus is on a critical set of fundamental linear algebra operations including direct and iterative solvers for dense and sparse linear systems of equations and eigenvalue problems. Achieving this requires a co-design effort due to the characteristics and overwhelming complexity and immense scale of such systems. Recognized experts in algorithm design and theory, parallelism, and auto-tuning will work together to explore and negotiate the necessary tradeoffs. The main research objectives are: (i) development of novel algorithms that expose as much parallelism as possible, exploit heterogeneity, avoid communication bottlenecks, respond to escalating fault rates, and help meet emerging power constraints; (ii) exploration of advanced scheduling strategies and runtime systems focusing on the extreme scale and strong scalability in multi/many-core and hybrid environments; (iii) design and evaluation of novel strategies and software support for both offline and online auto-tuning. The validation and dissemination of results will be done by integrating new software solutions into challenging scientific applications in materials science, power systems, study of energy solutions, and data analysis in astrophysics. The deliverables also include a sustainable set of methods and tools for cross-cutting issues such as scheduling, auto-tuning, and algorithm-based fault tolerance packaged into open-source library modules.

9.3.1.2. ERC Synergy grant EMC2

Title: Extreme-scale Mathematically-based Computational Chemistry (EMC2)

Programm: ERC

Duration: September 2019 - August 2025

PIs: E. Cances (ENPC), L. Grigori (Inria), Y. Maday (Sorbonne University), J. P. Piquemal (Sorbonne University)

Molecular simulation is one of the most dynamic areas of scientific computing. Its field of application is very broad, ranging from theoretical chemistry and drug design to materials science and nanotechnology. Its importance in modern science has been acknowledged by two Nobel Prizes (Kohn & Pople in 1998; Karplus, Levitt & Warshel in 2013). It is also a gold mine of exciting problems for mathematicians and computer scientists.

Molecular simulation can be used as a virtual microscope to study more or less complex molecules with atomic-scale space-time resolution. It can also be used as a tool for computer-aided design (CAD) and the engineering of new molecules, materials and nano-devices.

However, molecular simulation still has important limitations. In particular, the simulation of very large molecular systems, or smaller systems in which electrons interact strongly with each other, remains out of reach today. Overcoming these limitations is extremely difficult. This requires joint breakthroughs in several disciplines, and can, in our opinion, only be achieved through an intensive multidisciplinary effort such as those made possible by ERC-Synergy-type funding.

Our objective is to overcome some of the current limitations in this field and to provide academic communities and industrial companies with new generation, dramatically faster and quantitatively reliable molecular simulation software, to enable those communities to address major technological and societal challenges of the 21st century (in health, energy, and the environment, for example).

9.4. International Initiatives

9.4.1. Inria International Partners

9.4.1.1. Informal International Partners

- J. Demmel, UC Berkeley, USA
- M. Grote, Université de Bâle, Suisse
- F. Assous, Israel
- K.-M. Perfekt, Reading, UK
- T. Chacon, Seville, Spain

9.5. International Research Visitors

9.5.1. Visits to International Teams

9.5.1.1. Research Stays Abroad

- Visit of Laura Grigori to the group of J. Demmel at U.C. Berkeley, july-august 2019.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- Organization of the kick-off meeting of ERC Synergy project EMC2 by E. Cances, L. Grigori, Y. Maday, and J. P. Piquemal. For more details see the website.
- Organization of Horizon Maths 2019 day on Mathematics and Chemistry by E. Cances, L. Grigori, Y. Maday, and J. P. Piquemal, december 3 2019. For more details see the website.

10.1.2. Scientific Events: Selection

- Gene Golub Summer school on High Performance Data Analytics, June 17-28, 2019 in Aussois, France. Co-organized by L. Grigori with M. Knepley, O. Schenk, R. Vuduc. Sponsored by SIAM through an endowment from the estate of Gene Golub.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Laura Grigori, March 2014 – current. Member of the editorial board for the SIAM book series Software, Environments and Tools. See http://bookstore.siam.org/software-environments-and-tools/.
- Laura Grigori, January 2016 – current. Associate Editor, SIAM Journal on Scientific Computing.
- Laura Grigori, January 2017 – current. Associate Editor, SIAM Journal on Matrix Analysis and Applications.
- Laura Grigori, January 2016 – current. Editorial board, Numerical linear algebra with applications Journal, Wiley.
- Laura Grigori, October 2019 – current. Editorial board, Calcolo Journal, Springer.
- Frédéric Nataf, January 2015 – current, Editorial board, Journal of Numerical Mathematics, de Gruyter.

10.1.4. Invited Talks

- Laura Grigori was
  - Invited speaker, Advanced Solvers for modern Architectures Symposium, November 2019, Munster Germany.
  - Keynote speaker, workshop *Parallel solution methods for systems arising from PDEs*, September 2019, CIRM France.
  - Invited plenary speaker, SIAM Conference on Computational Science and Engineering (CSE19), February 2019, Washington USA, http://www.siam.org/meetings/cse19/.
- S. Hirstoaga gave a talk at the seminar of Laboratoire Jacques-Louis Lions, March 2019, Paris.
- Frédéric Nataf was invited speaker at RMR 2019: Rencontres Mathématiques de Rouen 2019, juin 2019, Rouen, France.
- Frédéric Hecht was
  - invited speaker at HPCSE 2019: HIGH PERFORMANCE COMPUTING IN SCIENCE AND ENGINEERING, University of Ostrava, Czech.
  - invited speaker at RMR 2019: Rencontres Mathématiques de Rouen 2019, juin 2019, Rouen, France.
  - invited speaker at COUPLED 2019: june 2019, Sigtes, Spain.
  - invited speaker, workshop *Parallel solution methods for systems arising from PDEs*, September 2019, CIRM France.

10.1.5. Leadership within the Scientific Community

- Laura Grigori, member elected of SIAM Council, January 2018 - December 2020, the committee supervising the scientific activities of SIAM. Nominated by a Committee and elected by the members of SIAM.
- Laura Grigori, member of the PRACE (Partnership for Advanced Computing in Europe) Scientific Steering Committee, September 2016 - current.
- Laura Grigori and Frédéric Hecht are coordinators of the High Performance in Scientific Computing Major of second year of Mathematics and Applications Master, Sorbonne University.

10.1.6. Scientific Expertise

- Laura Grigori: November 2015 - current, expert to the Scientific Commission of IFPEN (French Petroleum Institute). Evaluation of research programs, PhD theses, work representing a total of 5 days per year.

10.1.7. Research Administration

- Laura Grigori was vice-president of the committee CE46 of ANR, September 2017 - July 2019.
- Frédéric Nataf was president of the committee "Mathématiques et leurs interactions" (CE40) at ANR, September 2016 - August 2019.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master 1: Xavier Claeys, Initiation to C++, 36 hrs of programming tutorials in C++, SU.

Master 1: Xavier Claeys, Computational Linear Algebra, 32 hrs of lectures, SU.

Master 1: Xavier Claeys, Approximation of PDEs, 24 hrs of programming tutorials in Python, SU.

Master 2: Laura Grigori, Winter 2019, Participation in the course on High Performance Computing given at Sorbonne University, Computer Science, intervention for 8 hours per year.

Master 2: Laura Grigori, Course on *High performance computing for numerical methods and data analysis*, https://who.rocq.inria.fr/Laura.Grigori/TeachingDocs/UPMC_Master2/HPC_MN_DA.html, Master 2nd year, Mathematics & Applications, Sorbonne University, 24 hours per year.

Licence 1: S. Hirstoaga, analysis and matrix calculus, 24 HETD, INSEP, Sorbonne University.

Master 1: Frédéric Hecht, Initiation au C++, 24hrs, SU, France

Master 2: Frédéric Hecht, Des EDP à leur résolution par la méthode des éléments finis (MEF), 36hrs, M2, SU, France

Master 2: Frédéric Hecht, Numerical methods for fluid mechanics, 10hrs, SU, France

Master 2: Frédéric Hecht, Calcul scientifique 3 / projet industriel FreeFem++, 28hrs, M2, SU, France

Master 2: Frédéric Hecht, Ingénierie 1 / Logiciel pour la simulation (FreeFem++), 21hrs, SU, France

Master 2: Frédéric Hecht, Ingénierie 2 / Projet collaboratif, 21hrs, SU, France

Cours CIMPA: Frédéric Hecht, Modélisation, Analyse mathématique et calcul scientifique dans la gestion des déchets ménagers , 12h, Kenitra, Maroc

Master 2: Frédéric Nataf, Course on Domain Decomposition Methods, Sorbonne University.

10.2.2. Supervision

PhD: Sebastien Cayrols, Minimizing communication for incomplete factorizations and low-rank approximations on large scale computers, Sorbonne Université, February 2019, adivsor Laura Grigori.

PhD: Zakariae Jorti, Fast solution of sparse linear systems with adaptive choice of preconditioners, Sorbonne University, September 2019, advisor Laura Grigori.

PhD: Olivier Tissot, Solving linear systems on massively parallel architectures, Sorbonne University, January 2019, advisor Laura Grigori.

PhD: Julien Coulet, Sorbonne University, Méthode des Eléments Virtuels pour le calcul de la déformation mécanique couplée aux écoulements en milieux poreux, October 2019, advisors Frédéric Nataf, Vivette Girault (SU), Isabelle Faille (IFPEN) and Nicolas Guy (IFPEN).

PhD: Rim El Dbaissy, Sorbonne University, Discrétisation du problème de couplage instationnaire des équations de Navier- Stokes avec l'équation de la chaleur, Mars 2019, advisors Tony Sayah, Frédéric Hecht.

PhD in progress: Matthias Beaupere, since October 2019 (funded by ERC Synergy EMC2), advisor Laura Grigori.

PhD in progress: Igor Chollet, since October 2017 (funded by ICSD), advisors Xavier Claeys, Pierre Fortin, Laura Grigori.

PhD in progress: Thibault Cimic, since October 2017 (funded by ANR B3DCMB), advisor Laura Grigori.

PhD in progress: Pierre Marchand, since October 2016 (funded by ANR NonLocalDD project), advisors Xavier Claeys and Frédéric Nataf.

PhD in progress: Thanh Van Nguyen, since November 2017 (funded by ANR CinePara), advisor Laura Grigori.

PhD in progress: Jeremy Weisman, since October 2019 (funded by ERC Synergy EMC2), advisors Laura Grigori and J. P. Piquemal.

PhD in progress: Houssam Houssein, since January 2018 (CIFRE, funded by Airthium), advisor Frédéric Hecht.

10.2.3. Juries

- Xavier Claeys was reviewer and member of the jury for the PhD thesis of Federica CAFORIO, defended at Université Paris-Saclay, on January the 24th 2019.
- Xavier Claeys was member of the jury for the PhD thesis of Yohannes Tjandrawidjaja, defended at Université Paris-Saclay on December the 17th 2019.
- Laura Grigori was a member of the jury of Phd defense of Oleg Balabanov, September 2019.
- Laura Grigori was president of the jury of the HDR habilitation of Sever Hirstoaga, Université de Strasbourg, April 2019.
- Laura Grigori was a member of the jury of HDR habilitation of Stéphanie Chaillat, Examinateur, ENSTA, March 2019.
- Laura Grigori was a member of the jury of Phd defense of Fabio Verbosio, USI Lugano, February 2019.
- Frédéric Hecht was president of the HDR defense of Mourad Ismail, 2019, LIPhy, Université Grenoble Alpes.
- Frédéric Hecht was president of the PhD defense of Alexandre This, 2019, Sorbonne University.
- Frédéric Hecht was president of the PhD defense of Zakariae Jorti, 2019, Sorbonne University.
- Frédéric Hecht was president of the PhD defense of Sebastien Cayrols, 2019, Sorbonne University.
- Frédéric Nataf was president of the PhD defense of Rim ALDBAISSY, 2019, Sorbonne University.

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- Laura Grigori was vice-president of the Evaluation Commission of Inria, March 2018 - August 2019.

11. Bibliography

Major publications by the team in recent years

[1] X. CLAEYS.*Essential spectrum of local multi-trace boundary integral operators*, in "IMA J. Appl. Math.", 2016, vol. 81, n^{o} 6, p. 961–983, https://doi-org.accesdistant.sorbonne-universite.fr/10.1093/imamat/hxw019

[2] X. CLAEYS, R. HIPTMAIR.*Integral equations for electromagnetic scattering at multi-screens*, in "Integral Equations Operator Theory", 2016, vol. 84, n^{o} 1, p. 33–68, https://doi-org.accesdistant.sorbonne-universite.fr/10.1007/s00020-015-2242-5

[3] X. CLAEYS, R. HIPTMAIR, E. SPINDLER.*Second kind boundary integral equation for multi-subdomain diffusion problems*, in "Adv. Comput. Math.", 2017, vol. 43, n^{o} 5, p. 1075–1101, https://doi-org.accesdistant.sorbonne-universite.fr/10.1007/s10444-017-9517-0

[4] J. W. DEMMEL, L. GRIGORI, M. HOEMMEN, J. LANGOU.*Communication-optimal parallel and sequential QR and LU factorizations*, in "SIAM Journal on Scientific Computing", 2012, n^{o} 1, p. 206-239, short version of technical report UCB/EECS-2008-89 from 2008

[5] V. DOLEAN, P. JOLIVET, F. NATAF.*An Introduction to Domain Decomposition Methods: algorithms, theory and parallel implementation*, SIAM, 2015

[6] L. GRIGORI, J. W. DEMMEL, H. XIANG.*CALU: a communication optimal LU factorization algorithm*, in "SIAM Journal on Matrix Analysis and Applications", 2011, vol. 32, p. 1317-1350

[7] L. GRIGORI, S. MOUFAWAD, F. NATAF.*Enlarged Krylov Subspace Conjugate Gradient methods for Reducing Communication*, in "SIAM Journal on Matrix Analysis and Applications", 2016, vol. 37, n^{o} 2, p. 744-773

[8] R. HAFERSSAS, P. JOLIVET, F. NATAF.*An Additive Schwarz Method Type Theory for Lions's Algorithm and a Symmetrized Optimized Restricted Additive Schwarz Method*, in "SIAM J. Sci. Comput.", 2017, vol. 39, n^{o} 4, p. A1345–A1365, http://dx.doi.org/10.1137/16M1060066

[9] F. HECHT.*New development in FreeFem++*, in "J. Numer. Math.", 2012, vol. 20, n^{o} 3-4, p. 251–265

[10] P. JOLIVET, F. NATAF.*HPDDM: High-Performance Unified framework for Domain Decomposition methods, MPI-C++ library*, 2014, https://github.com/hpddm/hpddm

[11] N. SPILLANE, V. DOLEAN, P. HAURET, F. NATAF, C. PECHSTEIN, R. SCHEICHL.*Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps*, in "Numer. Math.", 2014, vol. 126, n^{o} 4, p. 741–770, http://dx.doi.org/10.1007/s00211-013-0576-y

Publications of the year

Doctoral Dissertations and Habilitation Theses

[12] R. ALDBAISSY.*Scalable domain decomposition preconditioner for Navier-Stokes equations coupled with the heat equation*, Sorbonne Université / Université Pierre et Marie Curie - Paris VI ; Université Saint-Joseph (Beyrouth), March 2019, https://hal.archives-ouvertes.fr/tel-02458872

[13] S. CAYROLS.*Minimizing communication for incomplete factorizations and low-rank approximations on large scale computers*, Sorbonne Universites, UPMC University of Paris 6, February 2019, https://tel.archives-ouvertes.fr/tel-02437769

[14] J. COULET.*Coupling Virtual Element Methods and finite volume schemes for computational geomechanics*, Sorbonne-Université, November 2019, https://tel.archives-ouvertes.fr/tel-02416511

[15] S. A. HIRSTOAGA.*Design and performant implementation of numerical methods for multiscale problems in plasma physics*, Université de Strasbourg, IRMA UMR 7501, April 2019, Habilitation à diriger des recherches, https://tel.archives-ouvertes.fr/tel-02081304

[16] Z. JORTI.*Fast solution of sparse linear systems with adaptive choice of preconditioners*, Sorbonne Université / Université Pierre et Marie Curie - Paris VI, October 2019, https://tel.archives-ouvertes.fr/tel-02425679

[17] O. TISSOT.*Iterative methods for solving linear systems on massively parallel architectures*, Sorbonne Université, January 2019, https://tel.archives-ouvertes.fr/tel-02428348

Articles in International Peer-Reviewed Journal

[18] A. ANCIAUX-SEDRAKIAN, L. GRIGORI, Z. JORTI, J. PAPEŽ, S. YOUSEF.*Adaptive solution of linear systems of equations based on a posteriori error estimators*, in "Numerical Algorithms", May 2019 [*DOI :* 10.1007/S11075-019-00757-Z], https://hal.inria.fr/hal-02425391

[19] A. AYALA, X. CLAEYS, L. GRIGORI.*ALORA: Affine Low-Rank Approximations*, in "Journal of Scientific Computing", May 2019, vol. 79, n^o 2, p. 1135–1160, https://hal.inria.fr/hal-01762882

[20] A. AYALA, X. CLAEYS, L. GRIGORI.*Linear-time CUR approximation of BEM matrices*, in "Journal of Computational and Applied Mathematics", April 2020, vol. 368, n^o 112528, https://hal.inria.fr/hal-01893036

[21] M. BONAZZOLI, V. DOLEAN, I. G. GRAHAM, E. A. SPENCE, P.-H. TOURNIER.*Domain decomposition preconditioning for the high-frequency time-harmonic Maxwell equations with absorption*, in "Mathematics of Computation", May 2019, https://arxiv.org/abs/1711.03789 , forthcoming [*DOI :* 10.1090/MCOM/3447], https://hal.archives-ouvertes.fr/hal-01644011

[22] X. CLAEYS, R. HIPTMAIR.*First-Kind Boundary Integral Equations for the Hodge-Helmholtz Operator*, in "SIAM Journal on Mathematical Analysis", January 2019, https://hal.archives-ouvertes.fr/hal-02425385

[23] X. CLAEYS, R. HIPTMAIR.*First-Kind Galerkin Boundary Element Methods for the Hodge-Laplacian in Three Dimensions*, in "Mathematical Methods in the Applied Sciences", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-02425493

[24] H. A. DAAS, L. GRIGORI.*A Class of Efficient Locally Constructed Preconditioners Based on Coarse Spaces*, in "SIAM Journal on Matrix Analysis and Applications", January 2019, vol. 40, n^o 1, p. 66-91 [*DOI :* 10.1137/18M1194365], https://hal.inria.fr/hal-02425450

[25] S. DIB, V. GIRAULT, F. HECHT, T. SAYAH.*A posteriori error estimates for Darcy's problem coupled with the heat equation*, in "ESAIM: Mathematical Modelling and Numerical Analysis", June 2019, vol. 53, n^o 6, p. 2121 - 2159 [*DOI :* 10.1051/M2AN/2019049], https://hal.archives-ouvertes.fr/hal-02427095

[26] M. GRAFF, M. J. GROTE, F. NATAF, F. ASSOUS.*How to solve inverse scattering problems without knowing the source term: a three-step strategy*, in "Inverse Problems", September 2019, vol. 35, n^o 10, 104001, Special Issue in Memory of Professor Armin Lechleiter [*DOI :* 10.1088/1361-6420/AB2D5F], https://hal.archives-ouvertes.fr/hal-02002931

[27] L. GRIGORI, Q. NIU, Y. XU.*Stabilized dimensional factorization preconditioner for solving incompressible Navier-Stokes equations*, in "Applied Numerical Mathematics", December 2019, vol. 146, p. 309-327 [*DOI :* 10.1016/J.APNUM.2019.05.026], https://hal.inria.fr/hal-02425402

[28] L. GRIGORI, O. TISSOT.*Scalable Linear Solvers Based on Enlarged Krylov Subspaces with Dynamic Reduction of Search Directions*, in "SIAM Journal on Scientific Computing", January 2019, vol. 41, n^{o} 5, p. C522-C547 [*DOI :* 10.1137/18M1196285], https://hal.inria.fr/hal-02425400

[29] F. NATAF.*Mathematical Analysis of Robustness of Two-Level Domain Decomposition Methods with respect to Inexact Coarse Solves*, in "Numerische Mathematik", February 2020, https://arxiv.org/abs/1912.13225 [*DOI :* 10.1007/S00211-020-01102-6], https://hal.archives-ouvertes.fr/hal-01573197

[30] P.-H. TOURNIER, I. ALIFERIS, M. BONAZZOLI, M. DE BUHAN, M. DARBAS, V. DOLEAN, F. HECHT, P. JOLIVET, I. EL KANFOUD, C. MIGLIACCIO, F. NATAF, C. PICHOT, S. SEMENOV.*Microwave Tomographic Imaging of Cerebrovascular Accidents by Using High-Performance Computing*, in "Parallel Computing", July 2019, vol. 85, p. 88-97, https://arxiv.org/abs/1607.02573 [*DOI :* 10.1016/J.PARCO.2019.02.004], https://hal.archives-ouvertes.fr/hal-01343687

International Conferences with Proceedings

[31] L. EYRAUD-DUBOIS, S. KUMAR.*Analysis of a List Scheduling Algorithm for Task Graphs on Two Types of Resources*, in "IPDPS 2020 - 34th IEEE International Parallel and Distributed Procesing Symposium", New Orleans, United States, May 2020, https://hal.inria.fr/hal-02431810

Other Publications

[32] H. AL DAAS, L. GRIGORI, P. JOLIVET, P.-H. TOURNIER.*A Multilevel Schwarz Preconditioner Based on a Hierarchy of Robust Coarse Spaces*, June 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02151184

[33] V. EHRLACHER, L. GRIGORI, D. LOMBARDI, H. SONG.*Adaptive hierarchical subtensor partitioning for tensor compression*, September 2019, working paper or preprint, https://hal.inria.fr/hal-02284456

[34] M. J. GROTE, F. NATAF, J. HOE TANG, P.-H. TOURNIER.*Parallel Controllability Methods For the Helmholtz Equation*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02046089

[35] H. HOUSSEIN, S. GARNOTEL, F. HECHT.*Frictionless contact problem for hyperelastic materials with interior point optimizer*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02355429

[36] P. MARCHAND, X. CLAEYS, P. JOLIVET, F. NATAF, P.-H. TOURNIER.*Two-level preconditioning for h-version boundary element approximation of hypersingular operator with GenEO*, July 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02188771

[37] F. NATAF.*Adaptive Domain Decomposition method for Saddle Point problem in Matrix Form*, November 2019, https://arxiv.org/abs/1911.01858 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02343808

[38] J. PAPEŽ, L. GRIGORI, R. STOMPOR.*Accelerating linear system solvers for time domain component separation of cosmic microwave background data*, February 2020, working paper or preprint, https://hal.inria.fr/hal-02470964

References in notes

[39] F. ASSOUS, M. KRAY, F. NATAF, E. TURKEL.*Time-reversed absorbing condition: application to inverse problems*, in "Inverse Problems", 2011, vol. 27, n^{o} 6, 065003, http://stacks.iop.org/0266-5611/27/i=6/a=065003

[40] M. J. GROTE, M. KRAY, F. NATAF, F. ASSOUS.*Time-dependent wave splitting and source separation*, in "Journal of Computational Physics", 2017, vol. 330, p. 981 - 996 [*DOI :* 10.1016/J.JCP.2016.10.021], http://www.sciencedirect.com/science/article/pii/S0021999116305198

Project-Team ANGE

Numerical Analysis, Geophysics and Environment

IN COLLABORATION WITH: Laboratoire Jacques-Louis Lions (LJLL)

IN PARTNERSHIP WITH:
CNRS

Centre d'expertise des risques, de l'environnement, des mobilités et de l'aménagement

Sorbonne Université (UPMC)

RESEARCH CENTER
Paris

THEME
Earth, Environmental and Energy Sciences

Table of contents

Project-Team ANGE

Creation of the Team: 2012 November 01, updated into Project-Team: 2014 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
A6.1.5. - Multiphysics modeling
A6.2. - Scientific computing, Numerical Analysis & Optimization
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.6. - Optimization
A6.3. - Computation-data interaction
A6.3.2. - Data assimilation
A6.3.4. - Model reduction
A6.3.5. - Uncertainty Quantification

Other Research Topics and Application Domains:

B3. - Environment and planet
B3.3. - Geosciences
B3.3.2. - Water: sea & ocean, lake & river
B3.3.3. - Nearshore
B3.4. - Risks
B3.4.1. - Natural risks
B3.4.3. - Pollution
B4. - Energy
B4.3. - Renewable energy production
B4.3.1. - Biofuels
B4.3.2. - Hydro-energy

1. Team, Visitors, External Collaborators

Research Scientists

Jacques Sainte-Marie [Team leader, Inria, Senior Researcher, HDR]
Vivien Mallet [Inria, Researcher]
Martin Parisot [Inria, Researcher, until Aug 2019]
Yohan Penel [Inria, Advanced Research Position]
Julien Salomon [Inria, Senior Researcher]

Faculty Members

Nina Aguillon [Sorbonne Université, Associate Professor]
Edwige Godlewski [Sorbonne Université, Professor, HDR]
Cindy Guichard [Sorbonne Université, Associate Professor]
Boris Haspot [Univ de Dauphine, Associate Professor, until Aug 2019]

Anne Mangeney [IPGP, Professor, until Aug 2019, HDR]

External Collaborators

Emmanuel Audusse [Univ Paris-Nord]
Bernard Di Martino [Univ. de Corse, HDR]

Technical Staff

Cédric Doucet [Inria, Engineer, until Mar 2019]
Apolline El Baz [Inria, Engineer, from Dec 2019]
Jérémy Ledoux [Univ de Dauphine, Engineer, until Sep 2019]
Guillaume Chérel [Inria, from Feb 2019]
Fabien Souillé [EDF, until Aug 2019]

PhD Students

Frederic Allaire [Inria, PhD Student]
Léa Boittin [Inria, PhD Student, until Feb 2019]
Nelly Boulos Al Makary [Univ Paris-Nord, PhD Student]
Virgile Dubos [Sorbonne Université, PhD Student]
Ngoc Le [IRSN, PhD Student, until Oct 2019]
Antoine Lesieur [Inria, PhD Student]
Liudi Lu [Sorbonne Université, PhD Student]
Hugo Martin [Sorbonne Université, PhD Student, until Apr 2019]
Mathieu Rigal [Sorbonne Université, PhD Student, from Oct 2019]

Post-Doctoral Fellow

Janelle Hammond [Inria, Post-Doctoral Fellow]

Visiting Scientist

Marie-Odile Bristeau [retired]

Administrative Assistants

Laurence Bourcier [Inria, Administrative Assistant]
Maryse Desnous [Inria, Administrative Assistant, until Apr 2019]
Julien Guieu [Inria, Administrative Assistant, from Apr 2019]

2. Overall Objectives

2.1. Presentation

Among all aspects of geosciences, we mainly focus on gravity driven flows arising in many situations such as

- hazardous flows (flooding, rogue waves, landslides...),
- sustainable energies (hydrodynamics-biology coupling, biofuel production, marine energies...),
- risk management and land-use planning (morphodynamic evolutions, early warning systems...)

There exists a strong demand from scientists and engineers in fluid mechanics for models and numerical tools able to simulate not only the water depth and the velocity field but also the distribution and evolution of external quantities such as pollutants or biological species and the interaction between flows and structures (seashores, erosion processes...). The key point of the researches carried out within ANGE is to answer this demand by the development of efficient, robust and validated models and numerical tools.

2.2. Scientific challenges

Due to the variety of applications with a wide range of spatial scales, reduced-size models like the shallow water equations are generally required. From the modelling point of view, the main issue is to describe the behaviour of the flow with a reduced-size model taking into account several physical processes such as non-hydrostatic terms, biological species evolution, topography and structure interactions within the flow. The mathematical analysis of the resulting model do not enter the field of hyperbolic equations anymore and new strategies have to be proposed. Moreover, efficient numerical resolutions of reduced-size models require particular attention due to the different time scales of the processes and in order to recover physical properties such as positivity, conservativity, entropy dissipation and equilibria.

The models can remain subject to uncertainties that originate from incomplete description of the physical processes and from uncertain parameters. Further development of the models may rely on the assimilation of observational data and the uncertainty quantification of the resulting analyses or forecasts.

3. Research Program

3.1. Overview

The research activities carried out within the ANGE team strongly couple the development of methodological tools with applications to real–life problems and the transfer of numerical codes. The main purpose is to obtain new models adapted to the physical phenomena at stake, identify the main properties that reflect the physical meaning of the models (uniqueness, conservativity, entropy dissipation, ...), propose effective numerical methods to approximate their solution in complex configurations (multi-dimensional, unstructured meshes, well-balanced, ...) and to assess the results with data in the purpose of potentially correcting the models.

The difficulties arising in gravity driven flow studies are threefold.

- Models and equations encountered in fluid mechanics (typically the free surface Navier-Stokes equations) are complex to analyze and solve.
- The underlying phenomena often take place over large domains with very heterogeneous length scales (size of the domain, mean depth, wave length, ...) and distinct time scales, *e.g.* coastal erosion, propagation of a tsunami, ...
- These problems are multi-physics with strong couplings and nonlinearities.

3.2. Modelling and analysis

Hazardous flows are complex physical phenomena that can hardly be represented by shallow water type systems of partial differential equations (PDEs). In this domain, the research program is devoted to the derivation and analysis of reduced complexity models compared to the Navier-Stokes equations, but relaxing the shallow water assumptions. The main purpose is then to obtain models well-adapted to the physical phenomena at stake.

Even if the resulting models do not strictly belong to the family of hyperbolic systems, they exhibit hyperbolic features: the analysis and discretisation techniques we intend to develop have connections with those used for hyperbolic conservation laws. It is worth noticing that the need for robust and efficient numerical procedures is reinforced by the smallness of dissipative effects in geophysical models which therefore generate singular solutions and instabilities.

On the one hand, the derivation of the Saint-Venant system from the Navier-Stokes equations is based on two approximations (the so-called shallow water assumptions), namely

- the horisontal fluid velocity is well approximated by its mean value along the vertical direction,
- the pressure is hydrostatic or equivalently the vertical acceleration of the fluid can be neglected compared to the gravitational effects.

As a consequence the objective is to get rid of these two assumptions, one after the other, in order to obtain models accurately approximating the incompressible Euler or Navier-Stokes equations.

On the other hand, many applications require the coupling with non-hydrodynamic equations, as in the case of micro-algae production or erosion processes. These new equations comprise non-hyperbolic features and a special analysis is needed.

3.2.1. Multilayer approach

As for the first shallow water assumption, *multi-layer* systems were proposed to describe the flow as a superposition of Saint-Venant type systems [26], [29], [30]. Even if this approach has provided interesting results, layers are considered separate and non-miscible fluids, which implies strong limitations. That is why we proposed a slightly different approach [27], [28] based on a Galerkin type decomposition along the vertical axis of all variables and leading, both for the model and its discretisation, to more accurate results.

A kinetic representation of our multilayer model allows to derive robust numerical schemes endowed with crucial properties such as: consistency, conservativity, positivity, preservation of equilibria, ... It is one of the major achievements of the team but it needs to be analyzed and extended in several directions namely:

- The convergence of the multilayer system towards the hydrostatic Euler system as the number of layers goes to infinity is a critical point. It is not fully satisfactory to have only formal estimates of the convergence and sharp estimates would provide an optimal number of layers.
- The introduction of several source terms due for instance to the Coriolis force or extra terms from changes of coordinates seems necessary. Their inclusion should lead to substantial modifications of the numerical scheme.
- Its hyperbolicity has not yet been proven and conversely the possible loss of hyperbolicity cannot be characterised. Similarly, the hyperbolic feature is essential in the propagation and generation of waves.

3.2.2. Non-hydrostatic models

The hydrostatic assumption consists in neglecting the vertical acceleration of the fluid. It is considered valid for a large class of geophysical flows but is restrictive in various situations where the dispersive effects (like wave propagation) cannot be neglected. For instance, when a wave reaches the coast, bathymetry variations give a vertical acceleration to the fluid that strongly modifies the wave characteristics and especially its height.

Processing an asymptotic expansion (w.r.t. the aspect ratio for shallow water flows) into the Navier-Stokes equations, we obtain at the leading order the Saint-Venant system. Going one step further leads to a vertically averaged version of the Euler/Navier-Stokes equations involving some non-hydrostatic terms. This model has several advantages:

- it admits an energy balance law (that is not the case for most dispersive models available in the literature),
- it reduces to the Saint-Venant system when the non-hydrostatic pressure term vanishes,
- it consists in a set of conservation laws with source terms,
- it does not contain high order derivatives.

3.2.3. Multi-physics modelling

The coupling of hydrodynamic equations with other equations in order to model interactions between complex systems represents an important part of the team research. More precisely, three multi-physics systems are investigated. More details about the industrial impact of these studies are presented in the following section.

- To estimate the risk for infrastructures in coastal zones or close to a river, the resolution of the shallow water equations with moving bathymetry is necessary. The first step consisted in the study of an additional equation largely used in engineering science: The Exner equation. The analysis enabled to exhibit drawbacks of the coupled model such as the lack of energy conservation or the strong variations of the solution from small perturbations. A new formulation is proposed to avoid these drawbacks. The new model consists in a coupling between conservation laws and an elliptic equation, like the Euler/Poisson system, suggesting to use well-known strategies for the analysis and the numerical resolution. In addition, the new formulation is derived from classical complex rheology models and allowed physical phenomena like threshold laws.

- Interaction between flows and floating structures is the challenge at the scale of the shallow water equations. This study requires a better understanding of the energy exchanges between the flow and the structure. The mathematical model of floating structures is very hard to solve numerically due to the non-penetration condition at the interface between the flow and the structure. It leads to infinite potential wave speeds that could not be solved with classical free surface numerical schemes. A relaxation model was derived to overcome this difficulty. It represents the interaction with the floating structure with a free surface model-type.
- If the interactions between hydrodynamics and biology phenomena are known through laboratory experiments, it is more difficult to predict the evolution, especially for the biological quantities, in a real and heterogeneous system. The objective is to model and reproduce the hydrodynamics modifications due to forcing term variations (in time and space). We are typically interested in phenomena such as eutrophication, development of harmful bacteria (cyanobacteria) and upwelling phenomena.

3.2.4. Data assimilation and inverse modelling

In environmental applications, the most accurate numerical models remain subject to uncertainties that originate from their parameters and shortcomings in their physical formulations. It is often desirable to quantify the resulting uncertainties in a model forecast. The propagation of the uncertainties may require the generation of ensembles of simulations that ideally sample from the probability density function of the forecast variables. Classical approaches rely on multiple models and on Monte Carlo simulations. The applied perturbations need to be calibrated for the ensemble of simulations to properly sample the uncertainties. Calibrations involve ensemble scores that compare the consistency between the ensemble simulations and the observational data. The computational requirements are so high that designing fast surrogate models or metamodels is often required.

In order to reduce the uncertainties, the fixed or mobile observations of various origins and accuracies can be merged with the simulation results. The uncertainties in the observations and their representativeness also need to be quantified in the process. The assimilation strategy can be formulated in terms of state estimation or parameter estimation (also called inverse modelling). Different algorithms are employed for static and dynamic models, for analyses and forecasts. A challenging question lies in the optimization of the observational network for the assimilation to be the most efficient at a given observational cost.

3.3. Numerical analysis

3.3.1. Non-hydrostatic scheme

The main challenge in the study of the non-hydrostatic model is to design a robust and efficient numerical scheme endowed with properties such as: positivity, wet/dry interfaces treatment, consistency. It must be noticed that even if the non-hydrostatic model looks like an extension of the Saint-Venant system, most of the known techniques used in the hydrostatic case are not efficient as we recover strong difficulties encountered in incompressible fluid mechanics due to the extra pressure term. These difficulties are reinforced by the absence of viscous/dissipative terms.

3.3.2. Space decomposition and adaptive scheme

In the quest for a better balance between accuracy and efficiency, a strategy consists in the adaptation of models. Indeed, the systems of partial differential equations we consider result from a hierarchy of simplifying assumptions. However, some of these hypotheses may turn out to be irrelevant locally. The adaptation of models thus consists in determining areas where a simplified model (*e.g.* shallow water type) is valid and where it is not. In the latter case, we may go back to the "parent" model (*e.g.* Euler) in the corresponding area.

This implies to know how to handle the coupling between the aforementioned models from both theoretical and numerical points of view. In particular, the numerical treatment of transmission conditions is a key point. It requires the estimation of characteristic values (Riemann invariant) which have to be determined according to the regime (torrential or fluvial).

3.3.3. Asymptotic-Preserving scheme for source terms

Hydrodynamic models comprise advection and sources terms. The conservation of the balance between source terms, typically viscosity and friction, has a significant impact since the overall flow is generally a perturbation around an equilibrium. The design of numerical schemes able to preserve such balances is a challenge from both theoretical and industrial points of view. The concept of Asymptotic-Preserving (AP) methods is of great interest in order to overcome these issues.

Another difficulty occurs when a term, typically related to the pressure, becomes very large compared to the order of magnitude of the velocity. At this regime, namely the so-called *low Froude* (shallow water) or *low Mach* (Euler) regimes, the difference between the speed of the gravity waves and the physical velocity makes classical numerical schemes inefficient: firstly because of the error of truncation which is inversely proportional to the small parameters, secondly because of the time step governed by the largest speed of the gravity wave. AP methods made a breakthrough in the numerical resolution of asymptotic perturbations of partial-differential equations concerning the first point. The second one can be fixed using partially implicit scheme.

3.3.4. Multi-physics models

Coupling problems also arise within the fluid when it contains pollutants, density variations or biological species. For most situations, the interactions are small enough to use a splitting strategy and the classical numerical scheme for each sub-model, whether it be hydrodynamic or non-hydrodynamic.

The sediment transport raises interesting issues from a numerical aspect. This is an example of coupling between the flow and another phenomenon, namely the deformation of the bottom of the basin that can be carried out either by bed load where the sediment has its own velocity or suspended load in which the particles are mostly driven by the flow. This phenomenon involves different time scales and nonlinear retroactions; hence the need for accurate mechanical models and very robust numerical methods. In collaboration with industrial partners (EDF–LNHE), the team already works on the improvement of numerical methods for existing (mostly empirical) models but our aim is also to propose new (quite) simple models that contain important features and satisfy some basic mechanical requirements. The extension of our 3D models to the transport of weighted particles can also be here of great interest.

3.3.5. Optimisation

Numerical simulations are a very useful tool for the design of new processes, for instance in renewable energy or water decontamination. The optimisation of the process according to a well-defined objective such as the production of energy or the evaluation of a pollutant concentration is the logical upcoming challenge in order to propose competitive solutions in industrial context. First of all, the set of parameters that have a significant impact on the result and on which we can act in practice is identified. Then the optimal parameters can be obtained using the numerical codes produced by the team to estimate the performance for a given set of parameters with an additional loop such as gradient descent or Monte Carlo method. The optimisation is used in practice to determine the best profile for turbine pales, the best location for water turbine implantation, in particular for a farm.

4. Application Domains

4.1. Overview

Sustainable development and environment preservation have a growing importance and scientists have to address difficult issues such as: management of water resources, renewable energy production, bio/geo-chemistry of oceans, resilience of society w.r.t. hazardous flows, urban pollutions, ...

As mentioned above, the main issue is to propose models of reduced complexity, suitable for scientific computing and endowed with stability properties (continuous and/or discrete). In addition, models and their numerical approximations have to be confronted with experimental data, as analytical solutions are hardly accessible for these problems/models. A. Mangeney (IPGP) and N. Goutal (EDF) may provide useful data.

4.2. Geophysical flows

Reduced models like the shallow water equations are particularly well-adapted to the modelling of geophysical flows since there are characterized by large time or/and space scales. For long time simulations, the preservation of equilibria is essential as global solutions are a perturbation around them. The analysis and the numerical preservation of non-trivial equilibria, more precisely when the velocity does not vanish, are still a challenge. In the fields of oceanography and meteorology, the numerical preservation of the so-called geostrophic state, which is the balance between the gravity field and the Coriolis force, can significantly improve the forecasts. In addition, data assimilation is required to improve the simulations and correct the dissipative effect of the numerical scheme.

The sediment transport modelling is of major interest in terms of applications, in particular to estimate the sustainability of facilities with silt or scour, such as canals and bridges. Dredging or filling-up operations are expensive and generally not efficient in the long term. The objective is to determine a configuration almost stable for the facilities. In addition, it is also important to determine the impact of major events like emptying dam which is aimed at evacuating the sediments in the dam reservoir and requires a large discharge. However, the downstream impact should be measured in terms of turbidity, river morphology and flood.

4.3. Hydrological disasters

It is a violent, sudden and destructive flow. Between 1996 and 2005, nearly 80% of natural disasters in the world have meteorological or hydrological origines. The main interest of their study is to predict the areas in which they may occur most probably and to prevent damages by means of suitable amenities. In France, floods are the most recurring natural disasters and produce the worst damages. For example, it can be a cause or a consequence of a dam break. The large surface they cover and the long period they can last require the use of reduced models like the shallow water equations. In urban areas, the flow can be largely impacted by the debris, in particular cars, and this requires fluid/structure interactions be well understood. Moreover, underground flows, in particular in sewers, can accelerate and amplify the flow. To take them into account, the model and the numerical resolution should be able to treat the transition between free surface and underground flows.

Tsunamis are another hydrological disaster largely studied. Even if the propagation of the wave is globally well described by the shallow water model in oceans, it is no longer the case close to the epicenter and in the coastal zone where the bathymetry leads to vertical accretions and produces substantial dispersive effects. The non-hydrostatic terms have to be considered and an efficient numerical resolution should be induced.

While viscous effects can often be neglected in water flows, they have to be taken into account in situations such as avalanches, debris flows, pyroclastic flows, erosion processes, ...*i.e.* when the fluid rheology becomes more complex. Gravity driven granular flows consist of solid particles commonly mixed with an interstitial lighter fluid (liquid or gas) that may interact with the grains and decrease the intensity of their contacts, thus reducing energy dissipation and favoring propagation. Examples include subaerial or subaqueous rock avalanches (e.g. landslides).

4.4. Biodiversity and culture

Nowadays, simulations of the hydrodynamic regime of a river, a lake or an estuary, are not restricted to the determination of the water depth and the fluid velocity. They have to predict the distribution and evolution of external quantities such as pollutants, biological species or sediment concentration.

The potential of micro-algae as a source of biofuel and as a technological solution for CO2 fixation is the subject of intense academic and industrial research. Large-scale production of micro-algae has potential for biofuel applications owing to the high productivity that can be attained in high-rate raceway ponds. One of the key challenges in the production of micro-algae is to maximize algae growth with respect to the exogenous energy that must be used (paddlewheel, pumps, ...). There is a large number of parameters that need to be optimized (characteristics of the biological species, raceway shape, stirring provided by the paddlewheel). Consequently our strategy is to develop efficient models and numerical tools to reproduce the flow induced by the paddlewheel and the evolution of the biological species within this flow. Here, mathematical models can greatly help us reduce experimental costs. Owing to the high heterogeneity of raceways due to gradients of temperature, light intensity and nutrient availability through water height, we cannot use depth-averaged models. We adopt instead more accurate multilayer models that have recently been proposed. However, it is clear that many complex physical phenomena have to be added to our model, such as the effect of sunlight on water temperature and density, evaporation and external forcing.

Many problems previously mentioned also arise in larger scale systems like lakes. Hydrodynamics of lakes is mainly governed by geophysical forcing terms: wind, temperature variations, ...

4.5. Sustainable energy

One of the booming lines of business is the field of renewable and decarbonated energies. In particular in the marine realm, several processes have been proposed in order to produce electricity thanks to the recovering of wave, tidal and current energies. We may mention water-turbines, buoys turning variations of the water height into electricity or turbines motioned by currents. Although these processes produce an amount of energy which is less substantial than in thermal or nuclear power plants, they have smaller dimensions and can be set up more easily.

The fluid energy has kinetic and potential parts. The buoys use the potential energy whereas the water-turbines are activated by currents. To become economically relevant, these systems need to be optimized in order to improve their productivity. While for the construction of a harbour, the goal is to minimize swell, in our framework we intend to maximize the wave energy.

This is a complex and original issue which requires a fine model of energy exchanges and efficient numerical tools. In a second step, the optimisation of parameters that can be changed in real-life, such as bottom bathymetry and buoy shape, must be studied. Eventually, physical experiments will be necessary for the validation.

4.6. Urban environment

The urban environment is essentially studied for air and noise pollutions. Air pollution levels and noise pollution levels vary a lot from one street to next. The simulations are therefore carried out at street resolution and take into account the city geometry. The associated numerical models are subject to large uncertainties. Their input parameters, e.g. pollution emissions from road traffic, are also uncertain. Quantifying the simulation uncertainties is challenging because of the high computational costs of the numerical models. An appealing approach in this context is the use of metamodels, from which ensembles of simulations can be generated for uncertainty quantification.

The simulation uncertainties can be reduced by the assimilation of fixed and mobile sensors. High-quality fixed monitoring sensors are deployed in cities, and an increasing number of mobile sensors are added to the observational networks. Even smartphones can be used as noise sensors and dramatically increase the spatial coverage of the observations. The processing and assimilation of the observations raises many questions regarding the quality of the measurements and the design of the network of sensors.

4.7. SmartCity

There is a growing interest for environmental problems at city scale, where a large part of the population is concentrated and where major pollutions can occur. Numerical simulation is well established to study the urban environment, *e.g.* for road traffic modelling. As part of the smartcity movement, an increasing number of sensors collect measurements, at traditional fixed observation stations, but also on mobile devices, like smartphones. They must properly be taken into account given their number but also their potential low quality.

Pratical applications include air pollution and noise pollution. These directly relate to road traffic. Data assimilation and uncertainty propagation are key topics in these applications.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Human ressources

Martin Parisot left the team in september 2019 to join the team Cardamom (Inria-Bordeaux). Apoline El Baz joined the team on an ADT position to developp Fresh Kiss 3D. Julien Guieu is the new assistant of the team (May). Mathieu Rigal and Chourouk El Hassanieh started a PhD. (September).

5.1.2. Travel policy

The team started to limit flies for professional activities. A deeper reflexion on the environmental impact of scientific activities has been initiated in the team. C. Guichard and Y. Penel are members of the CLDD of Inria Paris.

5.1.3. Awards and new grants

- J. Sainte-Marie was promoted Adjoint au Directeur Scientifique, in charge of the topic "Sciences de la planète, de l'environnement et de l'énergie". (July)
- M. Parisot was invited for a long term stay at Aachen University (Sept.-Nov.)
- The ANR project ALLOWAP supported by J. Salomon (with L. Halpern, F. Kwok and B. Delourme) was accepted for a start in Jan. 2020.
- M. Parisot and E. Audusse organized the CEMRACS project "Land Slide Tsunami" at CIRM (July-Aug.).
- Y. Penel organized the CEMRACS project "SGN-Num" at CIRM (July-Aug.).
- E. Audusse and Y. Penel organized the CEMRACS project "SW-Cor".

6. New Software and Platforms

6.1. Freshkiss

FREe Surface Hydrodynamics using KInetic SchemeS

KEYWORDS: Finite volume methods - Hydrostatic Navier-Stokes equations - Free surface flows

FUNCTIONAL DESCRIPTION: Freshkiss3D is a numerical code solving the 3D hydrostatic and incompressible Navier-Stokes equations with variable density.

- Participants: Fabien Souillé, Emmanuel Audusse, Jacques Sainte Marie and Marie-Odile Bristeau
- Partners: UPMC - CEREMA
- Contact: Jacques Sainte Marie

6.2. TSUNAMATHS

KEYWORDS: Modeling - Tsunamis

FUNCTIONAL DESCRIPTION: Tsunamaths is an educational platform aiming at simulating historical tsunamis. Real data and mathematical explanations are provided to enable people to better understand the overall process of tsunamis.

- Participants: Emmanuel Audusse, Jacques Sainte Marie and Raouf Hamouda
- Contact: Jacques Sainte Marie
- URL: http://tsunamath.paris.inria.fr/

6.3. Verdandi

KEYWORDS: HPC - Model - Software Components - Partial differential equation

FUNCTIONAL DESCRIPTION: Verdandi is a free and open-source (LGPL) library for data assimilation. It includes various such methods for coupling one or several numerical models and observational data. Mainly targeted at large systems arising from the discretization of partial differential equations, the library is devised as generic, which allows for applications in a wide range of problems (biology and medicine, environment, image processing, etc.). Verdandi also includes tools to ease the application of data assimilation, in particular in the management of observations or for a priori uncertainty quantification. Implemented in C++, the library may be used with models implemented in Fortran, C, C++ or Python.

- Participants: Dominique Chapelle, Gautier Bureau, Nicolas Claude, Philippe Moireau and Vivien Mallet
- Contact: Vivien Mallet
- URL: http://verdandi.gforge.inria.fr/

6.4. Polyphemus

KEYWORD: Simulation

FUNCTIONAL DESCRIPTION: Polyphemus is a modeling system for air quality. As such, it is designed to yield up-to-date simulations in a reliable framework: data assimilation, ensemble forecast and daily forecasts. Its completeness makes it suitable for use in many applications: photochemistry, aerosols, radionuclides, etc. It is able to handle simulations from local to continental scales, with several physical models. It is divided into three main parts:

libraries that gather data processing tools (SeldonData), physical parameterizations (AtmoData) and post-processing abilities (AtmoPy),

programs for physical pre-processing and chemistry-transport models (Polair3D, Castor, two Gaussian models, a Lagrangian model),

model drivers and observation modules for model coupling, ensemble forecasting and data assimilation.

- Participants: Sylvain Doré and Vivien Mallet
- Contact: Vivien Mallet
- URL: http://cerea.enpc.fr/polyphemus/

6.5. Urban noise analysis

KEYWORD: Environment perception

FUNCTIONAL DESCRIPTION: This software processes mobile observations collected by the application Ambiciti (previously known as SoundCity). It can merge simulated noise maps with the mobile observations.

- Authors: Raphaël Ventura, Vivien Mallet and Guillaume Chérel
- Contact: Vivien Mallet

6.6. Freshkiss3D

KEYWORDS: Python - Cython - Navier-Stokes

FUNCTIONAL DESCRIPTION: Tool for the numerical solution of free surface Navier-Stokes equations

- Participants: Cédric Doucet, Apolline El Baz and Jacques Sainte Marie
- Partner: UPMC
- Contact: Jacques Sainte Marie
- Publication: Numerical approximation of the 3d hydrostatic Navier-Stokes system with free surface

7. New Results

7.1. Numerical methods for fluid flows

7.1.1. PARAOPT: A parareal algorithm for optimality systems

Member: J. Salomon,
Coll.: Martin Gander, Felix Kwok

The time parallel solution of optimality systems arising in PDE constraint optimization could be achieved by simply applying any time parallel algorithm, such as Parareal, to solve the forward and backward evolution problems arising in the optimization loop. We propose in [21] a different strategy by devising directly a new time parallel algorithm, which we call ParaOpt, for the coupled forward and backward non-linear partial differential equations. ParaOpt is inspired by the Parareal algorithm for evolution equations, and thus is automatically a two-level method. We provide a detailed convergence analysis for the case of linear parabolic PDE constraints. We illustrate the performance of ParaOpt with numerical experiments both for linear and nonlinear optimality systems.

7.1.2. Dynamical Behavior of a Nondiffusive Scheme for the Advection Equation

Member: N. Aguillon,
Coll.: Pierre-Antoine Guihéneuf

In [16], we study the long time behaviour of a dynamical system strongly linked to the anti-diffusive scheme of Després and Lagoutiere for the 1-dimensional transport equation. This scheme is overcompressive when the Courant–Friedrichs–Levy number is 1/2: when the initial data is nondecreasing, the approximate solution becomes a Heaviside function. In a special case, we also understand how plateaus are formed in the solution and their stability, a distinctive feature of the Després and Lagoutiere scheme.

7.1.3. Convergence of numerical schemes for a conservation equation with convection and degenerate diffusion

Member: C. Guichard,
Coll.: Robert Eymard, Xavier Lhébrard

In [20], the approximation of problems with linear convection and degenerate nonlinear diffusion, which arise in the framework of the transport of energy in porous media with thermodynamic transitions, is done using a θ-scheme based on the centered gradient discretisation method. The convergence of the numerical scheme is proved, although the test functions which can be chosen are restricted by the weak regularity hypotheses on the convection field, owing to the application of a discrete Gronwall lemma and a general result for the time translate in the gradient discretisation setting. Some numerical examples, using both the Control Volume Finite Element method and the Vertex Approximate Gradient scheme, show the role of θ for stabilising the scheme.

7.1.4. Gradient-based optimization of a rotating algal biofilm process

Members: N. Aguillon, J. Sainte-Marie,
Coll.: Pierre-Olivier Lamare, Jérôme Grenier, Hubert Bonnefond, Olivier Bernard

Microalgae are microorganisms that have only very recently been used for bio-technological applications and more specifically for the production of bio-fuel. In the report [15] we focus on the shape optimization and optimal control of an innovative process where microalgae are fixed on a support. They are successively exposed to light and darkness. The resulting growth rate can be represented by a dynamic system describing the denaturation of key proteins due to excess light. A Partial Derivative Equation (PDE) model for the Rotary Algae Biofilm (RAB) is proposed. It represents the local growth of microalgae subjected to time-varying light. A gradient method based on the calculation of the model adjoint is proposed to identify the optimal (constant) folding of the process and the (time-varying) speed of the biofilm. Once this method is used in a realistic case, the optimization results in a configuration that significantly improves productivity compared to the case where the biofilm is fixed.

7.2. Modelling

7.2.1. Accurate steam-water equation of state for two-phase flow LMNC model with phase transition

Member: Y. Penel,
Coll.: Stéphane Dellacherie, Bérénice Grec, Gloria Faccanoni

The paper [9] is dedicated to the design of incomplete equations of state for a two-phase flow with phase transition that are specific to the low Mach number regime. It makes use of the fact that the thermodynamic pressure has small variations in this regime. These equations of state supplement the 2D LMNC model introduced in previous works. This innovative strategy relies on tabulated values and is proven to satisfy crucial thermodynamic requirements such as positivity, monotonicity, continuity. In particular, saturation values are exact. This procedure is assessed by means of analytical steady solutions and comparisons with standard analytical equations of state, and shows a great improvement in accuracy.

7.2.2. Numerical simulations of Serre - Green-Naghdi type models for dispersive free surface flows

Members: Y. Penel, J. Sainte-Marie
Coll.: Enrique D. Fernandez-Nieto, Tomas Morales de Luna, Cipriano Escalante Sanchez

The Serre - Green-Nagdhi equations are simulated under their non-hydrostatic formulation by means of a projection-correction method. This is then extended to the layerwise discretisation of the Euler equations with a special care to the computational cost. An original alternating direction method is used and relies on the tools designed for the monolayer case.

7.2.3. Entropy-satisfying scheme for a hierarchy of dispersive reduced models of free surface flow

Member: M. Parisot

The work [12] is devoted to the numerical resolution in the multidimensional framework of a hierarchy of reduced models of the water wave equations, such as the Serre-Green-Naghdi model. A particular attention is paid to the dissipation of mechanical energy at the discrete level, that act as a stability argument of the scheme, even with source terms such space and time variation of the bathymetry. In addition, the analysis leads to a natural way to deal with dry areas without leakage of energy. To illustrate the accuracy and the robustness of the strategy, several numerical experiments are carried out. In particular, the strategy is capable of treating dry areas without special treatment.

7.2.4. Congested shallow water model: on floating body

Members: E. Godlewski, M. Parisot, J. Sainte-Marie, F. Wahl

In [22], we are interested in the numerical modeling of body floating freely on the water such as icebergs or wave energy converters. The fluid-solid interaction is formulated using a congested shallow water model for the fluid and Newton's second law of motion for the solid. We make a particular focus on the energy transfer between the solid and the water since it is of major interest for energy production. A numerical approximation based on the coupling of a nite volume scheme for the fluid and a Newmark scheme for the solid is presented. An entropy correction based on an adapted choice of discretization for the coupling terms is made in order to ensure a dissipation law at the discrete level. Simulations are presented to verify the method and to show the feasibility of extending it to more complex cases.

7.2.5. Pseudo-compressibility, dispersive model and acoustic waves in shallow water flows

Members: E. Godlewski, M-O. Bristeau, J. Sainte-Marie

In this paper we study a dispersive shallow water type model derived from the compressible Navier-Stokes system. The compressible effects allow to capture the acoustic waves propagation and can be seen as a relaxation of an underlying incompressible model. Hence, the pseudo-compressibility terms circumvent the resolution of an elliptic equation for the non-hydrostatic part of the pressure. For the numerical approximation of shallow water type models, the hy- perbolic part, often approximated using explicit time schemes, is constrained by a CFL condition. Since the approximation of the dispersive terms – im- plicit in time – generally requires the numerical resolution of an elliptic equation, it is very costly. In this paper, we show that when considering the pseudo-compressibility terms a fully explicit in time scheme can be derived. This drastically reduces the cost of the numerical resolution of dispersive models especially in 2d and 3d.

7.2.6. Some quasi-analytical solutions for propagative waves in free surface Euler equations

Members: B. Di Martino, M-O. Bristeau, J. Sainte-Marie, A. Mangeney, F. Souillé

This note describes some quasi-analytical solutions for wave propagation in free surface Euler equations and linearized Euler equations. The obtained solutions vary from a sinusoidal form to a form with singularities. They allow a numerical validation of the free-surface Euler codes.

7.2.7. Challenges and prospects for dynamical cores of oceanic models across all scales

Members: E. Audusse, J. Sainte-Marie
Review paper, more than 30 co-authors

The paper [11] provides an overview of the recent evolution and future challenges of oceanic models dynamical cores used for applications ranging from global paleoclimate scales to short-term prediction in estuaries and shallow coastal areas. The dynamical core is responsible for the discrete approximation in space and time of the resolved processes, as opposed to the physical parameterizations which represent unresolved or under-resolved processes. The paper reviews the challenges and prospects outlined by the modeling groups that participated to the Community for the Numerical Modeling of the Global, Regional, and Coastal Ocean (COMMODORE) workshop. The topics discussed in the paper originate from the experience acquired during the development of 16 dynamical cores representative of the variety of numerical methods implemented in models used for realistic ocean simulations. The topics of interest include the choice of model grid and variables arrangement, vertical coordinate, temporal discretization, and more practical aspects about the evolution of code architecture and development practices.

7.2.8. The Navier-Stokes system with temperature and salinity for free surface flows Part I: Low-Mach approximation & layer-averaged formulation

Members: M-O. Bristeau, L. Boittin, A. Mangeney, J. Sainte-Marie, F. Bouchut

In this paper, we are interested in free surface flows where density variations coming e.g. from temperature or salinity differences play a significant role in the hydrodynamic regime. In water, acoustic waves travel much faster than gravity and internal waves, hence the study of models arising in compressible fluid mechanics often requires a decoupling between these waves. Starting from the compressible Navier-Stokes system, we derive the so-called Navier-Stokes-Fourier system in an incompressible context (the density does not depend on the fluid pressure) using the low-Mach scaling. Notice that a modified low-Mach scaling is necessary to obtain a model with a thermo-mechanical compatibility. The case where the density depends only on the temperature is studied first. Then the variations of the fluid density with respect to the temperature and the salinity are considered. We give a layer-averaged formulation of the obtained models in an hydrostatic context. Allowing to derive numerical schemes endowed with strong stability properties – that are presented in a companion paper – the layer-averaged formulation is very useful for the numerical analysis and the numerical simulations of the models. Several stability properties of the layer-averaged Navier-Stokes-Fourier system are proved.

7.2.9. The Navier-Stokes system with temperature and salinity for free surface flows - Part II: Numerical scheme and validation

Members: M-O. Bristeau, L. Boittin, A. Mangeney, J. Sainte-Marie, F. Bouchut

In this paper, we propose a numerical scheme for the layer-averaged Euler with variable density and the Navier-Stokes-Fourier systems presented in part I.These systems model hydrostatic free surface flows with density variations. We show that the finite volume scheme presented is well balanced with regards to the steady state of the lake at rest and preserves the positivity of the water height. A maximum principle on the density is also proved as well as a discrete entropy inequality in the case of the Euler system with variable density. Some numerical validations are finally shown with comparisons to 3D analytical solutions and experiments.

7.3. Functional analysis of PDE models in Fluid Mechanics

7.3.1. On the rigid-lid approximation of shallow water Bingham model

Member: J. Sainte-Marie
Coll.: Bilal Al Taki, Khawla Msheik

The paper [17] discusses the well posedness of an initial value problem describing the motion of a Bingham fluid in a basin with a degenerate bottom topography. A physical interpretation of such motion is discussed. The system governing such motion is obtained from the Shallow Water-Bingham models in the regime where the Froude number degenerates, i.e taking the limit of such equations as the Froude number tends to zero. Since we are considering equations with degenerate coefficients, then we shall work with weighted Sobolev spaces in order to establish the existence of a weak solution. In order to overcome the difficulty of the discontinuity in Bingham's constitutive law, we follow a similar approach to that introduced in [G. Duvaut and J.-L. Lions, Springer-Verlag, 1976]. We study also the behavior of this solution when the yield limit vanishes. Finally, a numerical scheme for the system in 1D is furnished.

7.3.2. Global bmo-1($\mathbb{R}^N$) radially symmetric solution for compressible Navier-Stokes equations with initial density in $\mathbb{L}^\infty(\mathbb{R}^N)$

Member: B. Haspot

In [24], we investigate the question of the existence of global weak solutionfor the compressible Navier Stokes equations provided that the initial momentum belongs to bmo-1($\mathbb{R}^N$) with $N = 2, 3$ and is radially symmetric. We prove then a equivalent of the so-called Koch-Tataru theorem for the compressible Navier-Stokes equations. In addition we assume that the initial density is only bounded in $\mathbb{L}^\infty(\mathbb{R}^N)$, it allows us in particular to consider initial density admitting shocks. Furthermore we show that if the coupling between the density and the velocity is sufficiently strong,then the initial density which admits initially shocks is instantaneously regularizinginasmuch as the density becomes Lipschitz. To finish we prove the global existence of strong solution for large initial data pro-vided that the initial data are radially symmetric and sufficiently regular in dimension $N = 2, 3$ for $\gamma-$law pressure.

7.3.3. New effective pressure and existence of global strong solution for compressible Navier-Stokes equations with general viscosity coefficient in one dimension

Member: Boris Haspot
Coll.: Cosmin Burtea

In this paper we prove the existence of global strong solution for the Navier-Stokesequations with general degenerate viscosity coefficients. The cornerstone of the proofis the introduction of a new effective pressure which allows to obtain an Oleinik-typeestimate for the so called effective velocity. In our proof we make use of additionalregularizing effects on the velocity which requires to extend the technics developedby Hoff for the constant viscosity case.

7.4. Assessments of models by means of experimental data and assimilation

7.4.1. Metamodeling corrected by observational data

Members: V. Mallet, J. Hammond

An air quality model at urban scale computes the air pollutant concentrations at street resolution based on various emissions, meteorology, imported pollution and city geometry. Because of the computational cost of such model, we previously designed a metamodel using dimension reduction and statistical emulation, and then corrected this metamodel with observational data. Novel work was dedicated to the error modeling for a more balanced integration of the observations. The work was also applied to air quality simulation over Paris using several months of data.

7.4.2. Metamodeling of a complete air quality simulation chain

Members: A. Lesieur, V. Mallet
Coll.: Ruiwei Chen

With the objective of uncertainty quantification, we worked on the generation of a metamodel for the simulation of urban air quality, using a complete simulation chain including dynamic traffic assignment, the computation of air pollutant emissions and the dispersion of the pollutant in a city. The traffic model and the dispersion model are computationally costly and operate in high dimension. We employed dimension reduction, and coupled it with Kriging in order to build a metamodel for the complete simulation chain.

7.4.3. Artificial neural networks for the modeling of air pollution

Member: V. Mallet

Air quality simulations at national, continental or global scales are subject to large uncertainties which are typically mitigated by data assimilation techniques. Another approach to improve the forecasts is to design an error model, learning from historical discrepancies between simulations and observations. Such a model was built using an artificial neural network trained with many meteorological and geographical data. Further studies showed that the technique could successfully generate not only an error model (to improve pre-existing simulations), but also a complete model (without the need for pre-existing simulations) whose forecasts are more accurate than those of traditional models.

7.4.4. Uncertainty quantification in atmospheric dispersion of radionuclides

Members: V. Mallet,
Coll.: Irène Korsakissok

In collaboration with IRSN (Institute of Radiation Protection and Nuclear Safety), we investigated the uncertainties of the atmospheric-dispersion forecasts that are used during an accidental release of radionuclides such as the Fukushima disaster. These forecasts are subject to considerable uncertainties which originate from inaccurate weather forecasts, poorly known source term and modeling shortcomings. In order to quantify the uncertainties, we designed a metamodel and carried out the calibration of the metamodel input distributions using Markov chain Monte Carlo.

7.4.5. Meta-modeling for urban noise mapping

Members: A. Lesieur, V. Mallet
Coll.: Pierre Aumond, Arnaud Can

Noise computing software can require several hours to produce a map over an urban center for a given set of input data. This computational cost makes the models unsuitable for applications like uncertainty quantification or data assimilation where thousands of simulations, or more, can be required. One solution is to replace the physical model with a meta-model which is very fast and yet fairly reproduces the results of the physical model. The strategy is first to reduce the dimension of both inputs and outputs of the physical model, which leads to a reduced model. This reduced model is then replaced by a statistical emulator. The emulator is trained with calls to the reduced model for a set of chosen inputs. The emulator relies on the interpolation between the training output values.

7.4.6. Data assimilation for urban noise maps generated with a meta-model

Members: A. Lesieur, V. Mallet
Coll.: Pierre Aumond, Arnaud Can

In an urban area, it is increasingly common to have access to both a simulated noise map and a sensor network. A data assimilation algorithm is developed to combine data from both a noise map simulator and a network of acoustic sensors. One-hour noise maps are generated with a meta-model fed with hourly traffic and weather data. The data assimilation algorithm merges the simulated map with the sound level measurements into an improved noise map. The performance of this method relies on the accuracy of the meta-model, the input parameters selection and the model of the error covariance that describes how the errors of the simulated sound levels are correlated in space. The performance of the data assimilation is obtained with a leave-one-out cross-validation method.

7.4.7. Uncertainty quantification in wildland fire propagation

Members: F. Allaire, V. Mallet
Coll.: Jean-Baptiste Filippi

We worked further on the Monte Carlo simulation of wildland fires. We calibrated the input distributions that represent the uncertainties in the inputs of our fire spread predictions by using the observations of the final contours for a number of fire cases. We used a new metric to measure the dissimilarity between two burned sufaces that relies on the Wasserstein distance. We designed a metamodel and carried out the calibration of the model input distributions using Markov chain Monte Carlo.

7.4.8. A non-intrusive reduced order data assimilation method applied to the monitoring of urban flows

Member: J. Hammond
Coll.: R. Chakir

In [13], we investigate a variational data assimilation method to rapidly estimate urban pollutant concentration around an area of interest using measurement data and CFD based models in a non-intrusive and computationally efficient manner. In case studies presented here, we used a sample of solutions from a dispersion model with varying meteorological conditions and pollution emissions to build a Reduced Basis approximation space and combine it with concentration observations. The method allows to correct for unmodeled physics, while significantly reducing online computational time.

7.5. Software Devlopments

Members: C., A. El Baz, J. Sainte-Marie

Several improvements of FreshKiss3D software have been made:

1. the code can now be used on RPM-based systems (Fedora, Redhat) thanks to a user contribution (Julien Jerphanion);
2. the conda-based installation steps are now automatically tested by means of a continuous integration process performed on Linux and Mac virtual machines provided by the Inria platform https://ci.inria.fr/;
3. third-party libraries have been updated so as to benefit from their very latest features;
4. software quality is now monitored by static analysis via the Inria SonarQube platform (https://sonarqube.inria.fr/);
5. code optimization including parallelization with MPI is currently under development.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- A contract between Institut Carnot SMILES and the corporation GTT involving (Jacques Sainte-Marie, Cindy Guichard, Yohan Penel and Julien Salomon) of 78 k€ has been approved. The aim is to improve the numerical modelling tool to simulate gas flows in the insulation spaces of LNG tankers. A Ph. D. thesis should start next year on that topic.

8.2. Bilateral Grants with Industry

- The ANR project Hyflo-Eflu relies on a collaboration with the company "HydroTube Energie", that ended in December has given rise to deep collaboration with Hydotube Energy.
- The ANR project Firecaster supports the Ph. D. Thesis of F. Allaire on the developpment of a fire decision support system at the national scale to estimate upcoming fire risk. The collaborations are CERFACS and CNRM (recherche de Météo-France).
- The ANR project Cense supports the Ph. D. Thesis of A. Lesieur on the developpment of a new methodology for the production of more realistic noise maps. The industrial collaborations include:
  - Bouygues Énergies & Services
  - Wi6Labs
  - Bruitparif (association)

  We refer to

  https://cense.ifsttar.fr/partenaires/entreprises-privees/

  https://cense.ifsttar.fr/partenaires/associations/

  for more details.

8.3. Other collaborations with Industry

- On the public operational side, ANGE team works with IRSN and its Modelling Bureau environmental transfers for the study of consequences of accidents (BMCA). This collaboration led to Bao Le's thesis. For more details, we refer to

  https://www.irsn.fr/FR/Larecherche/Organisation/equipes/radioprotection-homme/BMCA/Pages/bureau-modelisation-transferts-environnement-etude-consequences-accidents.aspx
- On the corporate side, ANGE collaborates with NUMTECH for pollution modelling. We mention also the long term collaboration with Ambiciti- co-founded by V. Mallet. In particular, the newspaper Ouest-France uses Ambiciti's air quality forecasts since the summer, so with the algorithms that come from ANGE.

  https://www.ouest-france.fr/meteo/

- Y. Penel obtained a partial support from EDF for the organisation of the project SGN-Num at CEMRACS 2019.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR MFG (2016-2021)

Participant: Julien Salomon.

Project acronym: MFG

Project title: Mean Field Games

Coordinator: Sébastien Boyaval (LHSV/ENPC)

Funding: 299 160 euros.

Mean field game theory (MFG) is a new and active field of mathematics, which analyses the dynamics of a very large number of agents. Introduced about ten years ago, MFG models have been used in different fields: economics, finance, social sciences, engineering,... MFG theory is at the intersection of mean field theory, mathematical game theory, optimal control, stochastic analysis, variation calculation, partial differential equations and scientific calculation. Drawing on an internationally recognized French team on the subject, the project seeks to obtain major contributions in 4 main directions: the "medium field" aspect (i.e., how to obtain macroscopic models from microscopic models); the analysis of new MFG systems; their numerical analysis; the development of new applications. In this period of rapid expansion of MFG models, the project seeks to foster French leadership in the field and attract new researchers from related fields.

9.1.2. ANR INFAMIE (2015-2019)

Participant: Boris Haspot.

Program: ANR Défi de tous les savoirs (DS10) 2015

Project acronym: INFAMIE

Project title: INhomogeneous Flows : Asymptotic Models and Interfaces Evolution

Coordinator: Raphaël Danchin (Univ. Paris-Est)

Funding: 232 960 euros.

Our project aims at a better mathematical understanding of several models for the evolution of inhomogeneous flows. Through three main lines of research (see below), we will pursue a twofold final objective. First, we want to develop the current theory of regular solutions for several equations for the evolution of fluids, proposing a new approach and developing tools that are likely to be efficient in various areas of PDEs. Second, for a few selected concrete systems that describe flows in the earth environment or in astrophysics, we wish to use this general approach to extract as much information as possible concerning the qualitative behavior of the solutions.

9.1.3. ANR SEDIFLO (2015-2019)

Participants: Emmanuel Audusse, Martin Parisot.

Program: ANR Défi 1 “Gestion sobre des ressources et adaptation au changement climatique” (JCJC)

Project acronym: SEDIFLO

Project title: Modelling and simulation of solid transport in rivers

Coordinator: Sébastien Boyaval (LHSV/ENPC)

Based on recent theoretical and experimental results, this project is aimed at modelling transport of sediments within rivers. It will rely on innovations from the point of view of rheology as well as advanced mathematical tools (asymptotic model reduction, PDE discretisation).

9.1.4. ANR Hyflo-Eflu (2016-2019)

Participants: Jérémy Ledoux, Martin Parisot, Jacques Sainte-Marie, Julien Salomon.

ANR project call: Energies marines renouvelables

Project acronym: Hyflo-Eflu

Project title: Hydroliennes flottantes et énergie fluviale

Coordinator: Julien Salomon

The project is a collaboration between the Inria-team ANGE, specialist of free surface flow and optimisation, and the industrial developers of the turbine, HydroTube Energie. The objective of the project HyFlo-EFlu is to deliver a numerical software able to simulate the dynamic of a floating water turbine in real context. For the academic partner, the main challenge is in the simulation of the floating structure at the scale of the river, and the modelling of the vertical and horisontal axis turbine. For the industrial partner, the objective is the validation of the stability of the structure and the performance in term of energy production.

9.1.5. ANR CHARMS (2016-2020)

Participant: Cindy Guichard.

ANR project call: Transformations et inter-conversions énergétiques

Project acronym: CHARMS

Project title: Modèles de réservoirs quantitatifs pour les systèmes hydrothermaux complexes

Coordinator: Simon Lopez (BRGM)

Funding: 73k euros for LJLL (in 767k euros for the whole project)

CHARMS ANR project is focused on the mathematical methods and software tools dedicated to the simulation of the physical models issued from geothermal engineering. The final objective is the achievement of a highly parallel code, validated on realistic cases.

9.1.6. GdR EGRIN (2017–2021)

Participants: Emmanuel Audusse, Bernard Di Martino, Nicole Goutal, Cindy Guichard, Anne Mangeney, Martin Parisot, Jacques Sainte-Marie.

EGRIN stands for Gravity-driven flows and natural hazards. J. Sainte-Marie is the head of the scientific committee of this CNRS research group and A. Mangeney is a member of the committee. Other members of the team involved in the project are local correspondents. The scientific goals of this project are the modelling, analysis and simulation of complex fluids by means of reduced-complexity models in the framework of geophysical flows.

9.1.7. ANR FireCaster (2017-2020)

Participants: Frédéric Allaire, Vivien Mallet.

ANR project call: DS0104

Project acronym: FireCaster

Project title: Plateforme de prévision incendie et de réponse d'urgence

Coordinator: Jean-Baptiste Filippi (Univ. Corse)

Funding: 442k euros

The goal of the FireCaster project is to prototype a fire decision support system at the national scale to estimate upcoming fire risk (H+24 to H+48) and in case of crisis, to predict fire front position and local pollution (H+1 to H+12).

9.1.8. ANR CENSE (2017-2020)

Participants: Antoine Lesieur, Vivien Mallet.

ANR project call: DS0601

Project acronym: CENSE

Project title: Caractérisation des environnements sonores urbains : vers une approche globale associant données libres, mesures et modélisations

Coordinator: Judicaël Picaut (IFSTTAR)

Funding: 856k euros

The CENSE project aims at proposing a new methodology for the production of more realistic noise maps, based on an assimilation of simulated and measured data through a dense network of low-cost sensors.

9.1.9. ANR RAVEX (2017-2020)

Participant: Anne Mangeney.

ANR project call: DS0106

Project acronym: RAVEX

Project title: Développement d'une approche intégrée pour la réduction des Risques Associés au Volcanisme EXplosif, de la recherche sur l'aléa aux outils de gestion de crise : le cas de la Martinique

Coordinator: Olivier Roche (IRD)

Funding: 619k euros

9.1.10. ANR CINE-PARA (2015-2019)

Participant: Julien Salomon.

ANR project call: DS0708

Project acronym: CINE-PARA

Project title: Méthodes de parallélisation pour cinétiques complexes

Coordinator: Yvon Maday (LJLL)

9.1.11. PGMO Project ORACLE (2019-2021)

Participant: Julien Salomon.

PGMO Call

Project acronym: Oracle

Project title: Optimal Resource Allocation in micro-organisms under Changing Environment

Coordinator: Térence Bayen

9.2. European Initiatives

9.2.1.

Participants: Martin Parisot, Yohan Penel, Jacques Sainte-Marie.

CNRS PICS NHML (2017-2019)

Program: CNRS PICS (projet international de collaboration scientifique)

Project acronym: NHML

Project title: non-hydrostatic multilayer models

Duration: 01/17-12/19

Coordinator: Yohan Penel (Inria)

Other partners: IMUS (Sevilla, Spain)

Other Participants: Enrique Fernández-Nieto (Sevilla), Tomas Morales de Luna (Cordoba)

Funding: 12k euros

Abstract: This collaboration aims at designing a hierarchy of multilayer models with a non-hydrostatic pressure as a discretisation along the vertical axis of the Euler equations. The hierarchy relies on the degree of approximation of the variables discretised with a Discontinuous Galerkin method for the vertical direction. These innovative models will imply a theoretical study and the development of numerical tools in dimensions 1 and 2 before the modelling of other physical phenomena (viscosity effects, ...).

9.3. International Initiatives

9.3.1. Inria International Partners

Three long-term collaborations with foreign colleagues have to be mentioned:

- **Spain** - A collaboration with spanish researchers has been initiated in 2016 to derive accurate models and effecient algorithms for free surface flows including non-hydrostatic effects.
- **Germany** A collaboration with researchers from the University of Constance is in progress about domain decomposition and identifaction algorithms (G. Ciaramella, S. Volkwein). The internship (Masterarbeit) of S. Buchwald has been co-supervised by J. Salomon.
- **Hong-Kong, Switzerland** A collaboration with F. Kwok and M. Gander on time parallelization for assimilation algorithm is in progress. A first paper has been submitted in october.

9.4. International Research Visitors

- Y. Penel made two two-week stay (March, October) at the university of Sevilla (Spain) to collaborate with E. Fernández-Nieto.
- M. Parisot spent two months (October, November) at the university of Aachen (Germany) to collaborate with S. Noelle.

9.4.1. Visits of International Scientists

- G. Ciaramella (Constance University) visited J. Salomon (20.05-23.05) to work on algorithms related to design of experiments and identification.
- F. Kwok (Baptist University, Hong-Kong) visited J. Salomon (8.07-11.07) to work on algorithms related to time parallelization for identification.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

Some members of ANGE were in the organizing committees of scientific events.

- J. Sainte-Marie took part of the organization of the 7th EGRIN summer school that took place at Le Lioran from 24th to 27th of June and that gathered 40 researchers [0].
- J. Sainte-Marie belongs to the "Groupe de travail Recherche et développement durable" at french ministry of research.

[0] https://indico.math.cnrs.fr/event/4391/

- Y. Penel and N. Aguillon organize the monthly ANGE seminar [0]. The speakers were : F. Allaire, M. Peruzzetto (IPGP), L. Boittin, S. Junca (Nice), H. Martin, A. Leprevost, M. Parisot, C. Donadello (Besançon), A. Lesieur, M. Rigal, M. Ancellin (Sorbonne Univeristé).
- J. Salomon co-organises the LJLL-Inria meetings [0] (twice a month).
- M. Parisot and J. Salomon organized the conference "EmrSim2019" [0] from 2nd to 4th July.
- Y. Penel and J. Sainte-Marie organized the session "challenges in geosciences" in the framework of Tarantola days [0] from 6 to 7 June.
- C. Guichard co-organized the "Journée d'Accueil des nouveaux recrutés en Mathématiques", at IHP [0]
- E. Godlewski is the president of the CFEM (commission française pour l'enseignement des mathématiques).
- E. Godlewski took part of the organization of the 50th birthday of the "laboratoire J-L. Lions" by giving a talk "Le rayonnement de l'école Lions. Vers une politique scientifique".

10.1.2. Journal

The summary of the reviewing activities of the team is given in the next table.

Member	Journals
JS	CRAS, SIAM SISC, JCP
CG	Computers and Mathematics with Applications, Numerische Mathematik
EG	M2AN, JCP, AIDE, JEP
MP	Journal of Hydraulic Research, DCDS, JCP, Journal of Computational Science
VM	JCP
JSM	M2AN, JCP, Applied Mathematical Modeling

10.1.3. Invited Talks

[0] https://team.inria.fr/ange/gdt-slides/
[0] https://project.inria.fr/rencontresljll/fr/
[0] https://emrsim2019.sciencesconf.org/
[0] https://ange-geophysics.sciencesconf.org/
[0] http://postes.smai.emath.fr/apres/accueil/

Member	Conference	Location	Date
JS	Colloque d'ouverture 50 ans du LJLL	Roscoff	4-8/3
NA	Séminaire Centrale Paris	Saclay	05/12
NA	journée de clôture MANON	Paris	March
NA	journée Tarantola : défis en géosciences	Paris	June
VD	CEMRACS	Marseille	13/08
VD	Congrès SMAI	Guidel	16/05
VD	EGRIN	Le Liorant	25/06
CG	journée Tarantola : défis en géosciences	Paris	June
FA	GdT ANGE	Inria Paris	23/01
YP	séminaire de l'équipe ACSIOM	Montpellier	19/02
YP	séminaire d'analyse appliquée	Amiens	14/01
YP	Congrès SMAI	Guidel	16/05
YP	Ecole EGRIN	Le Liorant	24/06
YP	ICIAM	Valencia (Spain)	16/07
MP	séminaire de l'IRMAR	Rennes	28/02
MP	Cours sur les modèles dispersif	Aachen	22-24/10
MP	Séminaire d'équipe d'analyse numérique	Aachen	31/10
VM	Nommer, modéliser, représenter le hasard dans les arts et les sciences	Arras	15/03
VM	AI for Climate	Paris	20/09
VM	Séminaire du laboratoire Jacques-Louis Lions	Paris	25/10
VM	AI, data and models for sustainable development, COP25	Madrid	3/12
JSM	"Geophysical flows - From hydrostatic to non-hydrostatic models"	Malaga	22/09

10.1.4. Research Administration

- J. Sainte-Marie is now "Adjoint au Directeur Scientifique" of Inria, in charge of the topic "Sciences de la planète, de l'environnement et de l'énergie".
- J. Salomon belongs to Inria's "Commission des emplois scientifiques".

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

10.2.2. Supervision and Jury

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

Table 3. Teaching (1/2)

Member	Topic	Duration	Level	Institution	Type (CM,TD,TP)
JS	Méthodes numériques pour des modèles incluants des EDP	45	M2	Université d'Abomey-Calavi, Bénin	CM
JS	Méthodes numériques pour des modèles incluants des EDP	45	M2	Univ. Paris-Dauphine	CM
VD	Mathématiques appliquées	32	L3	Polytech Sorbonne	TP
VD	Traitement numérique	14+22	M1	Polytech Sorbonne	CM+TP
VD	Initiation Latex	4	L3	Polytech Sorbonne	CM
LL	Méthodes numériques pour les équations différentielles	24	L3	Sorbonne Université	TP
LL	Mathématiques pour les études scientifiques I	102	L1	Sorbonne Université	TD
LL	Mathématiques pour DU RESPE S2	12	L0	Sorbonne Université	CM+TD
LL	Mathématiques pour DU RESPE S1	10	L0	Sorbonne Université	CM+TD
NA	Topologie et calcul différentielle	111	L2	Sorbonne Université	CM+TD
NA	Projet	15	L3	Sorbonne Université	TD
CG	Fondements des méthodes numériques	58	M1	SU	TP
CG	Méthodes numériques	31	M2	SU	CM+TD
CG	Co responsable de la majeure Ingénierie Mathématiques pour l'Entreprise		M2	SU	
BDM	Analyse et TP Python Sage algèbre et analyse	72	L2	Univ. Corse	C+TD+TP
BDM	Analyse numérique matricielle	54	L3	Univ. Corse	C+TD+TP
EG	Modèles hyperboliques d'écoulements complexes dans le domaine de l'énergie	10	M2	SU	CM
NB	Analyse2	15	L1	Paris13	TD
NB	Analayse3: Intégrales et séries	30	L2	Paris13	TD

Table 4. Teaching (2/2)

Member	Topic	Duration	Level	Institution	Type (CM,TD,TP)	
JS	VM	Simulation de la dispersion atmosphérique	3	M2	ENPC	TP
VM	Assimilation de données pour la géophysique	7	M2	ENPC	TP	
JH	Méthodes numériques pour les équations différentielles	25	L3	Sorbonne Université	TP	
JH	Analyse Appliquée	40	L3	Sorbonne Université	TD	
EA	Equations différentielles	30	L3	Paris13	TD	
EA	Optimisation différentiable	30	M1	Paris13	TD	
EA	Mécanique des fluides	18	M1	Paris13	CM	
EA	Calcul scientifique	30	L2	Paris13	CM+TD	
YP	Equations différentielles	15	L2	Paris 13	CM+TD	

- J. Salomon is a member of the "Commission Mentorat" in view of introducing a mentoring procedure in Inria-Paris,
- Y. Penel and C. Guichard belong to the "Groupe de travail Developpement Durable d'Inria Paris",
- J. Sainte-Marie took part of the committe MakeSenSe, that produced the report [14].

10.3.2. Education

- Congrès MATh.en.JEANS (CentraleSupélec, Modélisation de certains écoulements),
- E. Godlewski and J. Salomon dealt with secondary-school pupils in the framework of observation training course at LJLL and Inria, respectively,
- N. Aguillon was in the organizing committe of "Mathematic Park",
- N. Aguillon was in the organizing committe of "Cycle SMAI-Musée des arts et métiers",
- J. Salomon co-organized the stand "Logiciel Libre" at the "Salon Culture et Jeux mathématiques, 2019".

10.3.3. Interventions

- J. Sainte-Marie organized a round-table discussion about "Mathématiques et environnement" at the "Forum Emploi-Math 2019" (14/11).
- J. Sainte-Marie organized a debate "la sobriété numérique" at "cité des Sciences et de l'industrie" (15/11).

11. Bibliography

Table 5. Supervision of thesis and intership (1/2)

Member	Other advisors	Type	Student	Institution	Period	Title
JS		PhD	Sebastian Reyes-Riffo	Paris-Dauphine	2016-2019	Méthodes numériques pour les énergies marines renouvelables
JS, VM		PhD	Antoine Lesieur	Inria	2017-2020	Estimation d'état et modélisation inverse appliquées à la pollution sonore en milieu urbain
JS		PhD	Nadia Jbili	Paris-Dauphine	2016-2019	Contrôle optimal pour la résonance magnétique nucléaire
JS, JSM		PhD	Liudi Lu	Inria	2018-2021	Approches Lagrangiennes pour la modélisation et l'optimisation du couplage hydrodynamique-photosynthèse
JSM, VM		PhD	Frédéric Allaire	Inria	2017-2020	Quantification du risque incendie par méta-modélisation de la propagation de feux de forêt
YP, CG, JSM		PhD	Virgile Dubos	SU	2017-2020	Numerical methods for the elliptic/parabolic parts of non-hydrostatic fluid models
NA, EA, MP		PhD	Nelly BOULOS	Paris 13	2018-2021	Modélisation et simulation numérique de la dynamique d'un acquifère érodable
VM	Isabelle Herlin	PhD	Ngoc Bao Tran Le	Inria	2016-2019	Uncertainty quantification based on model reduction for atmospheric dispersion
VM		Post Doc	Janelle Hammond	Inria	2017-2020	Uncertainty quantification, metamodeling and data assimilation applied to urban air quality
EA,MP,JSM		PhD	Léa Boittin	inria	2016-2019	Modelling, analysis and efficient numerical resolution for erosion processes

Table 6. Supervision of thesis and intership (2/2)

NA		Stage M2	Mathieu Rigal	UPMC	2019	Schéma exact sur les chocs isolés pour les équations de St Venant avec topographie
NA, JSM	Nathalie Ayi	PhD	Mathieu Rigal	UPMC	2019-2022	Low Froude regime and dispersive effects in kinetic formulations
BDM, JSM		Stage M2	Chourouk El Hassanieh	Inria	04-08/2019	Multilayer Shallow Water Equations: Derivation, Properties, and Well Posedness
BDM, JSM	Samer Israwi	PhD	Chourouk El Hassanieh	Inria	2019-2022	Mathematical and numerical analysis of some dispersive models in fluids mechanics

Major publications by the team in recent years

[1] E. Audusse, M.-O. Bristeau, M. Pelanti, J. Sainte-Marie.*Approximation of the hydrostatic Navier-Stokes system for density stratified flows by a multilayer model. Kinetic interpretation and numerical validation*, in "J. Comput. Phys.", 2011, vol. 230, p. 3453-3478, http://dx.doi.org/10.1016/j.jcp.2011.01.042

[2] E. Audusse, M.-O. Bristeau, B. Perthame, J. Sainte-Marie.*A multilayer Saint-Venant system with mass exchanges for Shallow Water flows. Derivation and numerical validation*, in "ESAIM Math. Model. Numer. Anal.", 2011, vol. 45, p. 169-200, http://dx.doi.org/10.1051/m2an/2010036

[3] M.-O. Bristeau, A. Mangeney, J. Sainte-Marie, N. Seguin.*An energy-consistent depth-averaged Euler system: derivation and properties*, in "Discrete and Continuous Dynamical Systems - Series B", 2015, vol. 20, n^{o} 4, 28

[4] J. Sainte-Marie.*Vertically averaged models for the free surface Euler system. Derivation and kinetic interpretation*, in "Math. Models Methods Appl. Sci. (M3AS)", 2011, vol. 21, n^{o} 3, p. 459-490, http://dx.doi.org/10.1142/S0218202511005118

Publications of the year

Doctoral Dissertations and Habilitation Theses

[5] L. Boittin.*Modeling, analysis and simulation of two geophysical flows : Sediment transport and variable density flows*, Sorbonne Université, April 2019, https://tel.archives-ouvertes.fr/tel-02126695

[6] S. Riffo.*Mathematical methods for marine energy extraction*, Paris Sciences et Lettres ; Paris IX Dauphine, November 2019, https://hal.archives-ouvertes.fr/tel-02446450

Articles in International Peer-Reviewed Journal

Table 7. Participation to committees

Initiales	Mois	Type (PhD, HdR)	Rôle (rapp., prés., mbre)	Prénom Nom candidat	Etablissement	Titre
JS	Novembre	PhD	rapporteur	Pierre-Yves Wuillaume	Ecole Centrale de Nantes	Simulation numérique des opérations d'installation pourles fermes d'éoliennes offshore
NA	juillet	PhD	examinatrice	Mostafa Kadiri	Normandie Université	Optimisation de forme et applications aux ouvrages hydrauliques
CG	mars	PhD	examinatrice	Patrik Daniel	SU et Inria	Adaptive hp-finite elements with guaranteed errorcontraction and inexact multilevel solvers
CG	septembre	PhD	examinatrice	Pierrick Quemar	Paris 13 et EDF	MODELING AND NUMERICAL ANALYSIS OF FREE SURFACE FLOWS
CG	novembre	PhD	invitée	Karine LAURENT	IFPEN et Univ. Paris Saclay	Étude de nouveaux schémas numériques pour la simulation des écoulements à rapport de mobilités défavorable dans un contexte EOR
JSM	novembre	PhD	examinateur	Sebastian Reyes Riffo	PSL	Méthodes mathématiques pour l'extraction d'énergie marine
JSM	décembre	PhD	rapporteur	H. Boukili	univ. Marseille	Schémas de simulation d'un modèle à trois phases immiscibles pour application à l'explosion vapeur
EG	novembre	thèse (PhD)	présidente	Jean-François Abadie	SU	Estimations fiables d'une fonction et de ses dérivées & Étude théorique et numérique d'un problème de "shape from shading"
YP	novembre	PhD	rapporteur	Victor Osores	Univ. Concepcion, Chile	On multilayer shallow water systems for polydisperse sedimentation

[7] F. ALLAIRE, J.-B. FILIPPI, V. MALLET.*Generation and evaluation of an ensemble of wildland fire simulations*, in "International Journal of Wildland Fire", 2020, forthcoming [*DOI :* 10.1071/WF19073], https://hal.inria.fr/hal-02447187

[8] S. ALLGEYER, M.-O. BRISTEAU, D. FROGER, R. HAMOUDA, V. JAUZEIN, A. MANGENEY, J. SAINTE-MARIE, F. SOUILLÉ, M. VALLÉE.*Numerical approximation of the 3d hydrostatic Navier-Stokes system with free surface*, in "ESAIM: Mathematical Modelling and Numerical Analysis", August 2019, vol. 53, n^o^ 6, https://arxiv.org/abs/1709.06267 , https://hal.inria.fr/hal-01393147

[9] S. DELLACHERIE, G. FACCANONI, B. GREC, Y. PENEL.*Accurate steam-water equation of state for two-phase flow LMNC model with phase transition*, in "Applied Mathematical Modelling", 2019, vol. 65, p. 207-233 [*DOI :* 10.1016/J.APM.2018.07.028], https://hal.archives-ouvertes.fr/hal-01111730

[10] P.-O. LAMARE, N. AGUILLON, J. SAINTE-MARIE, J. GRENIER, H. BONNEFOND, O. BERNARD.*Gradient-based optimization of a rotating algal biofilm process*, in "Automatica", July 2019, vol. 105, p. 80-88 [*DOI :* 10.1016/J.AUTOMATICA.2019.02.043], https://hal.inria.fr/hal-02422853

[11] F. LEMARIÉ, H. BURCHARD, L. DEBREU, K. KLINGBEIL, J. SAINTE-MARIE.*Advancing dynamical cores of oceanic models across all scales*, in "Bulletin of the American Meteorological Society", 2019, vol. 100, p. ES109–ES115 [*DOI :* 10.1175/BAMS-D-18-0303.1], https://hal.inria.fr/hal-01939057

[12] M. PARISOT.*Entropy-satisfying scheme for a hierarchy of dispersive reduced models of free surface flow*, in "International Journal for Numerical Methods in Fluids", August 2019 [*DOI :* 10.1002/FLD.4766], https://hal.inria.fr/hal-01242128

National Conferences with Proceeding

[13] J. K. HAMMOND, R. CHAKIR.*A non-intrusive reduced order data assimilation method applied to the monitoring of urban flows*, in "CSMA2019 - 14ème Colloque National en Calcul des Structures", Presqu'île de Giens, France, January 2019, 7, CSMA2019 - 14ème Colloque National en Calcul des Structures, Presqu'île de Giens , FRANCE, 13-/05/2019 - 17/05/2019, https://hal.archives-ouvertes.fr/hal-02186298

Research Reports

[14] F. BERTHOUD, P. GUITTON, L. LEFÈVRE, S. QUINTON, A. ROUSSEAU, J. SAINTE-MARIE, C. SERRANO, J.-B. STEFANI, P. STURM, E. TANNIER.*Sciences, Environnements et Sociétés : Rapport long du groupe de travail MakeSEnS d'Inria*, Inria, October 2019, https://hal.inria.fr/hal-02340948

[15] P.-O. LAMARE, N. AGUILLON, J. SAINTE-MARIE, J. GRENIER, H. BONNEFOND, O. BERNARD.*Gradient-based optimization of a rotating algal biofilm process*, Inria - Sophia antipolis, January 2019, n^o^ RR-9250, https://hal.inria.fr/hal-01990002

Other Publications

[16] N. AGUILLON, P.-A. GUIHENEUF.*Dynamical behavior of a nondiffusive scheme for the advection equation*, October 2019, https://arxiv.org/abs/1910.03456 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02304798

[17] B. AL TAKI, K. MSHEIK, J. SAINTE-MARIE.*On the rigid-lid approximation of shallow water bingham model*, February 2019, working paper or preprint, https://hal.inria.fr/hal-02052055

[18] N. AÏSSIOUENE, M.-O. BRISTEAU, E. GODLEWSKI, A. MANGENEY, C. PARÉS, J. SAINTE-MARIE.*A two-dimensional method for a family of dispersive shallow water model*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01632522

[19] C. BURTEA, B. HASPOT.*New effective pressure and existence of global strong solution for compressible Navier-Stokes equations with general viscosity coefficient in one dimension*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02006843

[20] R. EYMARD, C. GUICHARD, X. LHÉBRARD.*Convergence of numerical schemes for a conservation equation with convection and degenerate diffusion*, August 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01939144

[21] M. J. GANDER, F. KWOK, J. SALOMON.*PARAOPT: A parareal algorithm for optimality systems*, October 2019, https://arxiv.org/abs/1911.01686 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02346535

[22] E. GODLEWSKI, M. PARISOT, J. SAINTE-MARIE, F. WAHL.*Congested shallow water model: on floating body*, September 2019, working paper or preprint, https://hal.inria.fr/hal-01871708

[23] J. K. HAMMOND, R. CHEN, V. MALLET.*Meta-modeling of a simulation chain for urban air quality*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02429687

[24] B. HASPOT.*Global bmo -1 (R N) radially symmetric solution for compressible Navier-Stokes equations with initial density in L ∞ (R N)*, January 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01976953

[25] N. JBILI, K. HAMRAOUI, S. J. GLASER, J. SALOMON, D. SUGNY.*Optimal periodic control of spin systems: Application to the maximization of the signal to noise ratio per unit time*, March 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02066372

References in notes

[26] E. AUDUSSE.*A multilayer Saint-Venant model : Derivation and numerical validation*, in "Discrete Contin. Dyn. Syst. Ser. B", 2005, vol. 5, n^o 2, p. 189-214

[27] E. AUDUSSE, M.-O. BRISTEAU, M. PELANTI, J. SAINTE-MARIE.*Approximation of the hydrostatic Navier-Stokes system for density stratified flows by a multilayer model. Kinetic interpretation and numerical validation*, in "J. Comput. Phys.", 2011, vol. 230, p. 3453-3478, http://dx.doi.org/10.1016/j.jcp.2011.01.042

[28] E. AUDUSSE, M.-O. BRISTEAU, B. PERTHAME, J. SAINTE-MARIE.*A multilayer Saint-Venant system with mass exchanges for Shallow Water flows. Derivation and numerical validation*, in "ESAIM Math. Model. Numer. Anal.", 2011, vol. 45, p. 169-200, http://dx.doi.org/10.1051/m2an/2010036

[29] F. BOUCHUT, V. ZEITLIN.*A robust well-balanced scheme for multi-layer shallow water equations*, in "Discrete Contin. Dyn. Syst. Ser. B", 2010, vol. 13, p. 739-758

[30] M. CASTRO, J. GARCÍA-RODRÍGUEZ, J. GONZÁLEZ-VIDA, J. MACÍAS, C. PARÉS, M. VÁZQUEZ-CENDÓN.*Numerical simulation of two-layer shallow water flows through channels with irregular geometry*, in "J. Comput. Phys.", 2004, vol. 195, n^o 1, p. 202–235

Project-Team ANTIQUE

Static Analysis by Abstract Interpretation

IN COLLABORATION WITH: Département d'Informatique de l'Ecole Normale Supérieure

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Paris

RESEARCH CENTER
Paris

THEME
Proofs and Verification

Table of contents

Project-Team ANTIQUE

Creation of the Team: 2014 January 01, updated into Project-Team: 2015 April 01

Keywords:

Computer Science and Digital Science:

A2. - Software
A2.1. - Programming Languages
A2.1.1. - Semantics of programming languages
A2.1.7. - Distributed programming
A2.1.12. - Dynamic languages
A2.2.1. - Static analysis
A2.3. - Embedded and cyber-physical systems
A2.3.1. - Embedded systems
A2.3.2. - Cyber-physical systems
A2.3.3. - Real-time systems
A2.4. - Formal method for verification, reliability, certification
A2.4.1. - Analysis
A2.4.2. - Model-checking
A2.4.3. - Proofs
A2.6.1. - Operating systems
A4.4. - Security of equipment and software
A4.5. - Formal methods for security

Other Research Topics and Application Domains:

B1.1. - Biology
B1.1.8. - Mathematical biology
B1.1.10. - Systems and synthetic biology
B5.2. - Design and manufacturing
B5.2.1. - Road vehicles
B5.2.2. - Railway
B5.2.3. - Aviation
B5.2.4. - Aerospace
B6.1. - Software industry
B6.1.1. - Software engineering
B6.1.2. - Software evolution, maintenance
B6.6. - Embedded systems

1. Team, Visitors, External Collaborators

Research Scientists

Xavier Rival [Team leader, Inria, Senior Researcher, HDR]
Vincent Danos [CNRS, Senior Researcher, HDR]
Cezara Drăgoi [Inria, Researcher]

Jérôme Feret [Inria, Researcher]
Caterina Urban [Inria, Researcher, from Feb 2019]

Technical Staff
Tie Cheng [Inria, Engineer, from May 2019]
Yves-Stan Le Cornec [Inria, Engineer, from Apr 2019]
Thierry Martinez [Inria, Engineer, from Mar 2019]

PhD Students
Andreea Beica [Inria, PhD Student, until Mar 2019]
Gaëlle Candel [Keymetrics, PhD Student]
Marc Chevalier [Ecole Normale Supérieure Paris, PhD Student]
Patricio Inzaghi [Inria, PhD Student, from Jun 2019]
Olivier Nicole [CEA, PhD Student, from Sep 2019]
Albin Salazar [Inria, PhD Student, from Sep 2019]

Visiting Scientist
Pierre Boutillier [Harvard Medical School]

Administrative Assistants
Chantal Chazelas [Inria, Administrative Assistant, until Jun 2019]
Nathalie Gaudechoux [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

Our group focuses on developing *automated* techniques to compute *semantic properties* of programs and other systems with a computational semantics in general. Such properties include (but are not limited to) important classes of correctness properties.

Verifying safety critical systems (such as avionics systems) is an important motivation to compute such properties. Indeed, a fault in an avionics system, such as a runtime error in the fly-by-wire command software, may cause an accident, with loss of life. As these systems are also very complex and are developed by large teams and maintained over long periods, their verification has become a crucial challenge. Safety critical systems are not limited to avionics: software runtime errors in cruise control management systems were recently blamed for causing *unintended acceleration* in certain Toyota models (the case was settled with a 1.2 billion dollars fine in March 2014, after years of investigation and several trials). Similarly, other transportation systems (railway), energy production systems (nuclear power plants, power grid management), medical systems (pacemakers, surgery and patient monitoring systems), and value transfers in decentralized systems (smart contracts), rely on complex software, which should be verified.

Beyond the field of embedded systems, other pieces of software may cause very significant harm in the case of bugs, as demonstrated by the Heartbleed security hole: due to a wrong protocol implementation, many websites could leak private information, over years.

An important example of semantic properties is the class of *safety* properties. A safety property typically specifies that some (undesirable) event will never occur, whatever the execution of the program that is considered. For instance, the absence of runtime error is a very important safety property. Other important classes of semantic properties include *liveness* properties (i.e., properties that specify that some desirable event will eventually occur) such as termination and *security* properties, such as the absence of information flows from private to public channels.

All these software semantic properties are *not decidable*, as can be shown by reduction to the halting problem. Therefore, there is no chance to develop any fully automatic technique able to decide, for any system, whether or not it satisfies some given semantic property.

The classic development techniques used in industry involve testing, which is not sound, as it only gives information about a usually limited test sample: even after successful test-based validation, situations that were untested may generate a problem. Furthermore, testing is costly in the long term, as it should be re-done whenever the system to verify is modified. Machine-assisted verification is another approach which verifies human specified properties. However, this approach also presents a very significant cost, as the annotations required to verify large industrial applications would be huge.

By contrast, the **antique** group focuses on the design of semantic analysis techniques that should be *sound* (i.e., compute semantic properties that are satisfied by all executions) and *automatic* (i.e., with no human interaction), although generally *incomplete* (i.e., not able to compute the best —in the sense of: most precise— semantic property). As a consequence of incompleteness, we may fail to verify a system that is actually correct. For instance, in the case of verification of absence of runtime error, the analysis may fail to validate a program, which is safe, and emit *false alarms* (that is reports that possibly dangerous operations were not proved safe), which need to be discharged manually. Even in this case, the analysis provides information about the alarm context, which may help disprove it manually or refine the analysis.

The methods developed by the **antique** group are not limited to the analysis of software. We also consider complex biological systems (such as models of signaling pathways, i.e. cascades of protein interactions, which enable signal communication among and within cells), described in higher level languages, and use abstraction techniques to reduce their combinatorial complexity and capture key properties so as to get a better insight in the underlying mechanisms of these systems.

3. Research Program

3.1. Semantics

Semantics plays a central role in verification since it always serves as a basis to express the properties of interest, that need to be verified, but also additional properties, required to prove the properties of interest, or which may make the design of static analysis easier.

For instance, if we aim for a static analysis that should prove the absence of runtime error in some class of programs, the concrete semantics should define properly what error states and non error states are, and how program executions step from a state to the next one. In the case of a language like C, this includes the behavior of floating point operations as defined in the IEEE 754 standard. When considering parallel programs, this includes a model of the scheduler, and a formalization of the memory model.

In addition to the properties that are required to express the proof of the property of interest, it may also be desirable that semantics describe program behaviors in a finer manner, so as to make static analyses easier to design. For instance, it is well known that, when a state property (such as the absence of runtime error) is valid, it can be established using only a state invariant (i.e., an invariant that ignores the order in which states are visited during program executions). Yet searching for trace invariants (i.e., that take into account some properties of program execution history) may make the static analysis significantly easier, as it will allow it to make finer case splits, directed by the history of program executions. To allow for such powerful static analyses, we often resort to a *non standard semantics*, which incorporates properties that would normally be left out of the concrete semantics.

3.2. Abstract interpretation and static analysis

Once a reference semantics has been fixed and a property of interest has been formalized, the definition of a static analysis requires the choice of an *abstraction*. The abstraction ties a set of *abstract predicates* to the concrete ones, which they denote. This relation is often expressed with a *concretization function* that maps each abstract element to the concrete property it stands for. Obviously, a well chosen abstraction should allow one to express the property of interest, as well as all the intermediate properties that are required in order to prove it (otherwise, the analysis would have no chance to achieve a successful verification). It should also lend itself to an efficient implementation, with efficient data-structures and algorithms for the representation and the manipulation of abstract predicates. A great number of abstractions have been proposed for all kinds of concrete data types, yet the search for new abstractions is a very important topic in static analysis, so as to target novel kinds of properties, to design more efficient or more precise static analyses.

Once an abstraction is chosen, a set of *sound abstract transformers* can be derived from the concrete semantics and that account for individual program steps, in the abstract level and without forgetting any concrete behavior. A static analysis follows as a result of this step by step approximation of the concrete semantics, when the abstract transformers are all computable. This process defines an *abstract interpretation* [22]. The case of loops requires a bit more work as the concrete semantics typically relies on a fixpoint that may not be computable in finitely many iterations. To achieve a terminating analysis we then use *widening operators* [22], which over-approximate the concrete union and ensure termination.

A static analysis defined that way always terminates and produces sound over-approximations of the programs behaviors. Yet, these results may not be precise enough for verification. This is where the art of static analysis design comes into play through, among others:

- the use of more precise, yet still efficient enough abstract domains;
- the combination of application-specific abstract domains;
- the careful choice of abstract transformers and widening operators.

3.3. Applications of the notion of abstraction in semantics

In the previous subsections, we sketched the steps in the design of a static analyzer to infer some family of properties, which should be implementable, and efficient enough to succeed in verifying non trivial systems.

The same principles can be applied successfully to other goals. In particular, the abstract interpretation framework should be viewed as a very general tool to *compare different semantics*, not necessarily with the goal of deriving a static analyzer. Such comparisons may be used in order to prove two semantics equivalent (i.e., one is an abstraction of the other and vice versa), or that a first semantics is strictly more expressive than another one (i.e., the latter can be viewed an abstraction of the former, where the abstraction actually makes some information redundant, which cannot be recovered). A classical example of such comparison is the classification of semantics of transition systems [21], which provides a better understanding of program semantics in general. For instance, this approach can be applied to get a better understanding of the semantics of a programming language, but also to select which concrete semantics should be used as a foundation for a static analysis, or to prove the correctness of a program transformation, compilation or optimization.

3.4. From properties to explanations

In many application domains, we can go beyond the proof that a program satisfies its specification. Abstractions can also offer new perspectives to understand how complex behaviors of programs emerge from simpler computation steps. Abstractions can be used to find compact and readable representations of sets of traces, causal relations, and even proofs. For instance, abstractions may decipher how the collective behaviors of agents emerge from the orchestration of their individual ones in distributed systems (such as consensus protocols, models of signaling pathways). Another application is the assistance for the diagnostic of alarms of a static analyzer.

Complex systems and software have often times intricate behaviors, leading to executions that are hard to understand for programmers and also difficult to reason about with static analyzers. Shared memory and distributed systems are notorious for being hard to reason about due to the interleaving of actions performed by different processes and the non-determinism of the network that might lose, corrupt, or duplicate messages. Reduction theorems, e.g., Lipton's theorem, have been proposed to facilitate reasoning about concurrency, typically transforming a system into one with a coarse-grained semantics that usually increases the atomic sections. We investigate reduction theorems for distributed systems and ways to compute the coarse-grained counter part of a system automatically. Compared with shared memory concurrency, automated methods to reason about distributed systems have been less investigated in the literature. We take a programming language approach based on high-level programming abstractions. We focus on partially-synchronous communication closed round-based models, introduced in the distributed algorithms community for its simpler proof arguments. The high-level language is compiled into a low-level (asynchronous) programming language. Conversely, systems defined under asynchronous programming paradigms are decompiled into the high-level programming abstractions. The correctness of the compilation/decompilation process is based on reduction theorems (in the spirit of Lipton and Elrad-Francez) that preserve safety and liveness properties.

In models of signaling pathways, collective behavior emerges from competition for common resources, separation of scales (time/concentration), non linear feedback loops, which are all consequences of mechanistic interactions between individual bio-molecules (e.g., proteins). While more and more details about mechanistic interactions are available in the literature, understanding the behavior of these models at the system level is far from easy. Causal analysis helps explaining how specific events of interest may occur. Model reduction techniques combine methods from different domains such as the analysis of information flow used in communication protocols, and tropicalization methods that comes from physics. The result is lower dimension systems that preserve the behavior of the initial system while focusing of the elements from which emerges the collective behavior of the system.

The abstraction of causal traces offer nice representation of scenarios that lead to expected or unexpected events. This is useful to understand the necessary steps in potential scenarios in signaling pathways; this is useful as well to understand the different steps of an intrusion in a protocol. Lastly, traces of computation of a static analyzer can themselves be abstracted, which provides assistance to classify true and false alarms. Abstracted traces are symbolic and compact representations of sets of counter-examples to the specification of a system which help one to either understand the origin of bugs, or to find that some information has been lost in the abstraction leading to false alarms.

4. Application Domains

4.1. Verification of safety critical embedded software

The verification of safety critical embedded software is a very important application domain for our group. First, this field requires a high confidence in software, as a bug may cause disastrous events. Thus, it offers an obvious opportunity for a strong impact. Second, such software usually have better specifications and a better design than many other families of software, hence are an easier target for developing new static analysis techniques (which can later be extended for more general, harder to cope with families of programs). This includes avionics, automotive and other transportation systems, medical systems ...

For instance, the verification of avionics systems represent a very high percentage of the cost of an airplane (about 30 % of the overall airplane design cost). The state of the art development processes mainly resort to testing in order to improve the quality of software. Depending on the level of criticality of a software (at the highest levels, any software failure would endanger the flight) a set of software requirements are checked with test suites. This approach is both costly (due to the sheer amount of testing that needs to be performed) and unsound (as errors may go unnoticed, if they do not arise on the test suite).

By contrast, static analysis can ensure higher software quality at a lower cost. Indeed, a static analyzer will catch all bugs of a certain kind. Moreover, a static analysis run typically lasts a few hours, and can be integrated in the development cycle in a seamless manner. For instance, ASTRÉE successfully verified the absence of runtime error in several families of safety critical fly-by-wire avionic software, in at most a day of computation, on standard hardware. Other kinds of synchronous embedded software have also been analyzed with good results.

In the future, we plan to greatly extend this work so as to verify *other families of embedded software* (such as communication, navigation and monitoring software) and *other families of properties* (such as security and liveness properties).

Embedded software in charge of communication, navigation, and monitoring typically relies on a *parallel* structure, where several threads are executed concurrently, and manage different features (input, output, user interface, internal computation, logging ...). This structure is also often found in automotive software. An even more complex case is that of *distributed* systems, where several separate computers are run in parallel and take care of several sub-tasks of a same feature, such as braking. Such a logical structure is not only more complex than the synchronous one, but it also introduces new risks and new families of errors (deadlocks, data-races...). Moreover, such less well designed, and more complex embedded software often utilizes more complex data-structures than synchronous programs (which typically only use arrays to store previous states) and may use dynamic memory allocation, or build dynamic structures inside static memory regions, which are actually even harder to verify than conventional dynamically allocated data structures. Complex data-structures also introduce new kinds of risks (the failure to maintain structural invariants may lead to runtime errors, non termination, or other software failures). To verify such programs, we will design additional abstract domains, and develop new static analysis techniques, in order to support the analysis of more complex programming language features such as parallel and concurrent programming with threads and manipulations of complex data structures. Due to their size and complexity, the verification of such families of embedded software is a major challenge for the research community.

Furthermore, embedded systems also give rise to novel security concerns. It is in particular the case for some aircraft-embedded computer systems, which communicate with the ground through untrusted communication media. Besides, the increasing demand for new capabilities, such as enhanced on-board connectivity, e.g. using mobile devices, together with the need for cost reduction, leads to more integrated and interconnected systems. For instance, modern aircrafts embed a large number of computer systems, from safety-critical cockpit avionics to passenger entertainment. Some systems meet both safety and security requirements. Despite thorough segregation of subsystems and networks, some shared communication resources raise the concern of possible intrusions. Because of the size of such systems, and considering that they are evolving entities, the only economically viable alternative is to perform automatic analyses. Such analyses of security and confidentiality properties have never been achieved on large-scale systems where security properties interact with other software properties, and even the mapping between high-level models of the systems and the large software base implementing them has never been done and represents a great challenge. Our goal is to prove empirically that the security of such large scale systems can be proved formally, thanks to the design of dedicated abstract interpreters.

The long term goal is to make static analysis more widely applicable to the verification of industrial software.

4.2. Static analysis of software components and libraries

An important goal of our work is to make static analysis techniques easier to apply to wider families of software. Then, in the longer term, we hope to be able to verify less critical, yet very commonly used pieces of software. Those are typically harder to analyze than critical software, as their development process tends to be less rigorous. In particular, we will target operating systems components and libraries. As of today, the verification of such programs is considered a major challenge to the static analysis community.

As an example, most programming languages offer Application Programming Interfaces (API) providing ready-to-use abstract data structures (e.g., sets, maps, stacks, queues, etc.). These APIs, are known under the name of containers or collections, and provide off-the-shelf libraries of high level operations, such as insertion, deletion and membership checks. These container libraries give software developers a way of abstracting from low-level implementation details related to memory management, such as dynamic allocation, deletion and pointer handling or concurrency aspects, such as thread synchronization. Libraries implementing data structures are important building bricks of a huge number of applications, therefore their verification is paramount. We are interested in developing static analysis techniques that will prove automatically the correctness of large audience libraries such as Glib and Threading Building Blocks.

4.3. Models of mechanistic interactions between proteins

Computer Science takes a more and more important role in the design and the understanding of biological systems such as signaling pathways, self assembly systems, DNA repair mechanisms. Biology has gathered large data-bases of facts about mechanistic interactions between proteins, but struggles to draw an overall picture of how these systems work as a whole. High level languages designed in Computer Science allow one to collect these interactions in integrative models, and provide formal definitions (i.e., semantics) for the behavior of these models. This way, modelers can encode their knowledge, following a bottom-up discipline, without simplifying *a priori* the models at the risk of damaging the key properties of the system. Yet, the systems that are obtained this way suffer from combinatorial explosion (in particular, in the number of different kinds of molecular components, which can arise at run-time), which prevents from a naive computation of their behavior.

We develop various analyses based on abstract interpretation, and tailored to different phases of the modeling process. We propose automatic static analyses in order to detect inconsistencies in the early phases of the modeling process. These analyses are similar to the analysis of classical safety properties of programs. They involve both forward and backward reachability analyses as well as causality analyses, and can be tuned at different levels of abstraction. We also develop automatic static analyses in order to identify key elements in the dynamics of these models. The results of these analyses are sent to another tool, which is used to automatically simplify models. The correctness of this simplification process is proved by the means of abstract interpretation: this ensures formally that the simplification preserves the quantitative properties that have been specified beforehand by the modeler. The whole pipeline is parameterized by a large choice of abstract domains which exploits different features of the high level description of models.

4.4. Consensus

Fault-tolerant distributed systems provide a dependable service on top of unreliable computers and networks. Famous examples are geo-replicated data-bases, distributed file systems, or blockchains. Fault-tolerant protocols replicate the system and ensure that all (unreliable) replicas are perceived from the outside as one single reliable machine. To give the illusion of a single reliable machine "consensus" protocols force replicas to agree on the "current state" before making this state visible to an outside observer. We are interested in (semi-)automatically proving the total correctness of consensus algorithms in the benign case (messages are lost or processes crash) or the Byzantine case (processes may lie about their current state). In order to do this, we first define new reduction theorems to simplify the behaviors of the system and, second, we introduce new static analysis methods to prove the total correctness of adequately simplified systems. We focus on static analysis based Satisfiability Modulo Theories (SMT) solvers which offers a good compromise between automation and expressiveness. Among our benchmarks are Paxos, PBFT (Practical Byzantine Fault-Tolerance), and blockchain algorithms (Red-Belly, Tendermint, Algorand). These are highly challenging benchmarks, with a lot of non-determinism coming from the interleaving semantics and from the adversarial environment in which correct processes execute, environment that can drop messages, corrupt them, etc. Moreover, these systems were originally designed for a few servers but today are deployed on networks with thousands of nodes. The "optimizations" for scalability can no longer be overlooked and must be considered as integral part of the algorithms, potentially leading to specifications weaker than the so much desired consensus.

4.5. Models of growth

In systems and synthetic biology (engineered systems) one would like study the environment of a given cellular process (such as signaling pathways mentioned earlier) and the ways in which that process interacts with different resources provided by the host. To do this, we have built coarse-grained models of cellular physiology which summarize fundamental processes (transcription, translation, transport, metabolism). such models describe global growth in mechanistic way and allow one to plug the model of one's process of interest into a simplified and yet realistic and reactive model of the process interaction with its immediate environment. A first ODE-based deterministic version of this model [26] explaining the famous bacterial growth laws and how the allocation of resources to different genomic sectors depends on the growth conditions- was published in 2015 and has already received nearly 150 citations. The model also allows one to bridge between population genetic models which describe cells in terms of abstract features and fitness and intra-cellular models. For instance, we find that fastest growing strategies are not evolutionary stable in competitive experiments. We also find that vastly different energy storage strategies exist [24]. In a recent article [25] in *Nature Communications* we build a stochastic version of the above model. We predict the empirical size and doubling time distributions as a function of growth conditions. To be able to fit the parameters of the model to available single-cell data (note that the fitting constraints are far tighter than in the deterministic case), we introduce new techniques for the approximation of reaction-division systems which generalize continuous approximations of Langevin type commonly used for pure reaction systems. We also use cross-correlations to visualize causality and modes in noise propagation in the model (in a way reminiscent to abstract computational traces mentioned earlier). In other work, we show how to connect our new class of models to more traditional ones stemming from "flux balance analysis" by introducing an allocation vector which allows one to assign a formal growth rate to a class of reaction systems [20].

4.6. Static analysis of data science software

Nowadays, thanks to advances in machine learning and the availability of vast amounts of data, computer software plays an increasingly important role in assisting or even autonomously performing tasks in our daily lives. As data science software becomes more and more widespread, we become increasingly vulnerable to programming errors. In particular, programming errors that do not cause failures can have serious consequences since code that produces an erroneous but plausible result gives no indication that something went wrong. This issue becomes particularly worrying knowing that machine learning software, thanks to its ability to efficiently approximate or simulate more complex systems, is slowly creeping into mission critical scenarios. However, programming errors are not the only concern. Another important issue is the vulnerability of machine learning models to adversarial examples, that is, small input perturbations that cause the model to misbehave in unpredictable ways. More generally, a critical issue is the notorious difficulty to interpret and explain machine learning software. Finally, as we are witnessing widespread adoption of software with far-reaching societal impact — i.e., to automate decision-making in fields such as social welfare, criminal justice, and even health care — a number of recent cases have evidenced the importance of ensuring software fairness as well as data privacy. Going forward, data science software will be subject to more and more legal regulations (e.g., the European General Data Protection Regulation adopted in 2016) as well as administrative audits. It is thus paramount to develop method and tools that can keep up with these developments and enhance our understanding of data science software and ensure it behaves correctly and reliably. In particular, we are interesting in developing new static analyses specifically tailored to the idiosyncrasies of data science software. This makes it a new and exciting area for static analysis, offering a wide variety of challenging problems with huge potential impact on various interdisciplinary application domains [13].

5. Highlights of the Year

5.1. Highlights of the Year

Caterina Urban joined the group as a CR in February 2019, and is opening new research directions towards static analysis for data-science software. She was invited to talk about her work in this area at SAS 2019 (Static Analysis Symposium) [13].

6. New Software and Platforms

6.1. APRON

SCIENTIFIC DESCRIPTION: The APRON library is intended to be a common interface to various underlying libraries/abstract domains and to provide additional services that can be implemented independently from the underlying library/abstract domain, as shown by the poster on the right (presented at the SAS 2007 conference. You may also look at:

FUNCTIONAL DESCRIPTION: The Apron library is dedicated to the static analysis of the numerical variables of a program by abstract interpretation. Its goal is threefold: provide ready-to-use numerical abstractions under a common API for analysis implementers, encourage the research in numerical abstract domains by providing a platform for integration and comparison of domains, and provide a teaching and demonstration tool to disseminate knowledge on abstract interpretation.

- Participants: Antoine Miné and Bertrand Jeannet
- Contact: Antoine Miné
- URL: http://apron.cri.ensmp.fr/library/

6.2. Astrée

The AstréeA Static Analyzer of Asynchronous Software

KEYWORDS: Static analysis - Static program analysis - Program verification - Software Verification - Abstraction

SCIENTIFIC DESCRIPTION: Astrée analyzes structured C programs, with complex memory usages, but without dynamic memory allocation nor recursion. This encompasses many embedded programs as found in earth transportation, nuclear energy, medical instrumentation, and aerospace applications, in particular synchronous control/command. The whole analysis process is entirely automatic.

Astrée discovers all runtime errors including:

undefined behaviors in the terms of the ANSI C99 norm of the C language (such as division by 0 or out of bounds array indexing),

any violation of the implementation-specific behavior as defined in the relevant Application Binary Interface (such as the size of integers and arithmetic overflows),

any potentially harmful or incorrect use of C violating optional user-defined programming guidelines (such as no modular arithmetic for integers, even though this might be the hardware choice),

failure of user-defined assertions.

FUNCTIONAL DESCRIPTION: Astrée analyzes structured C programs, with complex memory usages, but without dynamic memory allocation nor recursion. This encompasses many embedded programs as found in earth transportation, nuclear energy, medical instrumentation, and aerospace applications, in particular synchronous control/command. The whole analysis process is entirely automatic.

Astrée discovers all runtime errors including: - undefined behaviors in the terms of the ANSI C99 norm of the C language (such as division by 0 or out of bounds array indexing), - any violation of the implementation-specific behavior as defined in the relevant Application Binary Interface (such as the size of integers and arithmetic overflows), - any potentially harmful or incorrect use of C violating optional user-defined programming guidelines (such as no modular arithmetic for integers, even though this might be the hardware choice), - failure of user-defined assertions.

Astrée is a static analyzer for sequential programs based on abstract interpretation. The Astrée static analyzer aims at proving the absence of runtime errors in programs written in the C programming language.

- Participants: Antoine Miné, Jérôme Feret, Laurent Mauborgne, Patrick Cousot, Radhia Cousot and Xavier Rival
- Partners: CNRS - ENS Paris - AbsInt Angewandte Informatik GmbH
- Contact: Patrick Cousot
- URL: http://www.astree.ens.fr/

6.3. AstréeA

The AstréeA Static Analyzer of Asynchronous Software

KEYWORDS: Static analysis - Static program analysis

SCIENTIFIC DESCRIPTION: AstréeA analyzes C programs composed of a fixed set of threads that communicate through a shared memory and synchronization primitives (mutexes, FIFOs, blackboards, etc.), but without recursion nor dynamic creation of memory, threads nor synchronization objects. AstréeA assumes a real-time scheduler, where thread scheduling strictly obeys the fixed priority of threads. Our model follows the ARINC 653 OS specification used in embedded industrial aeronautic software. Additionally, AstréeA employs a weakly-consistent memory semantics to model memory accesses not protected by a mutex, in order to take into account soundly hardware and compiler-level program transformations (such as optimizations). AstréeA checks for the same run-time errors as Astrée , with the addition of data-races.

FUNCTIONAL DESCRIPTION: AstréeA is a static analyzer prototype for parallel software based on abstract interpretation. The AstréeA prototype is a fork of the Astrée static analyzer that adds support for analyzing parallel embedded C software.

- Participants: Antoine Miné, Jérôme Feret, Patrick Cousot, Radhia Cousot and Xavier Rival
- Partners: CNRS - ENS Paris - AbsInt Angewandte Informatik GmbH
- Contact: Patrick Cousot
- URL: http://www.astreea.ens.fr/

6.4. ClangML

KEYWORD: Compilation

FUNCTIONAL DESCRIPTION: ClangML is an OCaml binding with the Clang front-end of the LLVM compiler suite. Its goal is to provide an easy to use solution to parse a wide range of C programs, that can be called from static analysis tools implemented in OCaml, which allows to test them on existing programs written in C (or in other idioms derived from C) without having to redesign a front-end from scratch. ClangML features an interface to a large set of internal AST nodes of Clang , with an easy to use API. Currently, ClangML supports all C language AST nodes, as well as a large part of the C nodes related to C++ and Objective-C.

- Participants: Devin Mccoughlin, François Berenger and Pippijn Van Steenhoven
- Contact: Xavier Rival
- URL: https://github.com/Antique-team/clangml/tree/master/clang

6.5. FuncTion

SCIENTIFIC DESCRIPTION: FuncTion is based on an extension to liveness properties of the framework to analyze termination by abstract interpretation proposed by Patrick Cousot and Radhia Cousot. FuncTion infers ranking functions using piecewise-defined abstract domains. Several domains are available to partition the ranking function, including intervals, octagons, and polyhedra. Two domains are also available to represent the value of ranking functions: a domain of affine ranking functions, and a domain of ordinal-valued ranking functions (which allows handling programs with unbounded non-determinism).

FUNCTIONAL DESCRIPTION: FuncTion is a research prototype static analyzer to analyze the termination and functional liveness properties of programs. It accepts programs in a small non-deterministic imperative language. It is also parameterized by a property: either termination, or a recurrence or a guarantee property (according to the classification by Manna and Pnueli of program properties). It then performs a backward static analysis that automatically infers sufficient conditions at the beginning of the program so that all executions satisfying the conditions also satisfy the property.

- Participants: Antoine Miné and Caterina Urban
- Contact: Caterina Urban
- URL: http://www.di.ens.fr/~urban/FuncTion.html

6.6. HOO

Heap Abstraction for Open Objects

FUNCTIONAL DESCRIPTION: JSAna with HOO is a static analyzer for JavaScript programs. The primary component, HOO, which is designed to be reusable by itself, is an abstract domain for a dynamic language heap. A dynamic language heap consists of open, extensible objects linked together by pointers. Uniquely, HOO abstracts these extensible objects, where attribute/field names of objects may be unknown. Additionally, it contains features to keeping precise track of attribute name/value relationships as well as calling unknown functions through desynchronized separation.

As a library, HOO is useful for any dynamic language static analysis. It is designed to allow abstractions for values to be easily swapped out for different abstractions, allowing it to be used for a wide-range of dynamic languages outside of JavaScript.

- Participant: Arlen Cox
- Contact: Arlen Cox

6.7. MemCAD

The MemCAD static analyzer

KEYWORDS: Static analysis - Abstraction

FUNCTIONAL DESCRIPTION: MemCAD is a static analyzer that focuses on memory abstraction. It takes as input C programs, and computes invariants on the data structures manipulated by the programs. It can also verify memory safety. It comprises several memory abstract domains, including a flat representation, and two graph abstractions with summaries based on inductive definitions of data-structures, such as lists and trees and several combination operators for memory abstract domains (hierarchical abstraction, reduced product). The purpose of this construction is to offer a great flexibility in the memory abstraction, so as to either make very efficient static analyses of relatively simple programs, or still quite efficient static analyses of very involved pieces of code. The implementation consists of over 30 000 lines of ML code, and relies on the ClangML front-end. The current implementation comes with over 300 small size test cases that are used as regression tests.

- Participants: Antoine Toubhans, François Berenger, Huisong Li and Xavier Rival
- Contact: Xavier Rival
- URL: http://www.di.ens.fr/~rival/memcad.html

6.8. KAPPA

A rule-based language for modeling interaction networks

KEYWORDS: Systems Biology - Modeling - Static analysis - Simulation - Model reduction

SCIENTIFIC DESCRIPTION: OpenKappa is a collection of tools to build, debug and run models of biological pathways. It contains a compiler for the Kappa Language, a static analyzer (for debugging models), a simulator, a compression tool for causal traces, and a model reduction tool.

FUNCTIONAL DESCRIPTION: Kappa is provided with the following tools: - a compiler - a stochastic simulator - a static analyzer - a trace compression algorithm - an ODE generator.

RELEASE FUNCTIONAL DESCRIPTION: On line UI, Simulation is based on a new data-structure (see ESOP 2017), New abstract domains are available in the static analyzer (see SASB 2016), Local traces (see TCBB 2018), Reasoning on polymers (see SASB 2018).

- Participants: Jean Krivine, Jérôme Feret, Kim-Quyen Ly, Pierre Boutillier, Russ Harmer, Vincent Danos and Walter Fontana
- Partners: ENS Lyon - Université Paris-Diderot - HARVARD Medical School
- Contact: Jérôme Feret
- URL: http://www.kappalanguage.org/

6.9. QUICr

FUNCTIONAL DESCRIPTION: QUICr is an OCaml library that implements a parametric abstract domain for sets. It is constructed as a functor that accepts any numeric abstract domain that can be adapted to the interface and produces an abstract domain for sets of numbers combined with numbers. It is relational, flexible, and tunable. It serves as a basis for future exploration of set abstraction.

- Participant: Arlen Cox
- Contact: Arlen Cox

6.10. LCertify

KEYWORD: Compilation

SCIENTIFIC DESCRIPTION: The compilation certification process is performed automatically, thanks to a prover designed specifically. The automatic proof is done at a level of abstraction which has been defined so that the result of the proof of equivalence is strong enough for the goals mentioned above and so that the proof obligations can be solved by efficient algorithms.

FUNCTIONAL DESCRIPTION: Abstract interpretation, Certified compilation, Static analysis, Translation validation, Verifier. The main goal of this software project is to make it possible to certify automatically the compilation of large safety critical software, by proving that the compiled code is correct with respect to the source code: When the proof succeeds, this guarantees semantic equivalence. Furthermore, this approach should allow to meet some domain specific software qualification criteria (such as those in DO-178 regulations for avionics software), since it allows proving that successive development levels are correct with respect to each other i.e., that they implement the same specification. Last, this technique also justifies the use of source level static analyses, even when an assembly level certification would be required, since it establishes separately that the source and the compiled code are equivalent.ntees that no compiler bug did cause incorrect code to be generated.

- Participant: Xavier Rival
- Partners: CNRS - ENS Paris
- Contact: Xavier Rival
- URL: http://www.di.ens.fr/~rival/lcertify.html

6.11. Zarith

FUNCTIONAL DESCRIPTION: Zarith is a small (10K lines) OCaml library that implements arithmetic and logical operations over arbitrary-precision integers. It is based on the GNU MP library to efficiently implement arithmetic over big integers. Special care has been taken to ensure the efficiency of the library also for small integers: small integers are represented as Caml unboxed integers and use a specific C code path. Moreover, optimized assembly versions of small integer operations are provided for a few common architectures.

Zarith is currently used in the Astrée analyzer to enable the sound analysis of programs featuring 64-bit (or larger) integers. It is also used in the Frama-C analyzer platform developed at CEA LIST and Inria Saclay.

- Participants: Antoine Miné, Pascal Cuoq and Xavier Leroy
- Contact: Antoine Miné
- URL: http://forge.ocamlcore.org/projects/zarith

6.12. PYPPAI

Pyro Probabilistic Program Analyzer

KEYWORDS: Probability - Static analysis - Program verification - Abstraction

FUNCTIONAL DESCRIPTION: PYPPAI is a program analyzer to verify the correct semantic definition of probabilistic programs written in Pyro. At the moment, PYPPAI verifies consistency conditions between models and guides used in probabilistic inference programs.

PYPPAI is written in OCaml and uses the pyml Python in OCaml library. It features a numerical abstract domain based on Apron, an abstract domain to represent zones in tensors, and dedicated abstract domains to describe distributions and states in probabilistic programs.

- Contact: Xavier Rival
- URL: https://github.com/wonyeol/static-analysis-for-support-match

7. New Results

7.1. Relational Static Analysis

7.1.1. Relational abstraction for memory properties

Participants: Hugo Illous, Matthieu Lemerre, Xavier Rival [correspondant].

Static analyses aim at inferring semantic properties of programs. We can distinguish two important classes of static analyses: state analyses and relational analyses. While state analyses aim at computing an over-approximation of reachable states of programs, relational analyses aim at computing functional properties over the input-output states of programs. Several advantages of relational analyses are their ability to analyze incomplete programs, such as libraries or classes, but also to make the analysis modular, using input-output relations as composable summaries for procedures. In the case of numerical programs, several analyses have been proposed that utilize relational numerical abstract domains to describe relations. On the other hand, designing abstractions for relations over input-output memory states and taking shapes into account is challenging. We have proposed a set of novel logical connectives to describe such relations, which are inspired by separation logic. This logic can express that certain memory areas are unchanged, freshly allocated, or freed, or that only part of the memory was modified. Using these connectives, we have built an abstract domain and design a static analysis that over-approximates relations over memory states containing inductive structures. We implemented this analysis and evaluated it on a basic library of list manipulating functions.

This work was done as part of the Phd of Hugo Illous [10] and a journal paper is currently under submission.

7.2. Static Analysis of Probabilistic Programming Languages

7.2.1. Towards the verification of semantic assumptions required by probabilistic inference algorithms

Participants: Wonyeol Lee, Hangyeol Wu, Xavier Rival [correspondant], Hongseok Yang.

Probabilistic programming is the idea of writing models from statistics and machine learning using program notations and reasoning about these models using generic inference engines. Recently its combination with deep learning has been explored intensely, which led to the development of so called deep probabilistic programming languages, such as Pyro, Edward and ProbTorch. At the core of this development lie inference engines based on stochastic variational inference algorithms. When asked to find information about the posterior distribution of a model written in such a language, these algorithms convert this posterior-inference query into an optimisation problem and solve it approximately by a form of gradient ascent or descent. We analysed one of the most fundamental and versatile variational inference algorithms, called score estimator or REINFORCE, using tools from denotational semantics and program analysis. We formally expressed what this algorithm does on models denoted by programs, and exposed implicit assumptions made by the algorithm on the models. The violation of these assumptions may lead to an undefined optimisation objective or the loss of convergence guarantee of the optimisation process. We then describe rules for proving these assumptions, which can be automated by static program analyses. Some of our rules use nontrivial facts from continuous mathematics, and let us replace requirements about integrals in the assumptions, such as integrability of functions defined in terms of programs' denotations, by conditions involving differentiation or boundedness, which are much easier to prove automatically (and manually). Following our general methodology, we have developed a static program analysis for the Pyro programming language that aims at discharging the assumption about what we call model-guide support match. Our analysis is applied to the eight representative model-guide pairs from the Pyro webpage, which include sophisticated neural network models such as AIR. It found a bug in one of these cases, and revealed a non-standard use of an inference engine in another, and showed that the assumptions are met in the remaining six cases.

This work has been published in [12].

7.3. Static Analysis of JavaScript Code

7.3.1. Weakly Sensitive Analysis for Unbounded Iteration over JavaScript Objects

Participants: Yoonseok Ko, Xavier Rival [correspondant], Sukyoung Ryu.

In [23] and [11], we studied composite object abstraction for the analysis JavaScript.

JavaScript framework libraries like jQuery are widely use, but complicate program analyses. Indeed, they encode clean high-level constructions such as class inheritance via dynamic object copies and transformations that are harder to reason about. One common pattern used in them consists of loops that copy or transform part or all of the fields of an object. Such loops are challenging to analyze precisely, due to weak updates and as unrolling techniques do not always apply. In this work, we observe that precise field correspondence relations are required for client analyses (e.g., for call-graph construction), and propose abstractions of objects and program executions that allow to reason separately about the effect of distinct iterations without resorting to full unrolling. We formalize and implement an analysis based on this technique. We assess the performance and precision on the computation of call-graph information on examples from jQuery tutorials.

7.4. Rule-based Modeling with Arithmetics

7.4.1. Counters in Kappa: Semantics, Simulation, and Static Analysis.

Participants: Pierre Boutillier, Ioana Cristescu, Jérôme Feret.

Site-graph rewriting languages, such as Kappa or BNGL, offer parsimonious ways to describe highly combinatorial systems of mechanistic interactions among proteins. These systems may be then simulated efficiently. Yet, the modeling mechanisms that involve counting (a number of phosphorylated sites for instance) require an exponential number of rules in Kappa. In BNGL, updating the set of the potential applications of rules in the current state of the system comes down to the sub-graph isomorphism problem (which is NP-complete).

In [14], we extend Kappa to deal both parsimoniously and efficiently with counters. We propose a single push-out semantics for Kappa with counters. We show how to compile Kappa with counters into Kappa without counters (without requiring an exponential number of rules). We design a static analysis, based on affine relationships, to identify the meaning of counters and bound their ranges accordingly.

7.5. Reduced product

7.5.1. Sharing Ghost Variables in a Collection of Abstract Domains.

Participants: Marc Chevalier, Jérôme Feret.

In abstract interpretation, it is often necessary to be able to express complex properties while doing a precise analysis. A way to achieve that is to combine a collection of domains, each handling some kind of properties, using a reduced product. Separating domains allows an easier and more modular implementation, and eases soundness and termination proofs. This way, we can add a domain for any kind of property that is interesting. The reduced product, or an approximation of it, is in charge of refining abstract states, making the analysis precise.

In program verification, ghost variables can be used to ease proofs of properties by storing intermediate values that do not appear directly in the execution.

In [15], we propose a reduced product of abstract domains that allows domains to use ghost variables to ease the representation of their internal state. Domains must be totally agnostic with respect to other existing domains. In particular the handling of ghost variables must be entirely decentralized while still ensuring soundness and termination of the analysis.

7.6. Static Analysis of Neural Networks

7.6.1. Perfectly Parallel Fairness Certification.

Participants: Caterina Urban [correspondant], Maria Christakis, Valentin Wüestholz, Fuyuan Zhang.

Recently, there is growing concern that machine-learning models, which currently assist or even automate decision making, reproduce, and in the worst case reinforce, bias of the training data. The development of tools and techniques for certifying fairness of these models or describing their biased behavior is, therefore, critical.

In [19], we propose a perfectly parallel static analysis for certifying causal fairness of feed-forward neural networks used for classification tasks. When certification succeeds, our approach provides definite guarantees, otherwise, it describes and quantifies the biased behavior. We design the analysis to be sound, in practice also exact, and configurable in terms of scalability and precision, thereby enabling pay-as-you-go certification. We implement our approach in an open-source tool and demonstrate its effectiveness on models trained with popular datasets.

7.7. Reductions between synchronous and asynchronous programming abstractions

7.7.1. Communication closed asynchronous protocols.

Participants: Andrei Damien, Cezara Drăgoi, Alexandru Militaru, Josef Widder.

Fault-tolerant distributed systems are implemented over asynchronous networks, where performance emerges from the load of the system. Due to asynchronous communication and the occurrence of faults (e.g., process crashes or the network dropping messages) the implementations are hard to understand and analyze. In contrast, synchronous computation models simplify design and reasoning.

In [17], we defined the first algorithm that automatically transforms an asynchronous protocol into a synchronous one. The method is sound but not complete. The transformation is based on an axiomatization of the notion of communication closure introduce by Elrad and Frances. If the asynchronous protocol is communication-closed then the translator will successfully compute its synchronous counter-part. Checking communication closure is done locally without considering any interferences between processes. The translator was successfully applied to Multi-Paxos, ViewStamped, and the atomic broadcast of Chandra and Toueg, generating the first synchronous counterparts of these protocols. The transformation from asynchronous to synchronous preserves the local states process go through and the exchanged messages. The translator has been implemented in a prototype tool called Athos, i.e., Asynchronous To Heard-Of Synchronizer, that is open source. The tool takes as input protocols in an intermediate protocol languages that has an asynchronous semantics and it is very close to C. These results have been published in one of the main verification venues Computer Aided Verification, CAV 2019 (acceptance rate <25% out of >250 submissions). The impact of the translator from asynchronous protocols to equivalent synchronous ones is important for the verification community because such a transformation reduces dramatically the state space and the set of traces to explore in order to prove the program correct, independently of the used verification technique.

7.7.2. *Executable Rounds: a Programming Abstraction for Fault-Tolerant Protocols.*

Participants: Cezara Drăgoi, Josef Widder, Damien Zufferey.

Fault-tolerant distributed systems are notoriously difficult to design and implement. Although programming languages for distributed systems is an active research area, appropriate synchronization primitives for fault-tolerance and group communication remains an important challenge. In [18] we present a new programming abstraction, HSync, for implementing benign and Byzantine distributed protocols. HSync is based on communication-closed rounds. Round models offer a simple abstraction for group communication and communication-closed rounds simplify dealing with faults. Protocols are implemented in a modular way in HSync. The language separates the message reception from the process local computation. It extends classic rounds with language constructs that give to the programmer the possibility to implement network and algorithm-specific policies for message reception. We have implemented an execution platform for HSync that runs on top of commodity hardware. We evaluate experimentally its performance, by comparing consensus implementations in HSync with LibPaxos3 and Bft-SMaRt, two consensus libraries tolerant to benign, resp. Byzantine faults.

7.8. Introduction

Participant: Andreea Beica.

The PhD of Andreea Beica [9] aims at studying two aspects related to the modelling of Biochemical Reaction Net- works, in the context of Systems Biology.

In the first part, we analyse how scale-separation in biological systems can be exploited for model reduction. We first argue for the use of rule-based models for prototyping genetic circuits, and then show how the inherent multi-scaleness of such systems can be used to devise a general model approximation method for rule-based models of genetic regulatory networks. The reduction proceeds via static analysis of the rule system. Our method relies on solid physical justifications, however not unlike other scale-separation reduction techniques, it lacks precise methods for quantifying the approximation error, while avoiding to solve the original model. Consequently, we next propose an approximation method for deterministic models of biochemical networks, in which reduction guarantees represent the major requirement. This second method combines abstraction and numerical approximation, and aims at providing a better understanding of model reduction methods that are based on time- and concentration- scale separation.

In the second part of the thesis, we introduce a new re-parametrisation technique for differential equation models of biochemical networks, in order to study the effect of intracellular resource storage strategies on growth, in self-replicating mechanistic models. Finally, we aim towards the characterisation of cellular growth as an emergent property of a novel Petri Net model semantics of Biochemical Reaction Networks.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Follow up to the AnaStaSec project

Title: Analyse de propriété de sécurité

Type: Research contracts funded by AirBus France

Duration: March 2019 - August 2018 and November 2019 - March 2020

Inria contact: Jérôme Feret

Abstract: An emerging structure in our information processing-based society is the notion of trusted complex systems interacting via heterogeneous networks with an open, mostly untrusted world. This view characterises a wide variety of systems ranging from the information system of a company to the connected components of a private house, all of which have to be connected with the outside.

The goal of these constracts is to analyse an application that is used to filter messages from higher-level security regions to lower-level ones in trusted complex systems. This application shall check that messages are well-formed and that they match with existing requests. Moreover, so as to limit potential flows of information, one shall prove that the internal state of buffers are reset between the processing of each packet.

To certify these properties, the front-end of ASTRÉE has been upgraded with new directives to specify the properties of interest, and the analysis has been tuned to improve the analysis : 1) ghost variables are used to record the value of buffers between each packet processing so that already existing relational domains can prove that they are restored to the correct value, and 2) data-partitioning strategies have been implemented to separate the different modes of usage.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. AnaStaSec

Title: Static Analysis for Security Properties

Type: ANR générique 2014

Defi: Société de l'information et de la communication

Instrument: ANR grant

Duration: January 2015 - September 2019

Coordinator: Inria Paris-Rocquencourt (France)

Others partners: Airbus France (France), AMOSSYS (France), CEA LIST (France), Inria Rennes-Bretagne Atlantique (France), TrustInSoft (France)

Inria contact: Jérôme Feret

See also: http://www.di.ens.fr/ feret/anastasec/

Abstract: An emerging structure in our information processing-based society is the notion of trusted complex systems interacting via heterogeneous networks with an open, mostly untrusted world. This view characterises a wide variety of systems ranging from the information system of a company to the connected components of a private house, all of which have to be connected with the outside.

It is in particular the case for some aircraft-embedded computer systems, which communicate with the ground through untrusted communication media. Besides, the increasing demand for new capabilities, such as enhanced on-board connectivity, e.g. using mobile devices, together with the need for cost reduction, leads to more integrated and interconnected systems. For instance, modern aircrafts embed a large number of computer systems, from safety-critical cockpit avionics to passenger entertainment. Some systems meet both safety and security requirements. Despite thorough segregation of subsystems and networks, some shared communication resources raise the concern of possible intrusions.

Some techniques have been developed and still need to be investigated to ensure security and confidentiality properties of such systems. Moreover, most of them are model-based techniques operating only at architectural level and provide no guarantee on the actual implementations. However, most security incidents are due to attackers exploiting subtle implementation-level software vulnerabilities. Systems should therefore be analyzed at software level as well (i.e. source or executable code), in order to provide formal assurance that security properties indeed hold for real systems.

Because of the size of such systems, and considering that they are evolving entities, the only economically viable alternative is to perform automatic analyses. Such analyses of security and confidentiality properties have never been achieved on large-scale systems where security properties interact with other software properties, and even the mapping between high-level models of the systems and the large software base implementing them has never been done and represents a great challenge. The goal of this project is to develop the new concepts and technologies necessary to meet such a challenge.

The project ANASTASEC project will allow for the formal verification of security properties of software-intensive embedded systems, using automatic static analysis techniques at different levels of representation: models, source and binary codes. Among expected outcomes of the project will be a set of prototype tools, able to deal with realistic large systems and the elaboration of industrial security evaluation processes, based on static analysis.

9.1.2. DCore

Title: DCore - Causal Debugging for Concurrent Systems

Type: ANR générique 2018

Defi: Société de l'information et de la communication

Instrument: ANR grant

Duration: March 2019 - February 2023

Coordinator: Inria Grenoble - Rhône-Alpes (France)

Others partners: IRIF (France), Inria Paris (France)

Inria contact: Jérôme Feret

See also: https://project.inria.fr/dcore/

Abstract: As software takes over more and more functionalities in embedded and safety-critical systems, bugs may endanger the safety of human beings and of the environment, or entail heavy financial losses. In spite of the development of verification and testing techniques, debugging still plays a crucial part in the arsenal of the software developer. Unfortunately, usual debugging techniques do not scale to large concurrent and distributed systems: they fail to provide precise and efficient means to inspect and analyze large concurrent executions; they do not provide means to automatically reveal software faults that constitute actual causes for errors; and they do not provide succinct and relevant explanations linking causes (software bugs) to their effects (errors observed during execution).

The overall objective of the project is to develop a semantically well-founded, novel form of concurrent debugging, which we call "causal debugging", that aims to alleviate the deficiencies of current debugging techniques for large concurrent software systems.

Briefly, the causal debugging technology developed by the DCore project will comprise and integrate two main novel engines:

1. A reversible execution engine that allows programmers to backtrack and replay a concurrent or distributed program execution, in a way that is both precise and efficient (only the exact threads involved by a return to a target anterior or posterior program state are impacted);
2. a causal analysis engine that allows programmers to analyze concurrent executions, by asking questions of the form "what caused the violation of this program property?", and that allows for the precise and efficient investigation of past and potential program executions.

The project will build its causal debugging technology on results obtained by members of the team, as part of the past ANR project REVER, on the causal semantics of concurrent languages, and the semantics of concurrent reversible languages, as well as on recent works by members of the project on abstract interpretation, causal explanations and counterfactual causal analysis.

The project primarily targets multithreaded, multicore and multiprocessor software systems, and functional software errors, that is errors that arise in concurrent executions because of faults (bugs) in software that prevents it to meet its intended function. Distributed systems, which can be impacted by network failures and remote site failures are not an immediate target for DCore, although the technology developed by the project should constitute an important contribution towards full-fledged distributed debugging. Likewise, we do not target performance or security errors, which come with specific issues and require different levels of instrumentation, although the DCore technology should prove a key contribution in these areas as well.

9.1.3. REPAS

The project REPAS, Reliable and Privacy-Aware Software Systems via Bisimulation Metrics (coordination Catuscia Palamidessi, Inria Saclay), aims at investigating quantitative notions and tools for proving program correctness and protecting privacy, focusing on bisimulation metrics, the natural extension of bisimulation on quantitative systems. A key application is to develop mechanisms to protect the privacy of users when their location traces are collected. Partners: Inria (Comete, Focus), ENS Cachan, ENS Lyon, University of Bologna.

9.1.4. SAFTA

Title: SAFTA Static Analysis for Fault-Tolerant distributed Algorithms.

Type: ANR JCJC 2018

Duration: February 2018 - August 2022

Coordinator: Cezara Drăgoi, CR Inria

Abstract: Fault-tolerant distributed data structures are at the core distributed systems. Due to the multiple sources of non-determinism, their development is challenging. The project aims to increase the confidence we have in distributed implementations of data structures. We think that the difficulty does not only come from the algorithms but from the way we think about distributed systems. In this project we investigate partially synchronous communication-closed round based programming abstractions that reduce the number of interleavings, simplifying the reasoning about distributed systems and their proof arguments. We use partial synchrony to define reduction theorems from asynchronous semantics to partially synchronous ones, enabling the transfer of proofs from the synchronous world to the asynchronous one. Moreover, we define a domain specific language, that allows the programmer to focus on the algorithm task, it compiles into efficient asynchronous code, and it is equipped with automated verification engines.

9.1.5. TGFSYSBIO

Title: Microenvironment and cancer: regulation of TGF-β signaling

Type: Plan Cancer 2014-2019

Duration: December 2015 - September 2019

Coordinator: INSERM U1085-IRSET

Others partners: Inria Paris (France), Inria Rennes-Bretagne Atlantique (France),

Inria contact: Jérôme Feret

Abstract: Most cases of hepatocellular carcinoma (HCC) develop in cirrhosis resulting from chronic liver diseases and the Transforming Growth Factor β (TGF-β) is widely regarded as both the major pro-fibrogenic agent and a critical inducer of tumor progression and invasion. Targeting the deleterious effects of TGF-β without affecting its physiological role is the common goal of therapeutic strategies. However, identification of specific targets remains challenging because of the pleiotropic effects of TGF-β linked to the complex nature of its extracellular activation and signaling networks.

Our project proposes a systemic approach aiming at to identifying the potential targets that regulate the shift from anti- to pro-oncogenic effects of TGF-β. To that purpose, we will combine a rule-based model (Kappa language) to describe extracellular TGF-beta activation and large-scale state-transition based (Cadbiom formalism) model for TGF-β-dependent intracellular signaling pathways. The multi-scale integrated model will be enriched with a large-scale analysis of liver tissues using shotgun proteomics to characterize protein networks from tumor microenvironment whose remodeling is responsible for extracellular activation of TGF-β. The trajectories and upstream regulators of the final model will be analyzed with symbolic model checking techniques and abstract interpretation combined with causality analysis. Candidates will be classified with semantic-based approaches and symbolic bi-clustering technics. All efforts must ultimately converge to experimental validations of hypotheses and we will use our hepatic cellular models (HCC cell lines and hepatic stellate cells) to screen inhibitors on the behaviors of TGF-β signal.

The expected results are the first model of extracellular and intracellular TGF-β system that might permit to analyze the behaviors of TGF-β activity during the course of liver tumor progression and to identify new biomarkers and potential therapeutic targets.

9.1.6. VeriAMOS

Title: Verification of Abstract Machines for Operating Systems

Type: ANR générique 2018

Defi: Société de l'information et de la communication

Instrument: ANR grant

Duration: January 2019 - December 2022

Coordinator: Inria Paris (France)

Others partners: LIP6 (France), IRISA (France), UGA (France)

Inria contact: Xavier Rival

Abstract: Operating System (OS) programming is notoriously difficult and error prone. Moreover, OS bugs can have a serious impact on the functioning of computer systems. Yet, the verification of OSes is still mostly an open problem, and has only been done using user-assisted approaches that require a huge amount of human intervention. The VeriAMOS proposal relies on a novel approach to automatically and fully verifying OS services, that combines Domain Specific Languages (DSLs) and automatic static analysis. In this approach, DSLs provide language abstraction and let users express complex policies in high-level simple code. This code is later compiled into low level C code, to be executed on an abstract machine. Last, the automatic static analysis verifies structural and robustness properties on the abstract machine and generated code. We will apply this approach to the automatic, full verification of input/output schedulers for modern supports like SSDs.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

Type: IDEAS

Defi:

Instrument: ERC Proof of Concept Grant 2018

Objectif: Static Analysis for the VErification of Spreadsheets

Duration: January 2019 - June 2020

Coordinator: Inria (France)

Partner: None

Inria contact: Xavier Rival

Abstract: Spreadsheet applications (such as Microsoft Excel + VBA) are heavily used in a wide range of application domains including engineering, finance, management, statistics and health. However, they do not ensure robustness properties, thus spreadsheet errors are common and potentially costly. According to estimates, the annual cost of spreadsheet errors is around 7 billion dollars. For instance, in 2013, a series of spreadsheet errors at JPMorgan incurred 6 billion dollars trading losses. Yet, expert reports estimate about 90 % of the spreadsheets contain errors. The MemCAD ERC StG project opened the way to novel formal analysis techniques for spreadsheet applications. We propose to leverage these results into a toolbox able to safely *verify*, *optimize* and *maintain* spreadsheets, so as to reduce the likelihood of spreadsheet disasters. This toolbox will be commercialized by the startup MATRIXLEAD.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

Xavier Rival has a long standing collaboration with Bor-Yuh Evan Chang (University of Colorado, Boulder, USA), on the abstraction of symbolic properties and of complex memory data-structures.

Xavier Rival has a long standing collaboration with Sukyoung Ryu (KAIST, Daejeon, South Korea), on the analysis of dynamic programming languages. Xavier Rival has a set up a collaboration with Hongseok Yang (KAIST, Daejeon, South Korea), on the verification of probabilistic programs such as programs built in the Pyro framework.

Xavier Rival has started a collaboration with Shinya Katsumata, Jérémy Dubut, and Ichiro Hasuo (NII, Tokyo, Japan) on the formalization of abstract domains.

Xavier Rival has been working with Kwangkeun Yi on the writing of a book that should serve as an introduction to the field of static analysis, for students and engineers.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

9.4.1.1. Internships

Marc Chevalier and Jérôme Feret have supervised the L3 internship of Jérôme Boillot (L3 at ENS Lyon).

Jérôme Feret has ksupervised the M2 internship of Yvan Sraka (M2 UPMC).

Xavier Rival has supervised M1 Internships of Guillaume Reboullet and of Luc Chabassier (M1 at DIENS).

Xavier Rival has supervised M2 Internships of Josselin Giet (MPRI at ENS) and of Vincent Rébiscoul (M2 at ENS Lyon).

9.4.2. Visits to International Teams

9.4.2.1. Research Stays Abroad

Xavier Rival has visited KAIST and Seoul National University in November 2019.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- Jérôme Feret is a guest member of the Steering Committee of the Conference on Computational Methods in Systems Biology (CMSB).
- Jérôme Feret is a member of the Steering Committee of the Workshop on Static Analysis and Systems Biology (SASB).
- Xavier Rival is a member of the Steering Committee of the Static Analysis Symposium (SAS).
- Xavier Rival is a member of the Steering Committee of the Workshop on Tools for Automatic Program Analysis (TAPAS).
- Caterina Urban is a member of the Executive Board of ETAPS (European Joint Conferences on Theory & Practice of Software).

10.1.1.2. Member of the Organizing Committees

Caterina Urban serves as Chair of the Award Committee of the ETAPS Doctoral Dissertation Award 2020.

10.1.2. Event organization

- Cezara Drăgoi co-organize "VDS: Workshop on Verification of Distributed Systems" https://goto.ucsd.edu/vds/ with Klaus v. Gleissenthall from Univerisity of California, San Diego. This is a workshop that brings together two communities: verification and distributed systems by having participants from the two communities.

10.1.3. Scientific Events: Selection

10.1.3.1. Member of the Conference Program Committees

- Jérôme Feret served as a Member of the Program Committee of SASB 2019 (Static Analysis and Systems Biology).
- Jérôme Feret served as a Member of the Program Committee of CMSB 2019 (Computational Methods in Systems Biology).
- Jérôme Feret served as a Member of the Program Committee of CIBCB 2019 (Computational Intelligence in Bioinformatics and Computational Biology).
- Jérôme Feret served as a Member of the Program Committee of HSB 2019 (Hybrid Systems Biology).
- Jérôme Feret is serving as a member of the Program Committee of CMSB 2020 (Computational Methods in Systems Biology).
- Jérôme Feret is serving as a member of the Program Committee of HSB 2020 (Hybrid Systems Biology).
- Xavier Rival served as a Member of the Program Committee of SAS 2019 (Static Analysis Symposium).
- Xavier Rival is serving as a Member of the Program Committee of POPL 2020 (Symposium on Principles Of Programming Languages).

- Caterina Urban served as a member of the Program Committee of VMCAI 2019 (Verification, Model Checking, and Abstract Interpretation).
- Caterina Urban served as a member of the Artifact Evaluation Committee of POPL 2019 (Symposium on Principles of Programming Languages).
- Caterina Urban served as a member of the Program Committee of CAV 2019 (Computer Aided Verification).
- Caterina Urban served as a member of the Artifact Evaluation Committee of PLDI 2019 (Programming Languages Design and Implementation).
- Caterina Urban served as a member of the Program Committee of NSV 2019 (International Workshop on Numerical Software Verification).
- Caterina Urban served as a member of the Program Committee of LOPSTR 2019 (Logic-based Program Synthesis and Transformation).
- Caterina Urban served as a member of the Program Committee of TAPAS 2019 (Workshop on Tools for Automatic Program Analysis).
- Caterina Urban served as a member of the Program Committee of EMSOFT 2019 (Embedded Software).
- Caterina Urban served as a member of the Program Committee of iFM 2019 (integrated Formal Methods).
- Caterina Urban serves as a member of the Program Committee of VMCAI 2020 (Verification, Model Checking, and Abstract Interpretation).
- Caterina Urban serves as a member of the Program Committee of ESOP 2020 (European Symposium on Programming).
- Caterina Urban serves as a member of the Program Committee of CAV 2020 (Computer Aided Verification).
- Caterina Urban serves as a member of the Program Committee of SOAP 2020 (Workshop on the State Of the Art in Program Analysis).
- Cezara Drăgoi served as a member of the Program Committee of POPL 2020 (Symposium on Principles of Programming Languages)
- Cezara Drăgoi served as a member of the Program Committee of PLDI 2020 (Symposium on Programming Languages Design and Implementations)
- Cezara Drăgoi served as a member of the Program Committee of CAV 2019 (Computer Aided Verification).
- Cezara Drăgoi served as a member of the Program Committee of VMCAI 2019 (Verification, Model Checking, and Abstract Interpretation).
- Cezara Drăgoi served as a member of the Program Committee of NETYS 2019 (Conference on Networked Systems).

10.1.3.2. Reviewer

- Marc Chevalier served as Reviewer for SAS 2019 (Static Analysis Symposium).
- Marc Chevalier served as Reviewer for VMCAI 2020 (Verification, Model Checking, and Abstract Interpretation).
- Jérôme feret served as Reviewer for CAV 2019 (computer aided verification).
- Xavier Rival served as Reviewer for ESOP 2020 (European Symposium on Programming).
- Caterina Urban served as Reviewer for NFM 2019 (NASA Formal Methods).
- Caterina Urban served as Reviewer for POPL 2020 (Symposium on Principles of Programming Languages).

10.1.4. Journal

10.1.4.1. Member of the Editorial Boards

- Jérôme Feret serves as a Member of the Editorial Board of the Frontiers in Genetics journal and the Open Journal of Modeling and Simulation.
- Jérôme Feret served as co-Editor of an Issue of the Theoretical Computer Science journal, that is composed of papers from SASB 2016 and VEMP 2016, and appeared in 2019.
- Jérôme Feret served as co-Editor of an Issue of the IEEE/ACM Transactions on Computational Biology and Bioinformatics, that is composed of papers from CMSB 2017, and appeared in 2019.

10.1.4.2. Reviewer - Reviewing Activities

- Jérôme Feret served as a Reviewer for Fundamenta Informaticæ.
- Jérôme Feret served as a Reviewer for Theoretical Computer Science.
- Jérôme Feret served as a Reviewer for Transactions on Computational Biology and Bioinformatics.

10.1.5. Invited Talks

- Caterina Urban was invited to give a talk at the Gran Sasso Science Institute, L'Aquila, Italy.
- Caterina Urban was invited to give a talk at SAS 2019 (Static Analysis Symposium) [13].

10.1.6. Leadership within the Scientific Community

- Xavier Rival is a member of the IFIP Working Group 2.4 on Software implementation technology.

10.1.7. Scientific Expertise

- Jérôme Feret reviewed 5 proposals for the Prin 2017 (Research Projects of National Relevance) Program (Italian Ministry of Education and Research)

10.1.8. Research Administration

- Jérôme Feret and Xavier Rival are members of the Laboratory Council of DIENS.
- Jérôme Feret is member of the PhD Review Committee (CSD) of Inria Paris.
- Jérôme Feret is deputy head of study of the Department of Computer Science of École normale supérieure.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence:

- Marc Chevalier, Mathematics, 40h, L1, FDV Bachelor program (Frontiers in Life Sciences (FdV)), Université Paris-Descartes, France.
- Jérôme Feret and Xavier Rival (lectures), and Marc Chevalier (tutorials), "Semantics and Application to Verification", 36h, L3, at École Normale Supérieure, France.
- Xavier Rival, "Introduction to Static Analysis", 8h, L3, at École des Mines de Paris, France.
- Xavier Rival, "Functional Programming", 21h, L3, Bachelor Programme at at École Polytechnique, France.

Master:

- Xavier Rival, "Introduction to Static Analysis", 12h, Internet of Things Master (retraining curriculum, EXED), France.

- Cezara Drăgoi, Jérôme Feret, Antoine Miné, and Xavier Rival, “Abstract Interpretation: application to verification and static analysis”, 72h, M2. Parisian Master of Research in Computer Science (MPRI), France.
- Vincent Danos and Jérôme Feret (with Jean Krivine), Rule-based Modelling, 24h, M1. Interdisciplinary Approaches to Life Science (AIV), Master Program, Université Paris-Descartes, France.

Doctorat:

- Jérôme Feret, “Abstract Interpretation of Models of Intracellular Signalling Pathways", 3h, CNRS summer school “Formal modeling of Biological Regulation Networks”, June 2019, Île de Porquerolles, France.

10.2.2. Supervision

- PhD defended: Andreea Beica, Abstraction of Biochemical Reaction Networks, started in 2015 and supervised by Vincent Danos
- PhD defended: Hugo Illous, Relational Shape Abstraction Based on Separation Logic, started in 2015 and supervised by Xavier Rival and Matthieu Lemerre (CEA)
- PhD in progress
- PhD in progress: Marc Chevalier, Static analysis of Security Properties in Critical Embedded Software, started in 2017 and supervised by Jérôme Feret
- PhD in progress: Albin Salazar, Formal derivation of discrete models with separated time-scales, started in 2019 and supervised by Jérôme Feret
- PhD in progress: Olivier Nicole, Verification of micro-kernels, started in 2018 and supervised by Xavier Rival and Matthieu Lemerre (CEA)

10.2.3. Juries

- Jérôme Feret served as a member for the Jury of the PhD of Andreea Beica (Defense on the 12th of June 2019)
- Xavier Rival served as a member for the Habilitation Thesis Committee of Chien Chung Huang (Defense on the 28th of January 2019)
- Xavier Rival is serving as a member of the Review Committee for Vincenzo Arceri at University of Verona (to be defended in early 2020).
- Caterina Urban served as a member of the Jury of the PhD of Emilio Incerto at the Gran Sasso Science Institute in Italy (defense on the 5th of April 2019).
- Cezara Drăgoi servered as a member for the Jury of the PhD of Hugo Illous, Relational Shape Abstraction Based on Separation Logic (defence on the 16th avril 2019).

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- Cezara Drăgoi and Xavier Rival are elected members of the Inria Commision of Evaluation
- Jérôme Feret is a member of the jury of the SIF - Gilles Kahn PhD Award.
- Xavier Rival is member of the “Bureau du comité des projets” since October 2018.
- Xavier Rival is Chair of the Hiring Committee for Inria researchers at the center of Paris (CRCN) for 2019.
- Cezara Drăgoi served in the jury for part-time positions at LiX, January 2019.
- Cezara Drăgoi served in the Inria jury for CES - delegations and Phd scholarships in 2019.

10.3.2. Articles and contents

- Xavier Rival has been working with Kwangkeun Yi on the writing of a book that should serve as an introduction to the field of static analysis, for students and engineers, and this book is expected to be released by MIT Press in January 2020 [16].

10.3.3. Creation of media or tools for science outreach

Xavier Rival has given a talk as part of the “Mon équipe en 180 secondes” series at Inria Paris.

11. Bibliography

Major publications by the team in recent years

[1] J. BERTRANE, P. COUSOT, R. COUSOT, J. FERET, L. MAUBORGNE, A. MINÉ, X. RIVAL.*Static Analysis and Verification of Aerospace Software by Abstract Interpretation*, in "Proceedings of the American Institute of Aeronautics and Astronautics (AIAA Infotech@Aerospace 2010)", Atlanta, Georgia, USA, American Institute of Aeronautics and Astronautics, 2010

[2] B. BLANCHET, P. COUSOT, R. COUSOT, J. FERET, L. MAUBORGNE, A. MINÉ, D. MONNIAUX, X. RIVAL.*A Static Analyzer for Large Safety-Critical Software*, in "Proceedings of the ACM SIGPLAN 2003 Conference on Programming Language Design and Implementation (PLDI'03)", ACM Press, June 7–14 2003, p. 196–207

[3] A. BOUAJJANI, C. DRAGOI, C. ENEA, M. SIGHIREANU.*On inter-procedural analysis of programs with lists and data*, in "Proceedings of the 32nd ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI 2011, San Jose, CA, USA, June 4-8, 2011", 2011, p. 578–589, http://doi.acm.org/10.1145/1993498.1993566

[4] P. COUSOT.*Constructive Design of a Hierarchy of Semantics of a Transition System by Abstract Interpretation*, in "Theoretical Computer Science", 2002, vol. 277, n^o 1–2, p. 47–103

[5] J. FERET, V. DANOS, J. KRIVINE, R. HARMER, W. FONTANA.*Internal coarse-graining of molecular systems*, in "Proceeding of the national academy of sciences", Apr 2009, vol. 106, n^o 16

[6] L. MAUBORGNE, X. RIVAL.*Trace Partitioning in Abstract Interpretation Based Static Analyzers*, in "Proceedings of the 14th European Symposium on Programming (ESOP'05)", M. SAGIV (editor), Lecture Notes in Computer Science, Springer-Verlag, 2005, vol. 3444, p. 5–20

[7] A. MINÉ.*The Octagon Abstract Domain*, in "Higher-Order and Symbolic Computation", 2006, vol. 19, p. 31–100

[8] X. RIVAL.*Symbolic Transfer Functions-based Approaches to Certified Compilation*, in "Conference Record of the 31st Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages", ACM Press, New York, United States, 2004, p. 1–13

Publications of the year

Doctoral Dissertations and Habilitation Theses

[9] A. BEICA.*Abstractions of Biochemical Reaction Networks*, PSL University, June 2019, https://hal.inria.fr/tel-02411487

[10] H. ILLOUS.*Abstract Heap Relations for a Compositional Shape Analysis*, Ecole Normale Supérieure, April 2019, https://hal.inria.fr/tel-02399767

Articles in International Peer-Reviewed Journal

[11] Y. KO, X. RIVAL, S. RYU.*Weakly Sensitive Analysis for JavaScript Object-Manipulating Programs*, in "International Journal on Software - Practice and Experience", January 2019 [*DOI :* 10.1002/SPE], https://hal.archives-ouvertes.fr/hal-02399944

[12] W. LEE, H. YU, X. RIVAL, H. YANG.*Towards Verified Stochastic Variational Inference for Probabilistic Programs*, in "Proc. ACM Program. Lang", 2020, vol. 16, forthcoming [*DOI :* 10.1145/3371084], https://hal.archives-ouvertes.fr/hal-02399922

Invited Conferences

[13] C. URBAN.*Static Analysis of Data Science Software*, in "SAS 2019 - 26th Static Analysis Symposium", Porto, Portugal, B.-Y. E. CHANG (editor), Springer, October 2019, p. 17-23 [*DOI :* 10.1007/978-3-030-32304-2_2], https://hal.inria.fr/hal-02397699

International Conferences with Proceedings

[14] P. BOUTILLIER, I. CRISTESCU, J. FERET.*Counters in Kappa: Semantics, Simulation, and Static Analysis*, in "ESOP 2019 - 28th European Symposium on Programming", Prague, Czech Republic, Springer, April 2019, p. 176-204 [*DOI :* 10.1007/978-3-030-17184-1_7], https://hal.inria.fr/hal-02397876

[15] M. CHEVALIER, J. FERET.*Sharing Ghost Variables in a Collection of Abstract Domains*, in "VMCAI 2020 - 21st International Conference on Verification, Model Checking, and Abstract Interpretation", New Orleans, LA, United States, Proceedings of the 21st International Conference on Verification, Model Checking, and Abstract Interpretation (VMCAI 2020), January 2020, https://hal.inria.fr/hal-02378809

Scientific Books (or Scientific Book chapters)

[16] X. RIVAL, K. YI.*Introduction to Static Analysis*, MIT Press, February 2020, https://hal.archives-ouvertes.fr/hal-02402597

Other Publications

[17] A. DAMIAN, C. DRAGOI, A. MILITARU, J. WIDDER.*Communication-closed asynchronous protocols*, January 2019, working paper or preprint, https://hal.inria.fr/hal-01991415

[18] C. DRAGOI, J. WIDDER, D. ZUFFEREY.*Executable Rounds: a Programming Abstraction for Fault-Tolerant Protocols*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02317446

[19] C. URBAN, M. CHRISTAKIS, V. WÜSTHOLZ, F. ZHANG.*Perfectly Parallel Fairness Certification of Neural Networks*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02404036

References in notes

[20] A. BEICA, V. DANOS.*Synchronous Balanced Analysis*, in "Proceedings of the International Workshop on Hybrid Systems Biology", Springer-Verlag, Berlin, Germany, September 2016, p. 85-94, https://hal.archives-ouvertes.fr/hal-01974696

[21] P. COUSOT.*Constructive design of a hierarchy of semantics of a transition system by abstract interpretation*, in "Electr. Notes Theor. Comput. Sci.", 1997, vol. 6, p. 77–102, http://dx.doi.org/10.1016/S1571-0661(05)80168-9

[22] P. COUSOT, R. COUSOT.*Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints*, in "Conference Record of the Fourth Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages", ACM Press, New York, United States, 1977, p. 238–252

[23] Y. KO, X. RIVAL, S. RYU.*Weakly Sensitive Analysis for Unbounded Iteration over JavaScript Objects*, in "APLAS 2017 - 15th Asian Symposium on Programming Languages and Systems", Suzhou, China, LNCS, Springer, November 2017, vol. 10695, p. 148-168 [*DOI :* 10.1007/978-3-319-71237-6_8], https://hal.inria.fr/hal-01648680

[24] G. TERRADOT, A. BEICA, A. Y. WEISSE, V. DANOS.*Survival of the Fattest: Evolutionary Trade-offs in Cellular Resource Storage*, in "Electronic Notes in Theoretical Computer Science", April 2018, vol. 335, p. 91-112 [*DOI :* 10.1016/J.ENTCS.2018.03.010], https://hal.archives-ouvertes.fr/hal-01976385

[25] P. THOMAS, G. TERRADOT, V. DANOS, A. Y. WEISSE.*Sources, propagation and consequences of stochasticity in cellular growth*, in "Nature Communications", December 2018, vol. 9, n^{o} 1 [*DOI :* 10.1038/S41467-018-06912-9], https://hal.archives-ouvertes.fr/hal-01976406

[26] A. Y. WEISSE, D. A. OYARZUN, V. DANOS, P. S. SWAIN.*Mechanistic links between cellular trade-offs, gene expression, and growth*, in "Proceedings of the National Academy of Sciences of the United States of America ", March 2015, vol. 112, n^{o} 9, p. E1038-E1047 [*DOI :* 10.1073/PNAS.1416533112], https://hal.archives-ouvertes.fr/hal-01976394

Project-Team ARAMIS

Algorithms, models and methods for images and signals of the human brain

IN COLLABORATION WITH: Institut du Cerveau et de la Moelle Epinière

IN PARTNERSHIP WITH:
CNRS

INSERM

Sorbonne Université (UPMC)

RESEARCH CENTER
Paris

THEME
Computational Neuroscience and Medicine

Table of contents

Project-Team ARAMIS

Creation of the Team: 2012 October 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A3.4. - Machine learning and statistics
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.4. - Optimization and learning
A5.3. - Image processing and analysis
A5.4.4. - 3D and spatio-temporal reconstruction
A5.9. - Signal processing
A9. - Artificial intelligence
A9.2. - Machine learning
A9.3. - Signal analysis
A9.6. - Decision support

Other Research Topics and Application Domains:

B2. - Health
B2.2.6. - Neurodegenerative diseases
B2.6.1. - Brain imaging

1. Team, Visitors, External Collaborators

Research Scientists

Olivier Colliot [Team leader, CNRS, Senior Researcher, HDR]
Ninon Burgos [CNRS, Researcher]
Fabrizio de Vico Fallani [Inria, Researcher, HDR]
Stanley Durrleman [Inria, Senior Researcher, HDR]
Vincent Henry [Inria, Advanced Research Position, from Mar 2019]

Faculty Members

Benjamin Charlier [Univ de Montpellier, Associate Professor, until Aug 2019]
Didier Dormont [Univ Pierre et Marie Curie, Professor]

Technical Staff

Stephane Epelbaum [Hospital neurologist, Assistance publique/Hôpitaux de Paris, recipient of "Poste d'accueil" Inria since Nov 2019]
Marie-Constance Corsi [Inria, Engineer, until May 2019]
Mauricio Diaz Melo [Inria, Engineer]
Juliana Gonzalez Astudillo [Inria, Engineer, from Feb 2019 until Sep 2019]
Etienne Maheux [Inria, Engineer, from Nov 2019]
Arnaud Marcoux [Institut du Cerveau et de la Moelle Epinière, Engineer]
Benoit Martin [Inria, Engineer, until Oct 2019]
Thomas Nedelec [Inria, Engineer, from Sep 2019]
Alexandre Routier [Inria, Engineer]
Arnaud Valladier [Inria, Engineer]
Paul Vernhet [Inria, Engineer, from May 2019 until Aug 2019]

PhD Students

Manon Ansart [INSERM, PhD Student]
Wen Wei [Inria, PhD Student]
Federica Cacciamani [Institut du Cerveau et de la Moelle Epinière, PhD Student]
Giulia Bassignana [INSERM, PhD Student]
Simona Bottani [Inria, PhD Student]
Alexandre Bône [Univ Pierre et Marie Curie, PhD Student]
Tiziana Cattai [Inria, PhD Student]
Raphaël Couronné [Inria, PhD Student]
Johann Faouzi [Institut du Cerveau et de la Moelle Epinière, PhD Student]
Juliana Gonzalez Astudillo [Sorbonne Université, PhD Student, from Oct 2019]
Fanny Grosselin [MyBrainTechnologies, PhD Student, until May 2019]
Virgilio Kmetzsch Rosa E Silva [Inria, PhD Student, from Oct 2019]
Igor Koval [INSERM, PhD Student]
Thomas Lartigue [Inria, PhD Student]
Maxime Louis [Univ Pierre et Marie Curie, PhD Student, until Sep 2019]
Pascal Lu [Univ Pierre et Marie Curie, PhD Student, until Sep 2019]
Clement Mantoux [Inria, PhD Student, from Nov 2019]
Jorge Samper Gonzalez [Institut du Cerveau et de la Moelle Epinière, PhD Student, until Apr 2019]
Sophie Skriabine [Inria, PhD Student, from Oct 2019]
Elina Thibeau Sutre [Univ Pierre et Marie Curie, PhD Student]
Paul Vernhet [Inria, PhD Student, from Oct 2019]
Junhao Wen [China Scholarship Council, PhD Student, until Oct 2019]

Post-Doctoral Fellows

Marie-Constance Corsi [Institut du Cerveau et de la Moelle Epinière, Post-Doctoral Fellow, from Jun 2019]
Baptiste Couvy Duchesne [University of Queensland, Post-Doctoral Fellow]
Alexis Guyot [Institut du Cerveau et de la Moelle Epinière, Post-Doctoral Fellow]
Jorge Samper Gonzalez [INSERM, Post-Doctoral Fellow, from May 2019]

Visiting Scientist

Bruno Jedynak [University of Oregon, from Jun 2019 until Jul 2019]

Administrative Assistants

Helene Bessin Rousseau [Inria, Administrative Assistant, until Nov 2019]
Helene Milome [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Context

ARAMIS is an Inria project-team within the Brain and Spinal cord Institute (ICM - http://www.icm-institute.org) at the Pitié-Salpêtrière hospital in Paris. ARAMIS was created as a team of the Inria Paris Center in 2012 and became a project-team in 2014. ARAMIS has a joint affiliation to Inria, CNRS, Inserm and Sorbonne University.

The **Pitié-Salpêtrière hospital** is the largest adult hospital in Europe. It is a leading center for neurological diseases: in terms of size (around 20,000 neurological patients each year), level of clinical expertise and quality of the technical facilities. Created in 2010, the **Brain and Spinal cord Institute (ICM)** gathers all research activities in neuroscience and neurology of the Pitié-Salpêtrière hospital. The ICM is both a private foundation and a public research unit (affiliated to CNRS, Inserm and University Pierre and Marie Curie). It hosts 25 research teams as well as various high level technical facilities (neuroimaging, genotyping/sequencing, cell culture, cellular imaging, bioinformatics ...), and gathers over 600 personnel. In addition, the ICM hosts one of the six IHU (*Instituts Hospitalo-Universitaires*), which are 10-year research programs funded for 55M euros each.

ARAMIS is thus located both within a leading neuroscience institute and within a large hospital. This unique position has several advantages: direct contact with neuroscientists and clinicians allows us to foresee the emergence of new problems and opportunities for new methodological developments, provides access to unique datasets, and eases the transfer of our results to clinical research and clinical practice.

2.2. General aim

The ARAMIS team is devoted to the design of **computational, mathematical and statistical approaches for the analysis of multimodal patient data**, with an emphasis on neuroimaging data. The core methodological domains of our team are: statistical and machine learning, statistical modeling of complex geometric data, connectivity and network analysis. These new approaches are applied to clinical research in neurological diseases in collaboration with other teams of the ICM, clinical departments of the Pitié-Salpêtrière hospital and external partners. **The team has a pluridisciplinary composition**, bringing together researchers in mathematics, computer science and engineering (N. Burgos, O. Colliot, F. De Vico Fallani, S. Durrleman) and clinicians (D. Dormont, S. Epelbaum). This general endeavor is addressed within the five following main objectives.

3. Research Program

3.1. From geometrical data to multimodal imaging

Brain diseases are associated to alterations of brain structure that can be studied in vivo using anatomical and diffusion MRI. The anatomy of a given subject can be represented by sets of anatomical surfaces (cortical and subcortical surfaces) and curves (white matter tracks) that can be extracted from anatomical and diffusion MRI respectively. We aim to develop approaches that can characterize the variability of brain anatomy within populations of subjects. To that purpose, we propose methods to estimate population atlases that provide an average model of a population of subjects together with a statistical model of their variability. Finally, we aim to introduce representations that can integrate geometrical information (anatomical surfaces, white matter tracts) together with functional (PET, ASL, EEG/MEG) and microstructural information.

3.2. Models of brain networks

Functional imaging techniques (EEG, MEG and fMRI) allow characterizing the statistical interactions between the activities of different brain areas, i.e. functional connectivity. Functional integration of spatially distributed brain regions is a well-known mechanism underlying various cognitive tasks, and is disrupted in brain disorders. Our team develops a framework for the characterization of brain connectivity patterns, based on connectivity descriptors from the theory of complex networks. More specifically, we propose analytical tools to infer brain networks, chacterize their structure and integrate multiple networks (for instance from multiple frequency bands or multiple modalities). The genericity of this approach allows us to apply it to various types of data including functional and structural neuroimaging, as well as genomic data.

3.3. Spatiotemporal modeling from longitudinal data

Longitudinal data sets are collected to capture variable temporal phenomena, which may be due to ageing or disease progression for instance. They consist in the observation of several individuals, each of them being observed at multiple points in time. The statistical exploitation of such data sets is notably difficult since data of each individual follow a different trajectory of changes and at its own pace. This difficulty is further increased if observations take the form of structured data like images or measurements distributed at the nodes of a mesh, and if the measurements themselves are normalized data or positive definite matrices for which usual linear operations are not defined. We aim to develop a theoretical and algorithmic framework for learning typical trajectories from longitudinal data sets. This framework is built on tools from Riemannian geometry to describe trajectories of changes for any kind of data and their variability within a group both in terms of the direction of the trajectories and pace.

3.4. Decision support systems

We then aim to develop tools to assist clinical decisions such as diagnosis, prognosis or inclusion in therapeutic trials. To that purpose, we leverage the tools developed by the team, such as multimodal representations, network indices and spatio-temporal models which are combined with advanced classification and regression approaches. We also dedicate strong efforts to rigorous, transparent and reproducible validation of the decision support systems on large clinical datasets.

3.5. Clinical research studies

Finally, we aim to apply advanced computational and statistical tools to clinical research studies. These studies are often performed in collaboration with other researchers of the ICM, clinicians of the Pitié -Salpêtrière hospital or external partners. Notably, our team is very often involved ”ex-ante” in clinical research studies. As co-investigators of such studies, we contribute to the definition of objectives, study design and definition of protocols. This is instrumental to perform clinically relevant methodological development and to maximize their medical impact. A large part of these clinical studies were in the field of dementia (Alzheimer's disease, fronto-temporal dementia). Recently, we expanded our scope to other neurodegenerative diseases (Parkinson's disease, multiple sclerosis).

4. Application Domains

4.1. Introduction

We develop different applications of our new methodologies to brain pathologies, mainly neurodegenerative diseases. These applications aim at:

- better understanding the pathophysiology of brain disorders;
- designing systems to support clinical decisions such as diagnosis, prognosis and design of clinical trials;
- developing brain computer interfaces for clinical applications.

4.2. Understanding brain disorders

Computational and statistical approaches have the potential to help understand the pathophysiology of brain disorders. We first aim to contribute to better understand the relationships between pathological processes, anatomical and functional alterations, and symptoms. Moreover, within a single disease, there is an important variability between patients. The models that we develop have the potential to identify more homogeneous disease subtypes, that would constitute more adequate targets for new treatments. Finally, we aim to establish the chronology of the different types of alterations. We focus these activities on neurodegeneratives diseases: dementia (Alzheimer's disease, fronto-temporal dementia), Parkinson's disease, multiple sclerosis.

4.3. Supporting clinical decisions

We aim to design computational tools to support clinical decisions, including diagnosis, prognosis and the design of clinical trials. The differential diagnosis of neurodegenerative diseases can be difficult. Our tools have the potential to help clinicians by providing automated classification that can integrate multiple types of data (clinical/cognitive tests, imaging, biomarkers). Predicting the evolution of disease in individual patients is even more difficult. We aim to develop approaches that can predict which alterations and symptoms will occur and when. Finally, new approaches are needed to select participants in clinical trials. Indeed, it is widely recognized that, to have a chance to be successful, treatments should be administered at a very early stage.

4.4. Brain computer interfaces for clinical applications

A brain computer interface (BCI) is a device aiming to decode brain activity, thus creating an alternate communication channel between a person and the external environment. BCI systems can be categorized on the basis of the classification of an induced or evoked brain activity. The central tenet of a BCI is the capability to distinguish different patterns of brain activity, each being associated to a particular intention or mental task. Hence adaptation, as well as learning, is a key component of a BCI because users must learn to modulate their brainwaves to generate distinct brain patterns. Usually, a BCI is considered a technology for people to substitute some lost functions. However, a BCI could also help in clinical rehabilitation to recover motor functions. Indeed, in current neuroscience-based rehabilitation it is recognized that protocols based on mental rehearsal of movements (like motor imagery practicing) are a way to access the motor system because they can induce an activation of sensorimotor networks that were affected by lesions. Hence, a BCI based on movement imagery can objectively monitor patients' progress and their compliance with the protocol, monitoring that they are actually imagining movements. It also follows that feedback from such a BCI can provide patients with an early reinforcement in the critical phase when there is not yet an overt sign of movement recovery.

5. Highlights of the Year

5.1. Highlights of the Year

- F. De Vico Fallani was awarded an ERC Consolidator Grant.
- N. Burgos was awarded a Chair at the Paris Artificial Intelligence Research Institute (PRAIRIE).
- O. Colliot was awarded a Chair at the Paris Artificial Intelligence Research Institute (PRAIRIE).
- S. Durrleman was awarded a Chair at the Paris Artificial Intelligence Research Institute (PRAIRIE).
- S. Epelbaum holds a "Poste d'accueil" Inria since november 2019.

5.1.1. Awards

- N. Burgos received the Cor Baayen Young Researcher Award from ERCIM.
- V. Debavelaere was awarded a best paper award at the MICCAI conference in Shenzhen .

BEST PAPERS AWARDS :

[32]

V. DEBAVELAERE, A. BÔNE, S. DURRLEMAN, S. ALLASSONNIÈRE.*Clustering of longitudinal shape data sets using mixture of separate or branching trajectories*, in "Medical Image Computing and Computer Assisted Intervention – MICCAI", Shenzhen, China, Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, October 2019, p. 66-74, https://hal.archives-ouvertes.fr/hal-02103355

6. New Software and Platforms

6.1. Brain Networks Toolbox

KEYWORDS: Neuroimaging - Medical imaging

FUNCTIONAL DESCRIPTION: Brain Networks Toolbox is an open-source package of documented routines implementing new graph algorithms for brain network analysis. It mainly contains Matlab code of new methods developed by the team and associated to publications (e.g., brain network thresholding, extraction of the information redundancy, node accessibility, etc). It requires, as input, adjacency matrices representing brain connectivity networks. Thus, it is independent on the specific approach used to construct brain networks and it can be used to extract network properties from any neuroimaging modality in healthy and diseased subjects.

- Participants: Fabrizio de Vico Fallani, Jeremy Guillon and Mario Chavez
- Contact: Fabrizio de Vico Fallani
- URL: https://github.com/brain-network/bnt

6.2. Deformetrica

KEYWORDS: 3D modeling - C++ - Automatic Learning - Mesh - Anatomy - Image analysis

SCIENTIFIC DESCRIPTION: Deformetrica is a software for the statistical analysis of 2D and 3D shape data. It essentially computes deformations of the 2D or 3D ambient space, which, in turn, warp any object embedded in this space, whether this object is a curve, a surface, a structured or unstructured set of points, or any combination of them.

Deformetrica comes with two applications:

- Registration, which computes the best possible deformation between two sets of objects,

- Atlas construction, which computes an average object configuration from a collection of object sets, and the deformations from this average to each sample in the collection.

Deformetrica has very little requirements about the data it can deal with. In particular, it does not require point correspondence between objects!

- Participants: Alexandre Routier, Ana Fouquier, Barbara Gris, Benjamin Charlier, Cédric Doucet, Joan Alexis Glaunès, Marcel Prastawa, Michael Bacci, Pietro Gori and Stanley Durrleman
- Partners: University of Utah - Université de Montpellier 2 - Université Paris-Descartes
- Contact: Stanley Durrleman
- URL: http://www.deformetrica.org/

6.3. Clinica

KEYWORDS: Neuroimaging - Brain MRI - MRI - Clinical analysis - Image analysis - Machine learning

SCIENTIFIC DESCRIPTION: Clinica is a software platform for multimodal brain image analysis in clinical research studies. It makes it easy to apply advanced analysis tools to large scale clinical studies. For that purpose, it integrates a comprehensive set of processing tools for the main neuroimaging modalities: currently MRI (anatomical, functional, diffusion) and PET, in the future, EEG/MEG. For each modality, Clinica allows to easily extract various types of features (regional measures, parametric maps, surfaces, curves, networks). Such features are then subsequently used as input of machine learning, statistical modeling, morphometry or network analysis methods. Processing pipelines are based on combinations of freely available tools developed by the community. It provides an integrated data management specification to store raw and processing data. Clinica is written in Python. It uses the Nipype system for pipelining. It combines widely-used software for neuroimaging data analysis (SPM, Freesurfer, FSL, MRtrix...), morphometry (Deformetrica), machine learning (Scikit-learn) and the BIDS standard for data organization.

FUNCTIONAL DESCRIPTION: Clinica is a software platform for multimodal brain image analysis in clinical research studies. It makes it easy to apply advanced analysis tools to large scale clinical studies. For that purpose, it integrates a comprehensive set of processing tools for the main neuroimaging modalities: currently MRI (anatomical, functional, diffusion) and PET, in the future, EEG/MEG. For each modality, Clinica allows to easily extract various types of features (regional measures, parametric maps, surfaces, curves, networks). Such features are then subsequently used as input of machine learning, statistical modeling, morphometry or network analysis methods. Clinica also provides an integrated data management specification to store raw and processing data. Overall, Clinica helps to: i) apply advanced analysis tools to clinical research studies, ii) easily share data and results, iii) make research more reproducible.

NEWS OF THE YEAR: - Three clinical studies made with Clinica Clinica : Bertrand et al, JAMA Neurology, 2018 , Jacquemont et al, Neurobiol Aging, 2017, Wen et al, JNNP, 2018 - Clinica presented at OHBM 2018 conference - Clinica was the support for the tutorial "Pattern Recognition for Neuroimaging" at OHBM 2018

- Participants: Jeremy Guillon, Thomas Jacquemont, Pascal Lu, Arnaud Marcoux, Tristan Moreau, Alexandre Routier, Jorge Samper Gonzalez, Junhao Wen, Olivier Colliot, Stanley Durrleman, Michael Bacci, Simona Bottani, Ninon Burgos, Sabrina Fontanella, Pietro Gori, Mauricio Diaz and Elina Thibeau-Sutre

- Partners: Institut du Cerveau et de la Moelle épinière (ICM) - CNRS - INSERM - UPMC
- Contact: Olivier Colliot
- Publications: Amyloidosis and neurodegeneration result in distinct structural connectivity patterns in mild cognitive impairment - Yet Another ADNI Machine Learning Paper? Paving The Way Towards Fully-reproducible Research on Classification of Alzheimer's Disease - Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data - Neurite density is reduced in the presymptomatic phase of C9orf72 disease - Early cognitive, structural and microstructural changes in c9orf72 presymptomatic carriers before 40 years of age
- URL: http://www.clinica.run

6.4. Platforms

6.4.1. Platform Brain-computer interface

Our team coordinates the developments of the Brain-Computer Interface (BCI) platform at the Centre EEG/MEG of the neuroimaging core facility of the ICM. Several projects, including our NETBCI NSF/NIH/ANR and ATTACK Big-brain theory funded projects, as well as experiments by different researchers of the Institute, are currently being run. To reinforce the impact of the platform we have recently recruited an engineer (J. Gonzalez-Astudillo) and a master student (Tristan Venot) for the software and technical development.

7. New Results

7.1. Predicting PET-derived Demyelination from Multimodal MRI using Sketcher-Refiner Adversarial Training for Multiple Sclerosis

Participants: Wen Wei, Emilie Poirion, Benedetta Bodini, Stanley Durrleman, Nicholas Ayache, Bruno Stankoff, Olivier Colliot [Correspondant].

Multiple sclerosis (MS) is the most common demyelinating disease. In MS, demyelination occurs in the white matter of the brain and in the spinal cord. It is thus essential to measure the tissue myelin content to understand the physiopathology of MS, track progression and assess treatment efficacy. Positron emission tomography (PET) with [11C]PIB is a reliable method to measure myelin content in vivo. However, the availability of PET in clinical centers is limited. Moreover, it is expensive to acquire and invasive due to the injection of a radioactive tracer. By contrast, MR imaging is non-invasive, less expensive and widely available, but conventional MRI sequences cannot provide a direct and reliable measure of myelin. In this work, we therefore propose, to the best of our knowledge for the first time, a method to predict the PET-derived myelin content map from multimodal MRI. To that purpose, we introduce a new approach called Sketcher-Refiner generative adversarial networks (GANs) with specifically designed adversarial loss functions. The first network (Sketcher) generates global anatomical and physiological information. The second network (Refiner) refines and generates the tissue myelin content. A visual attention saliency map is also proposed to interpret the attention of neural networks. Our approach is shown to outperform the state-of-the-art methods in terms of image quality and myelin content prediction. Particularly, our prediction results show similar results to the PET-derived gold standard at both global and voxel-wise levels indicating the potential for clinical management of patients with MS.

More details in [25].

7.2. Reproducible evaluation of methods for predicting progression to Alzheimer's disease from clinical and neuroimaging data

Participants: Jorge Samper-González, Ninon Burgos, Simona Bottani, Marie-Odile Habert, Stéphane Epelbaum, Theodoros Evgeniou, Olivier Colliot [Correspondant].

Various machine learning methods have been proposed for predicting progression of patients with mild cognitive impairment (MCI) to Alzheimer's disease (AD) using neuroimaging data. Even though the vast majority of these works use the public dataset ADNI, reproducing their results is complicated because they often do not make available elements that are essential for reproducibility, such as selected participants and input data, image preprocessing and cross-validation procedures. Comparability is also an issue. Specially, the influence of different components like preprocessing, feature extraction or classification algorithms on the performance is difficult to evaluate. Finally, these studies rarely compare their results to models built from clinical data only, a critical aspect to demonstrate the utility of neuroimaging. In our previous work, 1, 2 we presented a framework for reproducible and objective classification experiments in AD, that included automatic conversion of ADNI database into the BIDS community standard, image preprocessing pipelines and machine learning evaluation. We applied this framework to perform unimodal classifications of T1 MRI and FDG-PET images. In the present paper, we extend this work to the combination of multimodal clinical and neuroimaging data. All experiments are based on standard approaches (namely SVM and random forests). In particular, we assess the added value of neuroimaging over using only clinical data. We first demonstrate that using only demographic and clinical data (gender, education level, MMSE, CDR sum of boxes, ADASCog) results in a balanced accuracy of 75% (AUC of 0.84). This performance is higher than that of standard neuroimaging-based classifiers. We then propose a simple trick to improve the performance of neuroimaging-based classifiers: training from AD patients and controls (rather than from MCI patients) improves the performance of FDG-PET classification by 5 percent points, reaching the level of the clinical classifier. Finally, combining clinical and neuroimaging data, prediction results further improved to 80% balanced accuracy and an AUC of 0.88). These prediction accuracies, obtained in a reproducible way, provide a base to develop on top of it and, to compare against, more sophisticated methods. All the code of the framework and the experiments is publicly available at https://github.com/aramis-lab/AD-ML.

More details in [34].

7.3. Disrupted core-periphery structure of multimodal brain networks in Alzheimer's disease

Participants: Jeremy Guillon, Mario Chavez, Federico Battiston, Yohan Attal, Valentina Corte, Michel Thiebaut de Schotten, Bruno Dubois, Denis Schwartz, Olivier Colliot, Fabrizio de Vico Fallani [Correspondant].

In Alzheimer's disease (AD), the progressive atrophy leads to aberrant network reconfigurations both at structural and functional levels. In such network reorganization, the core and peripheral nodes appear to be crucial for the prediction of clinical outcome because of their ability to influence large-scale functional integration. However, the role of the different types of brain connectivity in such prediction still remains unclear. Using a multiplex network approach we integrated information from DWI, fMRI, and MEG brain connectivity to extract an enriched description of the core-periphery structure in a group of AD patients and age-matched controls. Globally, the regional coreness-that is, the probability of a region to be in the multiplex core-significantly decreased in AD patients as result of a random disconnection process initiated by the neurodegeneration. Locally, the most impacted areas were in the core of the network-including temporal, parietal, and occipital areas-while we reported compensatory increments for the peripheral regions in the sensorimotor system. Furthermore, these network changes significantly predicted the cognitive and memory impairment of patients. Taken together these results indicate that a more accurate description of neurodegenerative diseases can be obtained from the multimodal integration of neuroimaging-derived network data.

More details in [20]

7.4. Network neuroscience for optimizing brain–computer interfaces

Participants: Fabrizio de Vico Fallani [Correspondant], Danielle Bassett.

Human-machine interactions are being increasingly explored to create alternative ways of communication and to improve our daily life. Based on a classification of the user's intention from the user's underlying neural activity, brain-computer interfaces (BCIs) allow direct interactions with the external environment while bypassing the traditional effector of the musculoskeletal system. Despite the enormous potential of BCIs, there are still a number of challenges that limit their societal impact, ranging from the correct decoding of a human's thoughts, to the application of effective learning strategies. Despite several important engineering advances, the basic neuroscience behind these challenges remains poorly explored. Indeed, BCIs involve complex dynamic changes related to neural plasticity at a diverse range of spatiotemporal scales. One promising antidote to this complexity lies in network science, which provides a natural language in which to model the organizational principles of brain architecture and function as manifest in its interconnectivity. Here, we briefly review the main limitations currently affecting BCIs, and we offer our perspective on how they can be addressed by means of network theoretic approaches. We posit that the emerging field of network neuroscience will prove to be an effective tool to unlock human-machine interactions.

More details in [13]

7.5. Quality Assessment of Single-Channel EEG for Wearable Devices

Participants: Fanny Grosselin, Xavier Navarro-Sune, Alessia Vozzi, Katerina Pandremmenou, Fabrizio de Vico Fallani, Yohan Attal, Mario Chavez [Correspondant].

The recent embedding of electroencephalographic (EEG) electrodes in wearable devices raises the problem of the quality of the data recorded in such uncontrolled environments. These recordings are often obtained with dry single-channel EEG devices, and may be contaminated by many sources of noise which can compromise the detection and characterization of the brain state studied. In this paper, we propose a classification-based approach to effectively quantify artefact contamination in EEG segments, and discriminate muscular artefacts. The performance of our method were assessed on different databases containing either artificially contaminated or real artefacts recorded with different type of sensors, including wet and dry EEG electrodes. Furthermore, the quality of unlabelled databases was evaluated. For all the studied databases, the proposed method is able to rapidly assess the quality of the EEG signals with an accuracy higher than 90

More details in [19]

7.6. Reduction of recruitment costs in preclinical AD trials. Validation of automatic pre-screening algorithm for brain amyloidosis

Participants: Manon Ansart [correspondant], Stéphane Epelbaum, Geoffroy Gagliardi, Olivier Colliot, Didier Dormont, Bruno Dubois, Harald Hampel, Stanley Durrleman.

We propose a method for recruiting asymptomatic Amyloid positive individuals in clinical trials, using a two-step process. We first select during a pre-screening phase a subset of individuals which are more likely to be amyloid positive based on the automatic analysis of data acquired during routine clinical practice, before doing a confirmatory PET-scan to these selected individuals only. This method leads to an increased number of recruitments and to a reduced number of PET-scans, resulting in a decrease in overall recruitment costs. We validate our method on 3 different cohorts, and consider 5 different classification algorithms for the pre-screening phase. We show that the best results are obtained using solely cognitive, genetic and socio-demographic features, as the slight increased performance when using MRI or longitudinal data is balanced by the cost increase they induce. We show that the proposed method generalizes well when tested on an independent cohort, and that the characteristics of the selected set of individuals are identical to the characteristics of a population selected in a standard way. The proposed approach shows how Machine Learning can be used effectively in practice to optimize recruitment costs in clinical trials.

More details in[7]

7.7. Learning low-dimensional representations of shape data sets with diffeomorphic autoencoders

Participants: Alexandre Bône [Correspondant], Maxime Louis, Olivier Colliot, Stanley Durrleman.

Contemporary deformation-based morphometry offers parametric classes of diffeomorphisms that can be searched to compute the optimal transformation that warps a shape into another, thus defining a similarity metric for shape objects. Extending such classes to capture the geometrical variability in always more varied statistical situations represents an active research topic. This quest for genericity however leads to computationally-intensive estimation problems. Instead, we propose in this work to learn the best-adapted class of diffeomorphisms along with its parametrization, for a shape data set of interest. Optimization is carried out with an auto-encoding variational inference approach, offering in turn a coherent model-estimator pair that we name diffeomorphic auto-encoder. The main contributions are: (i) an original network-based method to construct diffeomorphisms, (ii) a current-splatting layer that allows neural network architectures to process meshes, (iii) illustrations on simulated and real data sets that show differences in the learned statistical distributions of shapes when compared to a standard approach.

More details in[30]

7.8. Learning disease progression models with longitudinal data and missing values

Participants: Raphaël Couronné [correspondant], Marie Vidailhet, Jean-Christophe Corvol, Stéphane Lehéricy, Stanley Durrleman.

Statistical methods have been developed for the analysis of longitudinal data in neurodegenerative diseases. To cope with the lack of temporal markers- i.e. to account for subject-specific disease progression in regard to age- a common strategy consists in realigning the individual sequence data in time. Patient's specific trajectories can indeed be seen as spatiotemporal perturbations of the same normative disease trajectory. However, these models do not easily allow one to account for multimodal data, which more than often include missing values. Indeed, it is rare that imaging and clinical examinations for instance are performed at the same frequency in clinical protocols. Multimodal models also need to allow a different profile of progression for data with different structure and representation. We propose to use a generative mixed effect model that considers the progression trajectories as curves on a Riemannian Manifold. We use the concept of product manifold to handle multimodal data, and leverage the generative aspect of our model to handle missing values. We assess the robuste-ness of our methods toward missing values frequency on both synthetic and real data. Finally we apply our model on a real-world dataset to model Parkinson's disease progression from data derived from clinical examination and imaging.

More details in[31]

7.9. Learning the clustering of longitudinal shape data sets into a mixture of independent or branching trajectories

Participants: Vianney Debavelaere [correspondant], Stéphanie Allassonnière, Stanley Durrleman.

Given repeated observations of several subjects over time, i.e. a longitudinal data set, this work introduces a new model to learn a classification of the shapes progression in an unsupervised setting: we automatically cluster a longitudinal data set in different classes without labels. Our method learns for each cluster an average shape trajectory (or representative curve) and its variance in space and time. Representative trajectories are built as the combination of pieces of curves. This mixture model is flexible enough to handle independent trajectories for each cluster as well as fork and merge scenarios. The estimation of such non linear mixture models in high dimension is known to be difficult because of the trapping states effect that hampers the optimisation of cluster assignments during training. We address this issue by using a tempered version of the stochastic EM algorithm. Finally, we apply our algorithm on different data sets. First, synthetic data are used to show that a tempered scheme achieves better convergence. We then apply our method to different real data sets: 1D RECIST score used to monitor tumors growth, 3D facial expressions and meshes of the hippocampus. In particular, we show how the method can be used to test different scenarios of hippocampus atrophy in ageing by using an heteregenous population of normal ageing individuals and mild cognitive impaired subjects.

More details in[32]

7.10. Auto-encoding meshes of any topology with the current-splatting and exponentiation layers

Participants: Alexandre Bône [Correspondant], Olivier Colliot, Stanley Durrleman.

Deep learning has met key applications in image computing, but still lacks processing paradigms for meshes, i.e. collections of elementary geometrical parts such as points, segments or triangles. Meshes are both a powerful representation for geometrical objects, and a challenge for network architectures because of their inherent irregular structure. This work contributes to adapt classical deep learning paradigms to this particular type of data in three ways. First, we introduce the current-splatting layer which embeds meshes in a metric space, allowing the downstream network to process them without any assumption on their topology: they may be composed of varied numbers of elements or connected components, contain holes, or bear high levels of geometrical noise. Second, we adapt to meshes the exponentiation layer which, from an upstream image array, generates shapes with a diffeomorphic control over their topology. Third, we take advantage of those layers to devise a variational auto-encoding architecture, which we interpret as a generative statistical model that learns adapted low-dimensional representations for mesh data sets. An explicit norm-control layer ensures the correspondence between the latent-space Euclidean metric and the shape-space log-Euclidean one. We illustrate this method on simulated and real data sets, and show the practical relevance of the learned representation for visualization, classification and mesh synthesis.

More details in[29]

7.11. Riemannian Geometry Learning for Disease Progression Modelling

Participants: Maxime Louis, Raphael Couronne, Igor Koval, Benjamin Charlier, Stanley Durrleman.

The analysis of longitudinal trajectories is a longstanding problem in medical imaging which is often tackled in the context of Riemannian geometry: the set of observations is assumed to lie on an a priori known Riemannian manifold. When dealing with high-dimensional or complex data, it is in general not possible to design a Riemannian geometry of relevance. In this work, we perform Riemannian manifold learning in association with the statistical task of longitudinal trajectory analysis. After inference, we obtain both a submanifold of observations and a Riemannian metric so that the observed progressions are geodesics. This is achieved using a deep generative network, which maps trajectories in a low-dimensional Euclidean space to the observation space.

More details in[33]

7.12. How many patients are eligible for disease-modifying treatment in Alzheimer's disease? A French national observational study over 5 years.

Participants: Stéphane Epelbaum [Correspondant], Claire Paquet, Jacques Hugon, Julien Dumurgier, David Wallon, Didier Hannequin, Thérèse Jonveaux, Annick Besozzi, Stéphane Pouponneau, Caroline Hommet, Frédéric Blanc, Laetitia Berly, Adrien Julian, Marc Paccalin, Florence Pasquier, Julie Bellet, Claire Boutoleau-Bretonniere, Tiphaine Charriau, Olivier Rouaud, Olivier Madec, Aurélie Mouton, Renaud David, Samir Bekadar, Roxanne Fabre, Emmanuelle Liegey, Walter Deberdt, Philippe Robert, Bruno Dubois.

We aimed to study the epidemiology of the prodromal and mild stages of Alzheimer's disease (AD) patients who are eligible for clinical trials with disease-modifying therapies. We analyzed two large complementary databases to study the incidence and characteristics of this population on a nationwide scope in France from 2014 to 2018. The National Alzheimer Database contains data from 357 memory centres and 90 private neurologists. Data from 2014 to 2018 have been analyzed. Patients, 50–85 years old, diagnosed with AD who had an Mini-Mental State Exam (MMSE) score greater or equal to 20 were included. We excluded patients with mixed and non-AD neurocognitive disorders. Descriptive statistics of the population of interest was the primary measure. Results In the National Alzheimer Database, 550,198 patients were assessed. Among them, 72,174 (13.1%) were diagnosed with AD and had an MMSE greater or equal to 20. Using corrections for specificity of clinical diagnosis of AD, we estimated that about 50,000 (9.1%) had a prodromal or mild AD. In the combined electronic clinical records database of 11 French expert memory centres, a diagnosis of prodromal or mild AD, certified by the use of cerebrospinal fluid AD biomarkers, could be established in 195 (1.3%) out of 14 596 patients. AD was not frequently diagnosed at a prodromal or mild dementia stage in France in 2014 to 2018. Diagnosis rarely relied on a pathophysiological marker even in expert memory centres. National databases will be valuable to monitor early stage AD diagnosis efficacy in memory centres when a disease-modifying treatment becomes available. More details in [15]

7.13. EEG evidence of compensatory mechanisms in preclinical Alzheimer's disease

Participants: Sinead Gaubert, Federico Raimondo, Marion Houot, Marie-Constance Corsi, Jacobo Diego Sitt, Bertrand Hermann, Delphine Oudiette, Geoffroy Gagliardi, Marie Odile Habert, Bruno Dubois, Fabrizio de Vico Fallani, Hovagim Bakardjian, Stéphane Epelbaum [Correspondant].

Early biomarkers are needed to identify individuals at high risk of preclinical Alzheimer's disease and to better understand the pathophysiological processes of disease progression. Preclinical Alzheimer's disease EEG changes would be non-invasive and cheap screening tools and could also help to predict future progression to clinical Alzheimer's disease. However, the impact of amyloid-beta deposition and neurodegeneration on EEG biomarkers needs to be elucidated. We included participants from the INSIGHT-preAD cohort, which is an ongoing single-centre multimodal observational study that was designed to identify risk factors and markers of progression to clinical Alzheimer's disease in 318 cognitively normal individuals aged 70-85 years with a subjective memory complaint. We divided the subjects into four groups, according to their amyloid status (based on 18F-florbetapir PET) and neurodegeneration status (evidenced by 18F-fluorodeoxyglucose PET brain metabolism in Alzheimer's disease signature regions). The first group was amyloid-positive and neurodegeneration-positive, which corresponds to stage 2 of preclinical Alzheimer's disease. The second group was amyloid-positive and neurodegeneration-negative, which corresponds to stage 1 of preclinical Alzheimer's disease. The third group was amyloid-negative and neurodegeneration-positive, which corresponds to 'suspected non-Alzheimer's pathophysiology'. The last group was the control group, defined by amyloid-negative and neurodegeneration-negative subjects. We analysed 314 baseline 256-channel high-density eyes closed 1-min resting state EEG recordings. EEG biomarkers included spectral measures, algorithmic complexity and functional connectivity assessed with a novel information-theoretic measure, weighted symbolic mutual information. The most prominent effects of neurodegeneration on EEG metrics were localized in frontocentral regions with an increase in high frequency oscillations (higher beta and gamma power) and a decrease in low frequency oscillations (lower delta power), higher spectral entropy, higher complexity and increased functional connectivity measured by weighted symbolic mutual information in theta band. Neurodegeneration was associated with a widespread increase of median spectral frequency. We found a non-linear relationship between amyloid burden and EEG metrics in neurodegeneration-positive subjects, either following a U-shape curve for delta power or an inverted U-shape curve for the other metrics, meaning that EEG patterns are modulated differently depending on the degree of amyloid burden. This finding suggests initial compensatory mechanisms that are overwhelmed for the highest amyloid load. Together, these results indicate that EEG metrics are useful biomarkers for the preclinical stage of Alzheimer's disease.

More details in [17]

7.14. Latent class analysis identifies functional decline with Amsterdam IADL in preclinical Alzheimer's disease

Participants: Sarah-Christine Villeneuve, Marion Houot, Federica Cacciamani, Merike Verrijp, Marie Odile Habert, Bruno Dubois, Sietske Sikkes, Stéphane Epelbaum [Correspondant].

Trials in Alzheimer's disease (AD) now include participants at the earliest stages to prevent further decline. However, the lack of tools sensitive to subtle functional changes in early-stage AD hinders the development of new therapies as it is difficult to prove their clinical relevance. We assessed functional changes over three years in 289 elderly memory complainers from the Investigation of Alzheimer's Predictors in subjective memory complainers cohort using the Am-sterdam Instrumental-Activities-of-Daily-Living questionnaire (A-IADL-Q). No overall functional decline related to AD imaging markers was evidenced. However, five distinct classes of A-IADL-Q trajectories were identified. The largest class (212 [73.4%]) had stable A-IADL-Q scores over 3 years. A second group (23 [8.0%]) showed a persistent functional decline, higher amyloid load ($P<.0005$), and lower education ($P<.0392$). The A-IADL-Q identified a subtle functional decline in asymptomatic at-risk AD individuals. This could have important implications in the field of early intervention in AD

More details in [24]

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

8.1.1. Carthera

Participants: Stéphane Epelbaum [Correspondant], Alexandre Carpentier, Anne Bertrand, Marie Odile Habert.

Project title: Open label phase 1/2 study evaluating the safety and usefulness of transient opening of the blood-brain barrier using low intensity pulsed ultrasounds generated by the implantable device SONOCLOUD in patients with mild Alzheimer's disease

Started in 2016

Amount: 400 K€

Coordinator: Stéphane Epelbaum

Other partners: UPMC, AP-HP

Abstract: This project aims at opening the blood brain barrier (BBB) in 10 mild Alzheimer's disease patients in order to improve the clearance of beta-amyloid and tau deposits in their brain as suggested in mice models of the disease. This first in man study will evaluate the safety and efficacy of an implanted device, SONOCLOUD, to open the BBB 7 times in each participant. Efficacy will be evaluated on the ability of the method to decrease the amyloid load evidenced by AV45 Positron Emission Tomography (PET), increase the brain metabolism analyzed by Fluorodeoxyglucose PET and improve cognition. If successful, this study will pave the way for future trials in which drugs can be used in addition to BBB opening to maximize their effect.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. Health Data Hub

Participant: Stanley Durrleman.

Project acronym: Precise-PD-HDH

Project title: Modélisation et prédiction de la progression de la maladie de Parkinson

Duration: 1 year (pilot project)

Coordinator: Jean-Christophe Corvol

Other partners: Inserm, réseau NS-PARK, ICM

9.1.2. ANR

9.1.2.1. ANR-NIH-NSF CANDT

Participant: Fabrizio de Vico Fallani [Correspondant].

Project acronym: CANDT

Project title: Advancing neuroscientific discovery and training by lowering the barrier of entry to network neuroscience via open science

Duration: Oct 2019 - Sep 2023

Amount: 137k€

Coordinator: Fabrizio De Vico Fallani

Other partners: Indiana Univ., US; UPenn, US

Abstract: This project will use open science methods and cloud-computing, effectively lowering the barrier of entry to network neuroscience and increase the widespread availability of well-maintained and reproducible network neuroscience tools. We will use the platform brainlife.io as a digital marketplace for network neuroscience analysis methods; network neuroscience tools and software will be packaged into self-contained, standardized, reproducible Apps, shared with and modified by a burgeoning community of users, and seamlessly integrated into existing brainlife.io processing and analysis pipelines. This approach will engage both experts in network science, scientists from other domains, and users of the proposed methods. In addition, it will ensure correct implementation, a high level of reproducibility, and maximal reusability of network neuroscience methods. As a requirement, Apps will also be accompanied by links to primary sources, in-depth tutorials, and documentation, and worked-through examples, highlighting their correct usage and offering solutions for mitigating possible pitfalls. This proposed research lowers the barrier of entry to network neuroscience, standardizes the software sharing process, and provides a cloud-based repository of expertly-maintained network neuroscientific tools and software that is made available to the broader neuroscientific community.

9.1.2.2. ANR-NIH-NSF NETBCI

Participants: Fabrizio de Vico Fallani [Correspondant], Mario Chavez, Denis Schwartz.

Project acronym: NETBCI

Project title: Modeling and predicting brain-computer interface learning from dynamic networks

Duration: Avr 2016 - Avr 2020

Amount: 322k€

Coordinator: Fabrizio De Vico Fallani

Other partners: Complex system group, UPenn, USA

Abstract: This project will bring together expertise in computational and experimental neuroscience, signal processing and network science, statistics, modeling and simulation, to establish innovative methods to model and analyze temporally dynamic brain networks, and to apply these tools to develop predictive models of brain-computer interface (BCI) skill acquisition that can be used to improve performance. Leveraging experimental data and interdisciplinary theoretical techniques, this project will characterize brain networks at multiple temporal and spatial scales, and will develop models to predict the ability to control the BCI as well as methods to engineer BCI frameworks for adapting to neural plasticity. This project will enable a comprehensive understanding of the neural mechanisms of BCI learning, and will foster the design of viable BCI frameworks that improve usability and performance.

9.1.2.3. ANR-NIH-NSF HIPLAY7

Participants: Olivier Colliot [Correspondant], Marie Chupin, Stanley Durrleman, Anne Bertrand.

Project acronym: HIPLAY7

Project title: Hippocampal layers: advanced computational anatomy using very high resolution MRI at 7 Tesla in humans

Duration: Jan 2017 - Jan 2020

Amount: 770k€

Coordinator: Olivier Colliot and Pierre-François Van de Moortele

Other partners: University of Minnesota, Neurospin

Abstract: The overall goal of this proposal is to develop a coherent mathematical framework for computational anatomy of the internal structures of the hippocampus based on cutting edge MRI acquisition techniques at 7 Tesla. These mathematical and computational approaches are expected to significantly advance the field of computational anatomy of the human brain, breaking down the millimeter barrier of conventional brain morphometry and providing a coherent analysis framework for anatomical data at ultra-high spatial resolution.

9.1.2.4. ANR PREV-DEMALS

Participants: Olivier Colliot [Correspondant], Marie Chupin, Stanley Durrleman, Anne Bertrand.

Project acronym: PREV-DEMALS

Project title: Predict to prevent frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS)

Duration: Avr 2015 - Avr 2019

Amount: 487k€

Coordinator: Isabelle Le Ber

Other partners: ICM, AP-HP, CHR de Lille, CHU Limoges, CHU Rouen, Laboratory of Biomedical Imaging

Abstract: The project focuses on C9ORF72, the most frequent genetic form of frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS). Since 2006, major discoveries have helped elucidate the pathological bases and linked FTLD and ALS: 1) TDP-43 aggregates in neurons and 2) C9ORF72 mutations in both disorders. Two major pathological subtypes are now defined in FTLD, FTLD-TDP and FTLD-TAU. C9ORF72 mutations (associated to FTLD-TDP) are the most frequent genetic causes of FTLD (15%), FTLD-ALS (65%) and ALS (40%). No curative treatment actually exists, but therapeutics emerged against tau aggregation. The objectives of the project are to develop appropriate cognitive, brain imaging markers and peripheral biomarkers of the early phase of FTLD, to follow disease progression and to guide future targeted therapeutic trials. To address this questions, we will conduct a multimodal study (cognition, brain structural MRI, brain metabolism - FDG-PET) in C9ORF72 families. The cohort will be followed at 3-time points (M0, M18, M36). Longitudinal analyses will aim at characterizing the trajectory of decline across time. Brain structural changes will be evaluated by 1) morphometric analysis to assess global brain atrophy, cortical thickness and study of the cortical sulci; 2) functional connectivity analysis of resting-state MR data; 3) structural connectivity analysis of diffusion-weighted MRI. Brain metabolism will be evaluated with FDG-PET. We will use the most recent RNA sequencing technology to detect gene expression and RNA splicing alterations in lymphocytes of patients and presymptomatic carriers. The discovery of new markers involved in FTLD will have practical consequences for early and accurate diagnosis of FLD and ALS disease.

9.1.2.5. ANR IVMRS

Participants: Anne Bertrand [Correspondant], Alexandra Petiet, Mathieu Santin, Francesca Branzoli, Benoit Delatour, Marc Sanson.

Project acronym: IVMRS

Project title: Implantable miniaturized probe for In-vivo Magnetic Resonance Spectroscopy: Application to Murine models of Alzheimer's disease and Gliomas.

Duration: Oct 2016 - Oct 2020

Amount: 633k€

Coordinator: Luc Hebrard

Other partners: ICube - Unistra, Strasbourg; ISA Laboratory, Lyon; NYU School of Medicine, NY, USA.

Abstract: During the development of new therapeutics against brain diseases, the pre-clinical phase, i.e. the validation of treatment delivery, safety and efficacy in animal models of the disease, represents a crucial step. Magnetic Resonance Imaging (MRI) is a method of particular interest at this stage, as it provides non-invasive surrogate endpoints that can help selecting appropriate candidates during the process of drug development. Single Voxel Magnetic Resonance Spectroscopy (SVS) provides non-invasive, in-vivo quantitative measurements of brain metabolites, which reflects functional changes at the cellular and subcellular levels, and can be repeated longitudinally. As high-field MRI has become the benchmark in preclinical research on animal models, it appears possible to investigate the cerebral metabolomics changes in animals, and to use it as a surrogate marker in preclinical therapeutic trials. However, the number of relevant metabolites is much higher than the low number of measurable metabolites with conventional in-vivo high-field SVS. Moreover, considering also the subtle changes of these metabolites at the early stage of the disease, the use of conventional high-field SVS in preclinical studies remains strongly limited. The high volume of the Voxel-of-Interest (VOI), ranging from 10 to 30mm3, which is required to have a usable signal in conventional SVS, and the inherent variability of longitudinal SVS measurement due to the variable position of the VOI in the successive experiments, remain the two major issues when looking during time for small changes in metabolic concentrations and metabolites ratios in a specific small region of the animal brain. The IvMRS project aims at filling this gap by developing the first chronic implantable MRS micro-probe, minimally invasive, exhibiting very high signal sensitivity, and sharp spectral peaks, from sub-millimetric VOI. Such a probe will allow detecting a much higher number of metabolites than conventional in-vivo SVS. The probe will work at frequencies ranging from 300MHz to 500MHz in ultra-high field Magnetic Resonance Imaging scanners, 7T and 11.7T. It will embed a specific micro-coil antenna, a low-noise signal conditioning circuit designed in CMOS microelectronics technology, as well as an accurate on-chip positioning sensor. It will be dedicated to the study of changes in brain metabolite markers of two major diseases, Alzheimer's disease and cerebral gliomas, and to the assessment of effective therapeutic strategies.

9.1.3. Inria Project Labs

9.1.3.1. IPL Neuromarkers

Participants: Stanley Durrleman [Correspondant], Olivier Colliot [Correspondant], Fabrizio de Vico Fallani, Anne Bertrand, Stéphane Epelbaum.

Project acronym: Neuromarkers

Project title: Design of imaging biomarkers of neurodegenerative diseases for clinical trials and study of their genetic associations

Duration: 2017-2021

Coordinators: Stanley Durrleman and Olivier Colliot

Other partners: Inria GENSCALE, Inria BONSAI, Inria DYLISS, Inria XPOP, ICM, IHU/ICM iConics

Abstract: The Inria Project Lab Neuromarkers aims to develop new statistical and computational approaches to integrate multimodal imaging and omics data and to demonstrate their potential to identify early alterations and predict progression of neurodegenerative diseases. To tackle this challenge, the project brings together multidisciplinary expertise from Inria and ICM (Brain and Spine Institute) in the fields of statistical learning, brain imaging, bioinformatics, knowledge modeling, genomics and neurodegenerative diseases.

9.1.4. IHU

9.1.4.1. General program

Participants: Olivier Colliot, Stanley Durrleman, Didier Dormont, Ninon Burgos, Stéphane Epelbaum, Fabrizio de Vico Fallani.

Project acronym: IHU-A-ICM

Project title: Institute of Translational Neuroscience

Founded in 2011

General Director: Bertrand Fontaine

The IHU-A-ICM program was selected, in 2011, in a highly competitive national call for projects. A 10-year, 55M€ program, has been implemented by a recently created foundation for scientific cooperation. Based on the clinical and scientific strenghts of the ICM and the hospital Department of Nervous System Diseases, it mainly supports neuroscience research, but is also invested in improving care and teaching. ARAMIS is strongly involved in the IHU-A-ICM project, in particular in WP6 (neuroimaging and electrophysiology), WP7 (biostatistics), WP2 (Alzheimer) and WP5 (epilepsy). We have started collaborations with the new bioinformatics/biostatistics platform (IHU WP7, head: Ivan Moszer), in particular through a joint project on the integration of imaging and genomics data.

9.1.4.2. ICM-Internal Research projects

Participants: Anne Bertrand [Correspondant], Takoua Kaaouana, Benoit Delatour, Alexandra Petiet, Olivier Colliot, Arnaud Marcoux.

Project title: The Histo-MRI project: targeting MR signature of tauopathy from micro- to macroscopy

Started in 2014

Coordinator: Anne Bertrand

Identifying morphological MR signatures of brain diseases usually follows a top-down process, which starts by describing a pattern of MR signal changes in patients, hypothesizes an underlying pathological mechanism, and confirms this mechanism by correlating the observed MR signal changes with histological lesions on post-mortem examination. This top-down process, relevant for large, centimetric brain lesions, becomes inappropriate when targeting the MR signal intensity changes associated with microscopic lesions. Our project aims at developing an MR biomarker of NFT using a new bottom-up approach. We will start by identifying the MR signal changes associated with the presence of NFT at the level of the histological slice, and utilize these findings to develop a method of NFT quantification on clinical, millimetric 3D MR images. To achieve this goal, we will develop and implement a 11.7T histological coil dedicated to the scanning of histological slices, which allows both ultra-high resolution MR imaging (up to 33 microns in-plane) and perfect co-registration with histological staining, performed subsequently on the same slice. This method has the potential to provide a novel biomarker of tauopathy that could not have been identified using the usual top-down approach. It also envisions the possibility to describe and understand new MRI contrasts in other neurodegenerative diseases associated with microscopic deposition of various proteins.

9.1.4.3. ICM BBT Program - project PredictICD

Participants: Olivier Colliot [Correspondant], Jean-Christophe Corvol [Correspondant], Johann Faouzi.

Project title: Predict impulse control disorders in Parkinson's disease (PREDICT-ICD)

Started in 2018

Coordinators: Olivier Colliot and Jean-Christophe Corvol (ICM)

In Parkinson's disease (PD), the therapeutic strategy is based on the dopamine replacement therapy. Although available since the 1960s', it is only relatively recently that behavioral disorders associated with these drugs have been described. Gathered under the term of "behavioral addiction", they include impulse control disorders (ICDs), dopamine dysregulation syndrome (DDS), and punding. Interestingly, whereas addiction to L-dopa itself occurs quasi exclusively with L-dopa, ICDs appear electively under dopamine agonist (DA) therapy. The objectives of this project are: i) to elucidate the genetic basis of DA induced ICDs in PD patients from several international cohorts; ii) to develop and validate a machine learning model to predict the occurrence of ICDs from the combination of clinical and genetic data.

9.1.4.4. ICM BBT Program - project DYNAMO

Participants: Stanley Durrleman [Correspondant], Harald Hampel [Correspondant], Sabrina Fontanella, Simone Lista, Olivier Colliot, Stephanie Allassonniere, Jean-Baptiste Schiratti, Bruno Dubois, Hovagim Bakardjian, Remi Genthon, Enrica Cavedo, Katrine Rojkowa.

Project title: Dynamic models of disease progression across Alzheimer's disease stages informed by multimodal neuroimaging and biological data

Started in 2016

Coordinator: Stanley Durrleman and Harald Hampel

Other partners: Institut de la Mémoire et de la maladie d'Alzheimer

The estimation of data-driven models of disease progression for neurodegenerative diseases, including Alzheimer's disease (AD), is crucial to confirm, refine and extend the current hypothetical models. The estimation of such quantitative models from longitudinal data sets is notably difficult because of the lack of principled methodological frameworks for the analysis of spatiotemporal data.

The project builds on an innovative mathematical, statistical, and computational framework to automatically align the dynamics and the direction of individual trajectories of the evolving pathology, and then to infer a normative scenario of disease progression across different disease stages. The estimated scenario will combine spatiotemporal maps of lesion propagation, such as maps of amyloid deposition or cortical atrophy, and global measurements such as levels of CSF biomarkers. It will be possible to estimate not only a normative scenario but also the inter-individual variability in the values, dynamics and direction of both topographical and pathophysiological biomarkers changes during the course of the disease.

The application of this technology to publicly available and in-house longitudinal data sets of individuals from the asymptomatic at risk to the prodromal and dementia stages will yield new insights into the pathophysiology of AD from the preclinical to the AD dementia stages. This quantitative data-driven approach will be exploited to assess and refine the current qualitative hypothetical models of AD progression. Notably, it will complement these models with typical pathways of lesion propagation in the brain during disease progression. It will also highlight the effect of the known risk factors of AD such as apolipoprotein E genotype on the disease progression profile.

The project will open up the concrete possibility to derive a computer-aided diagnosis, staging, and prognosis tool for a better recruitment of patients in clinical studies and to assist clinicians in the diagnosis and the monitoring of both disease progression and treatment efficacy.

9.1.4.5. ICM BBT Program - project SEMAPHORE

Participants: Stanley Durrleman [Correspondant], Stéphane Lehéricy [Correspondant], Jean-Christophe Corvol, Marie Vidailhet, Raphael Couronné, Safia Said.

Project title: Personalized progression model of Parkinson's disease

Started in 2018

Coordinator: Stanley Durrleman and Stéphane Lehéricy

Other partners: Neurology and Neuro-radiology departments, Pitié-Salpêtrière Hospital, AP-HP

The aim of this project is to build a personalizable model of Parkinson's disease (PD) progression integrating the complex dynamical interplay between phenotypic, imaging, genetic and metabolic alterations. We will identify and validate markers for monitoring of progression of brain damage in early and prodromal PD and identify conversion markers in subjects at risk of PD (idiopathic rapid eye movement sleep behavior disorders iRBD, PD- related mutation carriers). We will describe the appearance, characterize clinical phenotypes of PD, and identify modifier genes of disease phenotype. To this aim, we will relie on a novel statistical learning method using Bayesian non-linear mixed-effects model allowing to combine and realign short term sequence data to estimate a long-term scenario of disease progression. This method is able to estimate individual stages of disease progression and to analyze automatically non-linear spatiotemporal patterns of data change. It estimates both a group-average scenario of PD progression as well as the inter-individual variability of this model in terms of age at onset, pace of disease progression and variability in the spatiotemporal trajectory of data changes. We will analyse the effect of genetic variants in the modulation of these non-linear progression patterns, and assess the statistical power of the individual parameters encoding for these patterns. The method will be applied to two sets of longitudinal data from the local prospective NUCLEIPARK (60 PD patients, 20 patients with iRBD, 60 controls) and ICEBERG studies (200 early idiopathic PD, 50 iRBD, 30 GBA and LRRK2 PD-related mutation carriers, 50 controls). Examinations included clinical, biological, and neurophysiological data, and multimodal 3T MRI, DATScan, and skin and salivary gland biopsies. The models of PD progression for each category of subjects will be released to the community, as well as the software for reproducibility purposes.

9.1.4.6. ICM BBT Program - project ATTACK

Participants: Fabrizio de Vico Fallani [Correspondant], Charlotte Rosso [Correspondant], Marie-Constance Corsi, Laurent Hugueville.

Project title: ATTACK Brain Network Models Of Motor Recovery After Stroke

Started in 2018

Coordinator: Fabrizio De Vico Fallani, Charlotte Rosso

Other partners: Neurology and Stroke departments, Pitié-Salpêtrière Hospital, AP-HP

Like in other connected systems, studying the structure of the interactions between different brain regions has profound implications in the comprehension of emergent complex phenomena as, for example, the capability of the human brain to functionally reorganize after cerebrovascular "attacks" or stroke. This dynamic skill, which is known in neuroscience as neural plasticity, is not only interesting from a network science perspective, but it also plays a crucial role in determining the motor/cognitive recovery of patients who survive a stroke. As a critical innovation, this project proposes to develop a systematic and rigorous approach based on neuroimaging techniques, signal processing, and network science for the modeling and analysis of temporally dynamic neural processes that characterize motor recovery after stroke. To achieve these goals, this project is organized around the following objectives: i) acquiring a comprehensive longitudinal dataset of brain and behavioral/clinical data after stroke, ii) developing new analytic tools to characterize and generate temporally dynamic brain networks, iii) building network-based models of motor recovery after stroke, accounting for individual patients. These objectives involve an intensive gathering of heterogeneous mass data, their processing, the subsequent outcome interpretation and statistical simulation, as well as the development of longitudinal models and network-based diagnostics of the patient's motor recovery progress. Results will be first characterized from pure network-theoretic and neuroscience perspectives, so as to highlight fundamental research challenges, and then validated to clarify the importance and the applicability to the clinical scenario. Our results will unveil multiscale properties of dynamic brain networks and identify predictive neuromarkers for motor recovery after stroke. This project has a two-fold impact on the society. On the one hand, it will provide new methods and robust tools to properly characterize and model temporally dynamic networks in neuroscience. On the other hand, it will provide longitudinal models of motor recovery in stroke patients that can potentially unveil the neural substrate that underpins rehabilitation, improve prognosis, and eventually lower cost of hospitalization time. From a broader perspective this interdisciplinary project proposes a transformative approach to analyze large-scale neural systems.

9.1.5. 3IA Institutes - PRAIRIE

Participants: Olivier Colliot, Stanley Durrleman, Ninon Burgos.

Project acronym: PRAIRIE

Project title: Paris Artificial Intelligence Research Institute

Founded in 2019

Director: Isabelle Ryl

Website: https://prairie-institute.fr/

PRAIRIE is one of the four selected French Institutes of AI. It was selected within a call for creation of interdisciplinary AI research institutes (or “3IAs” for “Instituts Interdisciplinaires d’Intelligence Artificielle”), as part of the national French initiative on Artificial Intelligence (AI). PRAIRIE aspires to become within five years a world leader in AI research and higher education, with an undeniable impact on economy and technology at the French, European and global levels. ARAMIS team members N. Burgos, O. Colliot and S. Durrleman hold a chair at PRAIRIE.

9.1.6. National Networks

- GdR Statistics and Medicine - http://gdr-stat-sante.math.cnrs.fr/spip/
- GdR (MaDICS) Masses de Données, Informations et Connaissances en Sciences Big Data - Data ScienceStatistics and Medicine - http://www.madics.fr/reseaux/
- F. De Vico Fallani participated to the GdR (HANDICAP) in the framework of the future strategy of Inria
- F. De Vico Fallani was founding member of the CORTICO national network for brain-computer interfaces

9.1.7. Other National Programs

9.1.7.1. Fondation Vaincre Alzheimer

Participants: Olivier Colliot, Vincent Henry, Martin Hoffman-Apitius.

Project title: Integrative multiscale knowledge model of Alzheimer’s disease pathophysiology

2019-2020

Amount: 100K€

Coordinator: Olivier Colliot

Other partners: Frauhofer SCAI (Germany)

Abstract: Alzheimer’s disease (AD) pathophysiology is still imperfectly understood. In particular, we currently lack an integrative view of the disease to interconnect knowledge about the molecular, cellular, clinical and systems levels that remain scattered. Computational knowledge models have the potential to provide such an integrative view. The aim of this project is to provide a multiscale knowledge model of AD pathophysiology by aggregating existing heterogeneous resources (disease maps, ontologies, databases) using Semantic Web standards. The resulting model and associated software tools will be made publicly available to the scientific community.

9.1.7.2. France Parkinson

Participants: Jean-Christophe Corvol, Olivier Colliot, Stanley Durrleman.

Project title: PRECISE-PD - From pathophysiology to precision medicine for Parkinson's disease

2019-2024

Amount: 3M€

Coordinator: Jean-Christophe Corvol

Other partners: Inserm CIC-1436, Inserm CIC-P1421, Inserm U1171, Université de Bordeaux (IMN), University of Glasgow, University of Calgary,

Abstract: Parkinson's disease (PD) is a complex neurodegenerative disease characterized by the progression of motor and non-motor symptoms resulting from the spreading of the disease into dopaminergic and non-dopaminergic areas. Clinical trials have failed to demonstrate efficacy to slow PD progression because the relationships between progression profiles and their underlying molecular mechanisms remain to be identified. The objective of PRECISE-PD is to propose a mechanismsbased progression model of PD by combining genetic and longitudinal clinical data from a large cohort of patients. We will implement a biobank to the NS-PARK/FCRIN cohort collecting motor and non-motor symptoms from >22,000 PD patients followed in the 24 expert centers in France. Genomic data will be generated by using a microarray platform developed for neurodegenerative diseases studies, and brain imaging will be obtained from a subgroup of patients. Computational and machine learning approaches will be developed to address the challenges of analyzing the high dimensionality and the mixture of data necessary to move beyond empirical stratification of patients. Replication will be performed in independent cohorts, and biological validation will combine biomarkers and preclinical research. PRECISE-PD is an unpreceded opportunity to open the path to the new era of precision and personalized medicine for PD.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. H2020 - Project EuroPOND

Participants: Olivier Colliot, Stanley Durrleman, Manon Ansart, Igor Koval, Alexandre Bône.

Project acronym: EuroPOND

Project title: Data-driven models for Progression Of Neurological Disease

Duration: Jan 2016 - Dec 2019

Amount: 6M€

Coordinator: Daniel Alexander

Other partners: University College London (UK), EMC Rotterdam (The Netherlands), VUMC (The Netherlands), Fate Bene Fratelli (Italy), Carol Besta Institute (Italy), Université de Genève (Switzerland), Icometrix (Belgium)

Abstract: EuroPOND will develop a data-driven statistical and computational modeling framework for neurological disease progression. This will enable major advances in differential and personalized diagnosis, prognosis, monitoring, and treatment and care decisions, positioning Europe as world leaders in one of the biggest societal challenges of 21st century healthcare. The inherent complexity of neurological disease, the overlap of symptoms and pathologies, and the high comorbidity rate suggests a systems medicine approach, which matches the specific challenge of this call. We take a uniquely holistic approach that, in the spirit of systems medicine, integrates a variety of clinical and biomedical research data including risk factors, biomarkers, and interactions. Our consortium has a multidisciplinary balance of essential expertise in mathematical/statistical/computational modelling;

clinical, biomedical and epidemiological expertise; and access to a diverse range of datasets for sporadic and well-phenotyped disease types.The project will devise and implement, as open-source software tools, advanced statistical and computational techniques for reconstructing long-term temporal evolution of disease markers from cross-sectional or short-term longitudinal data. We will apply the techniques to generate new and uniquely detailed pictures of a range of important diseases. This will support the development of new evidence-based treatments in Europe through deeper disease understanding, better patient stratification for clinical trials, and improved accuracy of diagnosis and prognosis. For example, Alzheimer's disease alone costs European citizens around €200B every year in care and loss of productivity. No disease modifying treatments are yet available. Clinical trials repeatedly fail because disease heterogeneity prevents bulk response. Our models enable fine stratification into phenotypes enabling more focussed analysis to identify subgroups that respond to putative treatments.

9.2.1.2. H2020 - Project VirtualBrainCloud

Participant: Stanley Durrleman.

Project acronym: TVBCloud

Project title: Personalized Recommendations for Neurodegenerative Disease

Duration: Jan 2019 - Dec 2022

Amount: 15 M€

Coordinator: Petra Ritter

Other partners: Charite Berlin, Université Aix Marseille, Fraunhofer Gesellschaft, University of Oxford, Forschungzentrum Juelich, Institut du Cerveau et de la Moëlle épinière, Inria, Fundacio institut de bioenginyeria de catalunya, Helsingin yliopisto, Universita degli studi di genova, Universidad complutense de Madrid, Codebox Computer-Dienste, Codemart, Eodyne Systems, Universität Wien, TP21, Alzheimer Europe

Abstract: The annual worldwide cost of Alzheimer's dementia was 777.81 billion Euro in 2015. This number will rise to 7.41 trillion Euro in 2050. Early diagnosis would save up to $7.9 trillion in medical and care costs by 2050 in the US alone. However, the emergent pathology is highly variable across people, necehighly variable across people, necessitating individualized diagnostics and interventions. The VirtualBrainCloud addresses this by bridging the gap between computational neuroscience and subcellular systems biology, integrating both research streams into a unifying computational model that supports personalized diagnostics and treatments in NDD. TheVirtualBrainCloud not only integrates existing software tools, it also merges the efforts of two big EU initiatives, namely The Virtual Brain large scale simulation platform of the EU Flagship Human Brain Project and IMI-EPAD initiative (European prevention of Alzheimer's dementia consortium). VirtualBrainCloud will develop and validate a decision support system that provides access to high quality multi-disciplinary data for clinical practice. The result will be a cloud-based brain simulation platform to support personalized diagnostics and treatments in NDD. The EU PRACE (Partnership for Advanced Computing in Europe) initiative, will provide the required computing infrastructure. TheVirtualBrainCloud will develop robust solutions for legal and ethical matters by interacting with EU projects such as European Open Science Cloud (EOSC), 'cloud4health', Alzheimer's Europe patient organizations and ELIXIR, an organization that manages and safeguards EU research data. Our software developers have already produced highly successful brain simulation and clinical decision support tools. The resulting software will be a cloud based computational modeling system that is tailored to the individual, and bridges multiple scales to identify key mechanisms that predict NDD progression and serves as Precision Decision Support System.

9.2.1.3. FET Flagship - Human Brain Project

Participants: Olivier Colliot, Stanley Durrleman.

Project acronym: HBP

Project title: Human Brain Project

Sub-project: SP8 - Medical Informatics Platform

Duration: 2016-

Abstract: The Human Brain Project (HBP) is a European Commission Future and Emerging Technologies Flagship. The HBP aims to put in place a cutting-edge, ICT-based scientific Research Infrastructure for brain research, cognitive neuroscience and brain-inspired computing. The Project promotes collaboration across the globe, and is committed to driving forward European industry. Our team is involved in the Subproject SP8 (Medical Informatics Platform). The Medical Informatics Platform (MIP) is an innovative data management system that gives researchers the means to access and analyse large amounts of anonymized clinical neuroscience data. Within that framework, we will develop and implement a method to construct disease progression models from longitudinal biomarkers. The method will use statistical learning techniques to infer a long-term disease progression model from multiple short term data from a series of individuals. The model will account for variability in age at disease onset, pace of disease progression and trajectories of biomarkers changes across individuals in the observed population.

9.2.1.4. ERC - LEASP

Participant: Stanley Durrleman.

Project acronym: LEASP

Project title: Learning Spatiotemporal Patterns in Longitudinal Image Data Sets of the Aging Brain

Duration: 2016-2021

Abstract: Time-series of multimodal medical images offer a unique opportunity to track anatomical and functional alterations of the brain in aging individuals. A collection of such time series for several individuals forms a longitudinal data set, each data being a rich iconic-geometric representation of the brain anatomy and function. These data are already extraordinary complex and variable across individuals. Taking the temporal component into account further adds difficulty, in that each individual follows a different trajectory of changes, and at a different pace. Furthermore, a disease is here a progressive departure from an otherwise normal scenario of aging, so that one could not think of normal and pathologic brain aging as distinct categories, as in the standard case-control paradigm.

Bio-statisticians lack a suitable methodological framework to exhibit from these data the typical trajectories and dynamics of brain alterations, and the effects of a disease on these trajectories, thus limiting the investigation of essential clinical questions. To change this situation, we propose to construct virtual dynamical models of brain aging by learning typical spatiotemporal patterns of alterations propagation from longitudinal iconic-geometric data sets.

By including concepts of the Riemannian geometry into Bayesian mixed effect models, the project will introduce general principles to average complex individual trajectories of iconic-geometric changes and align the pace at which these trajectories are followed. It will estimate a set of elementary spatiotemporal patterns, which combine to yield a personal aging scenario for each individual. Disease-specific patterns will be detected with an increasing likelihood.

This new generation of statistical and computational tools will unveil clusters of patients sharing similar lesion propagation profiles, paving the way to design more specific treatments, and care patients when treatments have the highest chance of success.

9.3. International Initiatives

9.3.1. Informal International Partners

- O. Colliot has an enduring collaboration with the Center for Magnetic Resonance Research, University of Minnesota, USA (P-F Van de Moortele, T. Henry).

- S. Durrleman and O. Colliot have a collaboration with the Center for Medical Image Computing (CMIC) at University College London (UCL), London, UK (D. Alexander, H. Zhang).
- F. De Vico Fallani has a collaboration with Penn University, US (Prof. D. Bassett) and Queen Mary University London, UK (Prof. Vito Latora).

9.4. International Research Visitors

9.4.1. Visits of International Scientists

We hosted Prof Bruno Jedynak from Portland State University (USA) in June and July 2019.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- S. Durrleman was general chair of the international Neuro Open Science Workhshop (ICM, Paris)

10.1.1.2. Member of the Organizing Committees

- N. Burgos organized the international workshop on Simulation and Synthesis in Medical Imaging (SASHIMI) 2019, a satellite workshop of MICCAI 2019.
- N. Burgos organized a PRACE training event (Introduction to machine learning in Python with Scikit-learn)
- F. De Vico Fallani co-organized the Network Neuroscience satellite within the NetSCI conference.
- S. Durrleman co-organized the international workshop on the Mathematical Foundations of Computational Anatomy (MFCA), a satellite workshop of MICCAI 2019.
- S. Durrleman co-organized the workshop GEANT (GEstion et pArtage de données en NeuroinformaTique), Marseille Medical School.

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

- O. Colliot served as Program Committee member for the international conference SPIE Medical Imaging (San Diego, USA, 2019) and for the international workshop PatchMI (Shenzen, China, 2019).
- O. Colliot served as Award Committee member for the international conference of the Organization for Human Brain Mapping (OHBM, meeting held in Rome, Italy, 2019).
- F. De Vico Fallani served as Program Committee member for the international conference Netsci, Complenet, Complex networks.
- S. Durrleman served as Program Committee member for the international workshop Medical Imaging meets NeurIPS.

10.1.2.2. Reviewer

- O. Colliot acted as a reviewer for the international conferences SPIE Medical Imaging, Annual meeting of the Organization for Human Brain Mapping (OHBM) and the international workshop PatchMI.
- N. Burgos acted as a reviewer for the international conference Annual meeting of the Organization for Human Brain Mapping (OHBM), for the international workshop on Simulation and Synthesis in Medical Imaging (SASHIMI), and for the GRETSI symposium of signal and image processing.

- F. De Vico Fallani acted as a reviewer for the international conference on Brain-computer interfaces (Graz BCI) and the international conferences Complenet, Netsci, Complex Networks.
- S. Durrleman acted as a reviewer for Information Processing in Medical Imaging (IPMI) and MIC meets NeurIPS.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- O. Colliot is a member of the Editorial Board of the journal Medical Image Analysis (Elsevier).
- F. De Vico Fallani is a member of the Editorial Board of PLoS One, Brain Topography, IEEE TNSRE.
- S. Epelbaum is member of the editorial board for the Medecine Cognition et Vieillissement Scientific Journal.
- S. Durrleman is member of the editorial board of IEEE Transactions on Medical Imaging (TMI), and Neuron, Behavior and Data Analysis (Scholastica)

10.1.3.2. Reviewer - Reviewing Activities

- O. Colliot acted as a reviewer for Medical Image Analysis, Neuroradiology and Clinical Epidemiology.
- N. Burgos acted as a reviewer for IEEE Transactions on Medical Imaging; Medical Image Analysis; Medical Physics; IEEE Transactions on Computational Imaging; Scientific Reports; Artificial Intelligence in Medicine; Physics in Medicine and Biology; NeuroImage: Clinical; EJNMMI Research; International Journal of Radiation Oncology, Biology, Physics.
- S. Epelbaum acted as a reviewer for Alzheimer's & Dementia, the Journal of Alzheimer's disease, Annals of Neurology and Annals of Clinical and Translational Neurology.
- S. Durrleman acted as a reviewer for Journal of Mathematical Imaging and Vision
- F. De Vico Fallani acted as a reviewer for Brain, PloS Biology, Network Neuroscience, J Neural Engineering, Neuroimage.

10.1.4. Invited Talks

- O. Colliot gave an invited seminar at the University of Southern California (Los Angeles, USA).
- O. Colliot gave an invited seminar at the meeting of the Scientific Council of the Institute of Biology of CNRS.
- O. Colliot gave an invited seminar at CentraleSupelec (Gif-sur-Yvette, France)
- N. Burgos gave an invited seminar at the Université de Tours (Tours, France)
- N. Burgos gave an invited presentation at the MaDICS Symposium (Rennes, France)
- F. De Vico Fallani was invited for a plenary talk at the Institut de neurosciences des systemes INS (Marseille, France)
- F. De Vico Fallani was invited for a plenary talk at the Stem-cell and brain research institute SBRI (Lyon, France)
- F. De Vico Fallani gave an invited seminar at Laboratoire Bordelais de la Recherche en informatique (LABRI), Bordeaux, France
- F. De Vico Fallani gave an invited seminar at the Ecole Polytechnique, Paris, France
- F. De Vico Fallani gave an invited seminar at the Ecole Normale Superieure, Paris, France
- S. Durrleman gave an invited lecture at the Institute for Pure and Applied Mathematics (IPAM), Los Angeles, USA
- S. Durrleman gave an invited lecture at the Pasadena Workshop, Sorbonne Université, Paris

- S. Durrleman gave an invited lecture at the Disease Progression Modeling workshop in Bonn, Germany
- S. Durrleman gave an invited seminar at CMIC, University College Longon, UK
- S. Durrleman gave an invited seminar at Bordeaux University, France
- S. Durrleman gave an invited seminar at Ecole Polytechnique, Paris, France

10.1.5. Scientific Expertise

- O. Colliot is a member of the "Commission des emplois scientifiques" of the Inria Paris Center, in charge of evaluating applications for PhD fellowships, postdoc fellowships and secondments.
- O. Colliot acts as an expert for GENCI (the national facility for high-performance computing).
- O. Colliot was a member of the recruitment committee ("jury d'admission") for the national competition for recruitment of permanent researchers at CNRS.
- S. Durrleman is a member of the Commission de Développement Technologique (CDT) of Inria Paris Center.
- S. Durrleman is director of the ICM Center for Neuroinformatics, and scientific manager of the ICM platform for data management and analytics.
- S. Durrleman serves as secretary of the MICCAI Special Interest Group SHAPE.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Master: Olivier Colliot coordinates the course "Methods for medical imaging" of the Master 2 in Computer Science of Sorbonne University and teaches 4.5 hours (eqTD).
- Master: F. De Vico Fallani, 3 hours (eqTD), "Methods for medical imaging" of the Master 2 in Computer Science of Sorbonne University.
- Master: Olivier Colliot coordinates the course "Artificial Intelligence" of the Master 2 Bioentrepreneur of Paris-Descartes University and teaches 30 hours (eqTD).
- Master: S. Epelbaum, Master in Neuroscience, 4 hours (eqTD), Sorbonne University
- Master: S. Durrleman, Master Mathématiques, Vision, Apprentissage (MVA), ENS Paris-Saclay (21 hrs cours magistral)
- Engineering school: Olivier Colliot, 3 hours (eqTD), Mines ParisTech
- Medical school: Didier Dormont is the Director of the University Diploma (DIU) “Diagnostic and Therapeutic Neuroradiology", Sorbonne University
- Medical school: Didier Dormont, Courses for Medical Students, Sorbonne University
- Medical school: Didier Dormont organizes and participates in the practical teaching of Neuroradiology for Medical Students in the Department of Diagnostic Neuroradiology of Pitié Salpêtrière University Hospital
- Medical school: Didier Dormont organizes and participates in the practical teaching of Neuroradiology for Radiology Specializing Residents in the Department of Diagnostic Neuroradiology of Pitié Salpêtrière University Hospital
- Medical school: S. Epelbaum gives lectures in Neurology on the topic of degenerative diseases for medical students of Sorbonne University (10 hours/year) and is regional supervisor of the national Inter University Diploma on Alzheimer's disease and Related disorders for Paris since 2015.

10.2.2. Supervision

PhD in progress : Juliana Gonzalez-Astudillo, “Network features for brain-computer interfaces”, UPMC, started in 2019, advisor: Fabrizio De Vico Fallani

PhD in progress : Giulia Bassignana, “Identification of driver nodes in biological networks”, Inserm, started in 2017, advisors: Fabrizio De Vico Fallani, Olivier Colliot, Violetta Zujovic

PhD in progress : Tiziana Cattai, “Leveraging brain connectivity networks to detect mental states in brain-computer interfaces”, Inria, started in 2017, advisor: Fabrizio De Vico Fallani

PhD in progress : Virgilio Kmetzsch, “CMultimodal analysis of neuroimaging and transcriptomic data in genetic fronto-temporal dementia”, Sorbonne University, Started in 2019, advisors: Olivier Colliot, Emmanuelle Becker and Olivier Dameron

PhD in progress: Wen Wei, “Learning brain alterations in multiple sclerosis from multimodal neuroimaging data”, Université de Nice Sophia-Antipolis, Started in 2016, advisors: N. Ayache, O. Colliot and S. Durrleman

PhD in progress: Alexandre Bône, “Learning methods for the spatiotemporal analysis of longitudinal image data: application to the diagnosis, prognosis and monitoring of Alzheimer’s disease”, started in 2016, advisors: O. Colliot and S. Durrleman

PhD in progress: Igor Koval, “Construction of disease progression models from multimodal longitudinal data”, started in 2016, advisors: S. Allassonnière and S. Durrleman

PhD in progress: Raphael Couronné, “Spatiotemporal analysis of the progression of the Parkinson’s Disease informed by multimodal longitudinal data”, started in 2018, advisor: S. Durrleman

PhD in progress: Thomas Lartigue, “Mixture Models in Gaussian Graphical Models”, started in 2017, advisors: S. Allassonnière and S. Durrleman

PhD in progress: Vianney Debavelaere, “Analysis of distribution of spatiotemporal trajectories in heterogeneous populations”, started in 2018, advisors: S. Allassonnière and S. Durrleman

PhD in progress: Johann Faouzi, “Machine learning approaches to predict impulse control disorders in Parkinson’s disease”, started in 2018, advisors: O. Colliot and J.-C. Corvol

PhD in progress: Simona Bottani, “Machine learning for differential diagnosis of neurodegenerative diseases from multimodal data”, started in 2018, advisors: O. Colliot and N. Burgos

PhD in progress: Elina Thibeau–Sutre, “Unsupervised learning from neuroimaging data to identify disease subtypes in Alzheimer’s disease and related disorders”, started in 2018, advisors: D. Dormont and N. Burgos

PhD in progress: Federica Cacciamani, “Awareness for cognitive decline in the earliest stages of Alzheimer’s disease”, started in 2018, advisor: S. Epelbaum

PhD in progress: Paul Vernhet, “Learning dynamical systems for disease progression modeling”, started in 2019, advisor: S. Durrleman

PhD in progress: Clément Mantoux, “Statistical analysis of graphs”, started in 2019, advisors: S. Durrleman and S. Allassonnière

10.2.3. Juries

- Olivier Colliot participated, as referee, to the PhD committee of Killian Hett (University of Bordeaux).
- Olivier Colliot participated, as referee, to the PhD committee of Florian Tilquin (University of Strasbourg).
- Olivier Colliot participated, as thesis director, to the PhD committee of Jorge Samper-Gonzalez (Sorbonne University).
- Olivier Colliot participated, as thesis director, to the PhD committee of Junhao Wen (Sorbonne University).
- Olivier Colliot participated, as thesis director, to the PhD committee of Pascal Lu (Sorbonne University).

- F. De Vico Fallani participated, as referee, to the PhD committee of Melodie Foullien (University of Lyon).
- F. De Vico Fallani participated, as examiner, to the PhD committee of Sebastien Campeon (Sorbonne University).
- S. Epelbaum participated, as thesis director, to the MD committee of Sinead Gaubert (AP-HP).
- S. Durrleman participated, as referee, to the PhD committee of Razwan Marinescu (University College London)
- S. Durrleman participated, as referee, to the PhD committee of Raphaël Sivera (Université Côte d'Azur)
- S. Durrleman participated, as thesis advisor, to the PhD committee of Maxime Louis (Sorbonne Université)
- S. Durrleman participated, as thesis advisor, to the PhD committee of Manon Ansart (Sorbonne Université)

10.3. Popularization

10.3.1. Articles and contents

- Olivier Colliot gave an interview for the popular science magazine "La Recherche".
- Stéphane Epelbaum participated to multiple events dedicated to general audience outreach including: articles in journals (L'Express, AFP, Pourquoi docteur) Radio shows and podcasts (France inter) and TV shows (BFMTV).
- S. Durrleman participated to several events dedicated to general audience outreach including 2 articles in Le Figaro Santé, 2 keynote lectures at AI for Health conference, and Sommet des start-up Challenges, and an article in Médecine, Cerveau et Cognition.

10.3.2. Interventions

- N. Burgos gave a presentation for the Rendez-vous des Jeunes Mathématiciennes (Inria Paris).

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] M. ANSART.*Design of data driven decision support systems for the early detection of subjects at risk to develop Alzheimer's disease*, Sorbonne Université, November 2019, https://tel.archives-ouvertes.fr/tel-02425735

[2] M. LOUIS.*Computational and statistical methods for trajectory analysis in a Riemannian geometry setting*, Sorbonnes universités, October 2019, https://hal.archives-ouvertes.fr/tel-02430450

[3] P. LU.*Statistical Learning from Multimodal Genetic and Neuroimaging data for prediction of Alzheimer's Disease*, Sorbonne Université, November 2019, https://hal.inria.fr/tel-02433613

[4] J. SAMPER-GONZÁLEZ.*Learning from multimodal data for classification and prediction of Alzheimer's disease*, Sorbonne Université, April 2019, https://tel.archives-ouvertes.fr/tel-02425827

[5] J. WEN.*Structural and microstructural neuroimaging for diagnosis and tracking of neurodegenerative diseases*, Sorbonne Universites, UPMC University of Paris 6, July 2019, https://tel.archives-ouvertes.fr/tel-02425625

Articles in International Peer-Reviewed Journal

[6] A. ANDROUIN, Y.-S. K. ABADA, M. LY, M. SANTIN, A. PETIET, S. EPELBAUM, A. BERTRAND, B. DELATOUR.*Activity-induced MEMRI cannot detect functional brain anomalies in the APPxPS1-Ki mouse model of Alzheimer's disease*, in "Scientific Reports", 2019, vol. 9, 1140 [*DOI :* 10.1038/S41598-018-37980-Y], https://hal.sorbonne-universite.fr/hal-02017552

[7] M. ANSART, S. EPELBAUM, G. GAGLIARDI, O. COLLIOT, D. DORMONT, B. DUBOIS, H. HAMPEL, S. DURRLEMAN.*Reduction of recruitment costs in preclinical AD trials. Validation of automatic pre-screening algorithm for brain amyloidosis*, in "Statistical Methods in Medical Research", January 2019, 096228021882303 [*DOI :* 10.1177/0962280218823036], https://hal.archives-ouvertes.fr/hal-01964942

[8] S. BENGHANEM, C. ROSSO, C. ARBIZU, E. MOULTON, D. DORMONT, A. LEGER, C. PIRES, Y. SAMSON.*Aphasia outcome: the interactions between initial severity, lesion size and location*, in "Journal of Neurology", June 2019, vol. 266, n^o 6, p. 1303-1309 [*DOI :* 10.1007/S00415-019-09259-3], https://hal.inria.fr/hal-02418331

[9] D. BERNARDESCHI, B. LAW-YE, F. BIELLE, B. HOCHET, O. STERKERS, D. DORMONT, N. PYATIGORSKAYA.*Bioactive glass granules for mastoid and epitympanic surgical obliteration: CT and MRI appearance*, in "European Radiology", October 2019, vol. 29, n^o 10, p. 5617-5626 [*DOI :* 10.1007/S00330-019-06120-Z], https://hal.inria.fr/hal-02418328

[10] M.-C. CORSI, M. CHAVEZ, D. SCHWARTZ, N. GEORGE, L. HUGUEVILLE, A. E. KAHN, S. DUPONT, D. BASSETT, F. DE VICO FALLANI.*Functional disconnection of associative cortical areas predicts performance during BCI training*, in "NeuroImage", January 2020, 116500, forthcoming [*DOI :* 10.1016/J.NEUROIMAGE.2019.116500], https://hal.inria.fr/hal-02438794

[11] C. CURY, S. DURRLEMAN, D. CASH, M. LORENZI, J. M. NICHOLAS, M. BOCCHETTA, J. C. VAN SWIETEN, B. BORRONI, D. GALIMBERTI, M. MASELLIS, M. C. TARTAGLIA, J. ROWE, C. GRAFF, F. TAGLIAVINI, G. B. FRISONI, R. LAFORCE, E. FINGER, A. DE MENDONÇA, S. SORBI, S. OURSELIN, J. ROHRER, M. MODAT, C. ANDERSSON, S. ARCHETTI, A. ARIGHI, L. BENUSSI, S. BLACK, M. COSSEDDU, M. FALLSTRM, C. G. FERREIRA, C. FENOGLIO, N. FOX, M. FREEDMAN, G. FUMAGALLI, S. GAZZINA, R. GHIDONI, M. GRISOLI, V. JELIC, L. JISKOOT, R. KEREN, G. LOMBARDI, C. MARUTA, L. MEETER, R. VAN MINKELEN, B. NACMIAS, L. IJERSTEDT, A. PADOVANI, J. PANMAN, M. PIEVANI, C. POLITO, E. PREMI, S. PRIONI, R. RADEMAKERS, V. REDAELLI, E. ROGAEVA, G. ROSSI, M. ROSSOR, E. SCARPINI, D. TANG-WAI, H. THONBERG, P. TIRABOSCHI, A. VERDELHO, J. WARREN.*Spatiotemporal analysis for detection of pre-symptomatic shape changes in neurodegenerative diseases: Initial application to the GENFI cohort*, in "NeuroImage", March 2019, vol. 188, p. 282-290 [*DOI :* 10.1016/J.NEUROIMAGE.2018.11.063], https://www.hal.inserm.fr/inserm-01958916

[12] C. CURY, M. A. SCELZI, R. TORO, V. FROUIN, E. ARTIGES, A. HEINZ, H. LEMAÎTRE, J.-L. MARTINOT, J.-B. POLINE, M. N. SMOLKA, H. WALTER, G. SCHUMANN, A. ALTMANN, O. COLLIOT.*Genome wide association study of incomplete hippocampal inversion in adolescents*, in "PLoS ONE", 2020, forthcoming, https://hal.inria.fr/hal-02416654

[13] F. DE VICO FALLANI, D. BASSETT.*Network neuroscience for optimizing brain–computer interfaces*, in "Physics of Life Reviews", December 2019, vol. 31, p. 304-309 [*DOI :* 10.1016/J.PLREV.2018.10.001], https://hal.archives-ouvertes.fr/hal-02428684

[14] D. DORMONT, N. BEN-ELIEZER, N. BURGOS, O. COLLIOT, F. DE VICO FALLANI, M. DHENAIN, S. DURRLEMAN, S. EPELBAUM, M.-O. HABERT, I. LE BER, J.-F. MÉDER, Y. MENU, C. OPPENHEIM, Y. WADGHIRI.*Tribute to Anne Bertrand (1978–2018): Neuroradiologist, scientist, teacher and friend*, in "Journal de Neuroradiologie / Journal of Neuroradiology", March 2019, vol. 46, n^o 2, p. 155-159 [*DOI :* 10.1016/J.NEURAD.2019.01.001], https://hal.inria.fr/hal-02265305

[15] S. EPELBAUM, C. PAQUET, J. HUGON, J. DUMURGIER, D. WALLON, D. HANNEQUIN, T. JONVEAUX, A. BESOZZI, S. POUPONNEAU, C. HOMMET, F. BLANC, L. BERLY, A. JULIAN, M. PACCALIN, F. PASQUIER, J. BELLET, C. BOUTOLEAU-BRETONNIÈRE, T. CHARRIAU, O. ROUAUD, O. MADEC, A. MOUTON, R. DAVID, S. BEKADAR, R. FABRE, E. LIEGEY, W. DEBERDT, P. ROBERT, B. DUBOIS.*How many patients are eligible for disease-modifying treatment in Alzheimer's disease? A French national observational study over 5 years*, in "BMJ Open", June 2019, vol. 9, n^o 6, e029663 [*DOI :* 10.1136/BMJOPEN-2019-029663], https://hal.inria.fr/hal-02165245

[16] C. FOURNIER, M. BARBIER, A. CAMUZAT, V. ANQUETIL, S. LATTANTE, F. CLOT, C. CAZENEUVE, M. SABATELLI, S. FORLANI, L. JORNEA, E. E. LEGUERN, A. BRICE, S. AURIACOMBE, S. BELLIARD, F. BLANC, C. BOUTOLEAU-BRETONNIÈRE, M. CECCALDI, M. DIDIC, C. DUYCKAERTS, F. ETCHARRY-BOUYX, V. GOLFIER, L. LACOMBLEZ, B.-F. MICHEL, C. THOMAS-ANTÉRION, J. PARIENTE, F. SELLAL, M. VERCELLETTO, E. BENCHETRIT, H. BERTIN, A. BERTRAND, A. BISSERY, S. BOMBOIS, M.-P. BONCOEUR, P. CASSAGNAUD, M. CHASTAN, Y. CHEN, M. CHUPIN, O. COLLIOT, P. COURATIER, X. DELBEUCQ, V. DERAMECOURT, C. DELMAIRE, E. GÉRARDIN, C. HOSSEIN-FOUCHER, B. DUBOIS, M. HABERT, D. HANNEQUIN, G. LAUTRETTE, T. LEBOUVIER, I. LE BER, S. LEHÉRICY, B. LE TOULLEC, R. LÉVY, M. MARTINEAU, A. MACKOWIAK, J. MONTEIL, F. PASQUIER, G. PETYT, P.-F. PRADAT, A.-H. OYA, D. RINALDI, A. ROLLIN-SILLAIRE, F. SALACHAS, S. SAYAH, D. WALLON.*Relations between C9orf72 expansion size in blood, age at onset, age at collection and transmission across generations in patients and presymptomatic carriers*, in "Neurobiology of Aging", February 2019, vol. 74, p. 234.e1-234.e8 [*DOI :* 10.1016/J.NEUROBIOLAGING.2018.09.010], https://hal-unilim.archives-ouvertes.fr/hal-02025976

[17] S. GAUBERT, F. RAIMONDO, M. HOUOT, M.-C. CORSI, J. D. SITT, B. HERMANN, D. OUDIETTE, G. GAGLIARDI, M.-O. HABERT, B. DUBOIS, F. DE VICO FALLANI, H. BAKARDJIAN, S. EPELBAUM.*EEG evidence of compensatory mechanisms in preclinical Alzheimer's disease*, in "Brain - A Journal of Neurology ", June 2019 [*DOI :* 10.1093/BRAIN/AWZ150], https://hal.inria.fr/hal-02164586

[18] B. GEERTS, B. L. YE, D. GALANAUD, D. DORMONT, N. PYATIGORSKAYA.*Potential Effect of Fetal Origin of Posterior Cerebral Artery on the Arterial Spin Labeling Sequence*, in "American Journal of Neuroradiology", November 2019 [*DOI :* 10.1016/J.NEURAD.2019.08.006], https://hal.inria.fr/hal-02418326

[19] F. GROSSELIN, X. NAVARRO-SUNE, A. VOZZI, K. PANDREMMENOU, F. DE VICO FALLANI, Y. ATTAL, M. CHAVEZ.*Quality Assessment of Single-Channel EEG for Wearable Devices*, in "Sensors", 2019, vol. 19, n^o 3, 601 [*DOI :* 10.3390/S19030601], https://hal.sorbonne-universite.fr/hal-02017731

[20] J. GUILLON, M. CHAVEZ, F. BATTISTON, Y. ATTAL, V. L. CORTE, M. THIEBAUD DE SCHOTTEN, B. DUBOIS, D. SCHWARTZ, O. COLLIOT, F. DE VICO FALLANI.*Disrupted core-periphery structure of multimodal brain networks in Alzheimer's disease*, in "Network Neuroscience", January 2019, vol. 3, n^o 2, p. 635-652 [*DOI :* 10.1162/NETN_A_00087], https://hal.inria.fr/hal-02265308

[21] M. LACOUR, O. QUENEZ, A. ROVELET-LECRUX, B. SALOMON, S. ROUSSEAU, A.-C. RICHARD, M. QUILLARD-MURAINE, F. PASQUIER, A. ROLLIN-SILLAIRE, O. MARTINAUD, A. ZAREA, V. DE LA SAYETTE, C. BOUTOLEAU-BRETONNIÈRE, F. ETCHARRY-BOUYX, V. CHAUVIRÉ, M. SARAZIN, I. LE BER, S. EPELBAUM, T. JONVEAUX, O. ROUAUD, M. CECCALDI, O. GODEFROY, M. FORMAGLIO, B.

CROISILE, S. AURIACOMBE, E. MAGNIN, M. SAUVÉE, C. MARELLI, A. GABELLE, J. PARIENTE, C. PAQUET, A. BOLAND, J.-F. DELEUZE, D. CAMPION, D. HANNEQUIN, G. NICOLAS, D. WALLON.*Causative Mutations and Genetic Risk Factors in Sporadic Early Onset Alzheimer's Disease Before 51 Years*, in "Journal of Alzheimer's Disease", September 2019, vol. 71, n^{o} 1, p. 227-243 [*DOI :* 10.3233/JAD-190193], https://hal-normandie-univ.archives-ouvertes.fr/hal-02332506

[22] G. QUERIN, P. BEDE, M. M. EL MENDILI, M. LI, M. PÉLÉGRINI-ISSAC, D. RINALDI, M. CATALA, D. SARACINO, F. SALACHAS, A. CAMUZAT, V. MARCHAND-PAUVERT, J. COHEN-ADAD, O. COLLIOT, I. LE BER, P.-F. PRADAT.*Presymptomatic spinal cord pathology in c9orf72 mutation carriers: a longitudinal neuroimaging study*, in "Annals of Neurology", June 2019, vol. 86, n^{o} 2, p. 158-167 [*DOI :* 10.1002/ANA.25520], https://hal.inria.fr/hal-02265301

[23] M. THIEBAUD DE SCHOTTEN, H. BAKARDJIAN, H. BENALI, H. BERTIN, J. BONHEUR, L. BOUKADIDA, N. BOUKERROU, O. COLLIOT, M. DUBOIS, S. EPELBAUM, G. GAGLIARDI, R. GENTHON, M. HOUOT, A. KAS, F. LAMARI, M. LEVY, C. METZINGER, F. MOCHEL, F. NYASSE, C. POISSON, M.-C. POTIER, M. RÉVILLON, A. SANTOS, K. S. ANDRADE, M. SOLE, M. SURTEE, N. YOUNSI, M. AFSHAR, L. F. AGUILAR, L. AKMAN-ANDERSON, J. ARENAS, J. AVILA, C. BABILONI, F. BALDACCI, R. BATRLA, N. BENDA, K. BLACK, A. L. BOKDE, U. BONUCCELLI, K. BROICH, F. CACCIOLA, F. CARACI, J. CASTRILLO, E. CAVEDO, R. CERAVOLO, P. A. CHIESA, J.-C. CORVOL, A. C. CUELLO, J. CUMMINGS, H. DEPYPERE, B. DUBOIS, A. DUGGENTO, E. EMANUELE, V. ESCOTT-PRICE, H. FEDEROFF, M. T. FERRETTI, M. FIANDACA, R. FRANK, F. GARACI, H. GEERTS, F. GIORGI, E. J. GOETZL, M. GRAZIANI, M. HABERKAMP, M.-O. HABERT, H. HAMPEL, K. HERHOLZ, F. HERNANDEZ, D. KAPOGIANNIS, E. KARRAN, S. J. KIDDLE, S. KIM, Y. KORONYO, M. KORONYO-HAMAOUI, T. LANGEVIN, S. LEHÉRICY, A. LUCIA, S. LISTA, J. LORENCEAU, D. MANGO, M. MAPSTONE, C. NERI, R. NISTICÒ, S. O'BRYANT, G. PALERMO, G. L. W. PERRY, C. RITCHIE, S. ROSSI, A. SAIDI, E. SANTARNECCHI, L. S. SCHNEIDER, O. SPORNS, N. TOSCHI, S. R. VERDOONER, A. VERGALLO, N. VILLAIN, L. WELIKOVITCH, J. WOODCOCK, E. YOUNESI.*Differential default mode network trajectories in asymptomatic individuals at risk for Alzheimer's disease*, in "Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring", July 2019, vol. 15, n^{o} 7, p. 940-950 [*DOI :* 10.1016/J.JALZ.2019.03.006], https://hal.archives-ouvertes.fr/hal-02368904

[24] S.-C. VILLENEUVE, M. HOUOT, F. CACCIAMANI, M. VERRIJP, B. DUBOIS, S. SIKKES, S. EPELBAUM.*Latent class analysis identifies functional decline with Amsterdam IADL in preclinical Alzheimer's disease*, in "Alzheimer's and Dementia", 2019, vol. 5, p. 553-562 [*DOI :* 10.1016/J.TRCI.2019.08.009], https://hal.archives-ouvertes.fr/hal-02333456

[25] W. WEI, E. POIRION, B. BODINI, S. DURRLEMAN, N. AYACHE, B. STANKOFF, O. COLLIOT.*Predicting PET-derived Demyelination from Multimodal MRI using Sketcher-Refiner Adversarial Training for Multiple Sclerosis*, in "Medical Image Analysis", 2019, forthcoming [*DOI :* 10.1016/J.MEDIA.2019.101546], https://hal.archives-ouvertes.fr/hal-02276634

[26] W. WEI, E. POIRION, B. BODINI, S. DURRLEMAN, O. COLLIOT, B. STANKOFF, N. AYACHE.*Fluid-attenuated inversion recovery MRI synthesis from multisequence MRI using three-dimensional fully convolutional networks for multiple sclerosis*, in "Journal of Medical Imaging", February 2019, vol. 6, n^{o} 01 [*DOI :* 10.1117/1.JMI.6.1.014005], https://hal.inria.fr/hal-02042526

[27] J. WEN, H. ZHANG, D. C. ALEXANDER, S. DURRLEMAN, A. ROUTIER, D. RINALDI, M. HOUOT, P. COURATIER, D. HANNEQUIN, F. PASQUIER, J. ZHANG, O. COLLIOT, I. LE BER, A. BERTRAND.*Neurite density is reduced in the presymptomatic phase of C9orf72 disease*, in "Journal of Neurology, Neurosurgery

and Psychiatry", April 2019, vol. 90, n^{o} 4, p. 387-94 [*DOI :* 10.1136/JNNP-2018-318994], https://hal.inria.fr/hal-01907482

[28] L. XICOTA, F. ICHOU, F.-X. LEJEUNE, B. COLSCH, A. TENENHAUS, I. LEROY, G. FONTAINE, M. LHOMME, H. BERTIN, M.-O. HABERT, S. EPELBAUM, B. DUBOIS, F. MOCHEL, M.-C. POTIER.*Multi-omics signature of brain amyloid deposition in asymptomatic individuals at-risk for Alzheimer's disease: The INSIGHT-preAD study*, in "EBioMedicine", 2019, vol. 47, p. 518-528 [*DOI :* 10.1016/J.EBIOM.2019.08.051], https://hal.sorbonne-universite.fr/hal-02305609

International Conferences with Proceedings

[29] A. BÔNE, O. COLLIOT, S. DURRLEMAN.*Auto-encoding meshes of any topology with the current-splatting and exponentiation layers*, in "Geometry Meets Deep Learning @ ICCV 2019", Séoul, South Korea, October 2019, https://hal.inria.fr/hal-02087586

[30] A. BÔNE, M. LOUIS, O. COLLIOT, S. DURRLEMAN.*Learning low-dimensional representations of shape data sets with diffeomorphic autoencoders*, in "IPMI 2019 : Information Processing in Medical Imaging", Hong-Kong, China, June 2019, Auteur collectif : Alzheimer's Disease Neuroimaging Initiative, https://hal.inria.fr/hal-01963736

[31] R. COURONNE, M. VIDAILHET, J.-C. CORVOL, S. LEHÉRICY, S. DURRLEMAN.*Learning disease progression models with longitudinal data and missing values*, in "ISBI 2019 - International Symposium on Biomedical Imaging", Venice, Italy, April 2019, https://hal.archives-ouvertes.fr/hal-02091571

[32] *Best Paper*
V. DEBAVELAERE, A. BÔNE, S. DURRLEMAN, S. ALLASSONNIÈRE.*Clustering of longitudinal shape data sets using mixture of separate or branching trajectories*, in "Medical Image Computing and Computer Assisted Intervention – MICCAI", Shenzhen, China, Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, October 2019, p. 66-74, https://hal.archives-ouvertes.fr/hal-02103355.

[33] M. LOUIS, R. COURONNE, I. KOVAL, B. CHARLIER, S. DURRLEMAN.*Riemannian Geometry Learning for Disease Progression Modelling*, in "Information Processing in Medical Imaging - IPMI 2019", Hong-Kong, China, May 2019, p. 542-553 [*DOI :* 10.1007/978-3-030-20351-1_42], https://hal.inria.fr/hal-02429839

[34] J. SAMPER-GONZÁLEZ, N. BURGOS, S. BOTTANI, M.-O. HABERT, T. EVGENIOU, S. EPELBAUM, O. COLLIOT.*Reproducible evaluation of methods for predicting progression to Alzheimer's disease from clinical and neuroimaging data*, in "SPIE Medical Imaging 2019", San Diego, United States, February 2019, https://hal.inria.fr/hal-02025880

[35] E. THIBEAU-SUTRE, O. COLLIOT, D. DORMONT, N. BURGOS.*Visualization approach to assess the robustness of neural networks for medical image classification*, in "SPIE Medical Imaging 2020", Houston, United States, February 2020, https://arxiv.org/abs/1911.08264 , https://hal.archives-ouvertes.fr/hal-02370532

Conferences without Proceedings

[36] M. ANSART, N. BURGOS, O. COLLIOT, D. DORMONT, S. DURRLEMAN.*Prediction of future cognitive scores and dementia onset in Mild Cognitive Impairment patients*, in "OHBM 2019 - Organization for Human Brain Mapping Conference", Rome, Italy, June 2019, https://hal.inria.fr/hal-02098427

[37] G. BASSIGNANA, J. FRANSSON, O. COLLIOT, V. ZUJOVIC, F. DE VICO FALLANI.*Target controllability in genetic networks of macrophage activation*, in "Netsci 2019", Burlington, United States, May 2019, https://hal.archives-ouvertes.fr/hal-02314001

[38] A. BÔNE, B. MARTIN, M. LOUIS, O. COLLIOT, S. DURRLEMAN.*Hierarchical modeling of Alzheimer's disease progression from a large longitudinal MRI data set*, in "OHBM 2019 - 25th Annual Meeting of the Organization for Human Brain Mapping", Roma, Italy, June 2019, https://hal.inria.fr/hal-02090275

[39] M.-C. CORSI, M. CHAVEZ, D. P. SCHWARTZ, N. GEORGE, L. HUGUEVILLE, A. E. KAHN, S. DUPONT, D. S. BASSETT, F. DE VICO FALLANI.*Looking for cortical patterns of successful motor imagery-based BCI learning*, in "8th Graz Brain-Computer Interface Conference 2019", Graz, Austria, September 2019, https://hal.archives-ouvertes.fr/hal-02123582

[40] M.-C. CORSI, M. CHAVEZ, D. SCHWARTZ, N. GEORGE, L. HUGUEVILLE, A. E. KAHN, S. DUPONT, D. S. BASSETT, F. DE VICO FALLANI.*Looking for neurophysiological correlates of brain-computer interface learning*, in "OHBM 2019 - Annual Meeting on Organization for Human Brain Mapping", Rome, Italy, June 2019, https://hal.archives-ouvertes.fr/hal-02157179

[41] R. COURONNE, M. VIDAILHET, G. MANGONE, S. LEHÉRICY, S. DURRLEMAN.*Motor abnormalities progression in prodromal PD*, in "AD/PD 2019 - 14th International Conference on Alzheimer's and Parkinson's Diseases and related neurological disorders", Lisbon, Portugal, March 2019, https://hal.archives-ouvertes.fr/hal-02276297

[42] C. CURY, M. A. SCELZI, R. TORO, V. FROUIN, E. ARTIGES, A. HEINZ, H. W. WALTER, H. LEMAÎTRE, J.-L. MARTINOT, J.-B. POLINE, M. N. SMOLKA, G. SCHUMANN, A. ALTMANN, O. COLLIOT.*On the genetic bases of incomplete hippocampal inversion: a genome-wide association study*, in "OHBM 2019 - Annual Meeting on Organization for Human Brain Mapping", Rome, Italy, June 2019, 1, https://www.hal.inserm.fr/inserm-02074616

[43] J. FAOUZI, S. BEKADAR, O. COLLIOT, J.-C. CORVOL.*Predicting Impulse Control Disorders in Parkinson's Disease: A Challenging Task*, in "International Congress of Parkinson's Disease and Movement Disorders", Nice, France, September 2019, https://hal.inria.fr/hal-02315533

[44] V. J. HENRY, G. BASSIGNANA, V. ZUJOVIC, F. DE VICO FALLANI, O. DAMERON, I. MOSZER, O. COLLIOT.*Conciliation of process description and molecular interaction networks using logical properties of ontology*, in "JOBIM 2019 - Journées Ouvertes Biologie, Informatique et Mathématiques", Nantes, France, July 2019, https://hal.archives-ouvertes.fr/hal-02301702

[45] I. KOVAL, A. MARCOUX, N. BURGOS, S. ALLASSONNIÈRE, O. COLLIOT, S. DURRLEMAN.*Deciphering the progression of PET alterations using surface-based spatiotemporal modeling*, in "OHBM 2019 - Annual meeting of the Organization for Human Brain Mapping", Rome, Italy, June 2019, https://hal.inria.fr/hal-02134909

[46] A. ROUTIER, M.-O. HABERT, A. BERTRAND, A. KAS, P.-M. DAVID, H. BERTIN, O. GODEFROY, F. ETCHARRY-BOUYX, O. MOREAUD, F. PASQUIER, P. COURATIER, K. BENNYS, C. BOUTOLEAU-BRETONNIÈRE, O. MARTINAUD, B. LAURENT, J. PARIENTE, M. PUEL, S. BELLIARD, R. MIGLIACCIO, B. DUBOIS, O. COLLIOT, M. TEICHMANN.*New advances in the Clinica software platform for clinical neuroimaging studies*, in "Annual meeting of the Organization for Human Brain Mapping - OHBM 2019", Roma, Italy, June 2019, https://hal.inria.fr/hal-01764289

[47] A. ROUTIER, A. MARCOUX, M. DIAZ MELO, J. GUILLON, J. SAMPER-GONZÁLEZ, J. WEN, S. BOTTANI, A. GUYOT, E. THIBEAU-SUTRE, M. TEICHMANN, M.-O. HABERT, S. DURRLEMAN, N. BURGOS, O. COLLIOT.*New advances in the Clinica software platform for clinical neuroimaging studies*, in "OHBM 2019 - Annual Meeting on Organization for Human Brain Mapping", Roma, Italy, June 2019 [*DOI :* 10.1016/J.NEUROIMAGE.2011.09.015], https://hal.inria.fr/hal-02132147

[48] J. SAMPER-GONZÁLEZ, N. BURGOS, S. BOTTANI, M.-O. HABERT, T. EVGENIOU, S. EPELBAUM, O. COLLIOT.*Predicting progression to Alzheimer's disease from clinical and imaging data: a reproducible study*, in "OHBM 2019 - Organization for Human Brain Mapping Annual Meeting 2019", Rome, Italy, June 2019, https://hal.inria.fr/hal-02142315

[49] J. WEN, J. SAMPER-GONZÁLEZ, A. M. ROUTIER, S. BOTTANI, S. DURRLEMAN, N. BURGOS, O. COLLIOT.*Beware of feature selection bias! Example on Alzheimer's disease classification from diffusion MRI*, in "2019 OHBM Annual Meeting - Organization for Human Brain Mapping", Rome, Italy, June 2019, https://hal.inria.fr/hal-02105134

[50] J. WEN, E. THIBEAU-SUTRE, J. SAMPER-GONZÁLEZ, A. M. ROUTIER, S. BOTTANI, D. DORMONT, S. DURRLEMAN, O. COLLIOT, N. BURGOS.*How serious is data leakage in deep learning studies on Alzheimer's disease classification?*, in "2019 OHBM Annual meeting - Organization for Human Brain Mapping", Rome, Italy, June 2019, https://hal.inria.fr/hal-02105133

Scientific Books (or Scientific Book chapters)

[51] N. BURGOS, A. A. GOOYA, D. SVOBODA.*Simulation and Synthesis in Medical Imaging : 4th International Workshop, SASHIMI 2019, held in conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings*, Springer, 2019, vol. Lecture Notes in Computer Science, n^{o} 11827 [*DOI :* 10.1007/978-3-030-32778-1], https://hal.inria.fr/hal-02422136

Books or Proceedings Editing

[52] D. ZHU, J. YAN, S. DURRLEMAN, S. SOMMER, H. HUANG, L. SHEN, P. M. THOMPSON, C.-F. WESTIN, X. PENNEC, S. JOSHI, M. NIELSEN, T. FLETCHER (editors). *Multimodal Brain Image Analysis and Mathematical Foundations of Computational Anatomy : 4th International Workshop, MBIA 2019, and 7th International Workshop, MFCA 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings*, Lecture Notes in Computer Science (LNCS), Springer, October 2019, vol. 11846 [*DOI :* 10.1007/978-3-030-33226-6], https://hal.inria.fr/hal-02341877

Scientific Popularization

[53] O. COLLIOT.*Interpretable and reliable artificial intelligence systems for brain diseases*, in "ERCIM News", July 2019, vol. 118, https://hal.inria.fr/hal-02178901

Other Publications

[54] M. ANSART, S. EPELBAUM, G. BASSIGNANA, A. BÔNE, S. BOTTANI, T. CATTAI, R. COURONNE, J. FAOUZI, I. KOVAL, M. LOUIS, E. THIBEAU-SUTRE, J. WEN, A. WILD, N. BURGOS, D. DORMONT, O. COLLIOT, S. DURRLEMAN.*Predicting the Progression of Mild Cognitive Impairment Using Machine Learning: A Systematic and Quantitative Review*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02337815

[55] A. BÔNE, O. COLLIOT, S. DURRLEMAN.*Learning the spatiotemporal variability in longitudinal shape data sets*, October 2019, working paper or preprint, https://hal.inria.fr/hal-02091549

[56] R. COURONNE, M. LOUIS, S. DURRLEMAN.*Longitudinal autoencoder for multi-modal disease progression modelling*, April 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02090886

[57] C. CURY, J. A. GLAUNÈS, R. TORO, M. CHUPIN, G. D. SCHUMANN, V. FROUIN, J. B. POLINE, O. COLLIOT.*Statistical shape analysis of large datasets based on diffeomorphic iterative centroids*, February 2019, working paper or preprint [*DOI :* 10.1101/363861], https://hal.inria.fr/hal-01832191

[58] V. DEBAVELAERE, S. DURRLEMAN, S. ALLASSONNIÈRE.*Learning the clustering of longitudinal shape data sets into a mixture of independent or branching trajectories*, September 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02283747

[59] A. GUYOT, A. B. GRACIANO FOUQUIER, E. GERARDIN, M. CHUPIN, J. GLAUNÈS, L. MARRAKCHI-KACEM, J. GERMAIN, C. BOUTET, C. CURY, L. HERTZ-PANNIER, A. VIGNAUD, S. DURRLEMAN, T. HENRY, P.-F. VAN DE MOORTELE, A. TROUVÉ, O. COLLIOT.*A Diffeomorphic Vector Field Approach to Analyze the Thickness of the Hippocampus from 7T MRI*, November 2019, working paper or preprint, https://hal.inria.fr/hal-02359660

[60] I. KOVAL, S. ALLASSONNIÈRE, S. DURRLEMAN.*Simulation of virtual cohorts increases predictive accuracy of cognitive decline in MCI subjects*, April 2019, https://arxiv.org/abs/1904.02921 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02089881

[61] M. LOUIS, R. COURONNE, I. KOVAL, B. CHARLIER, S. DURRLEMAN.*Riemannian geometry learning for disease progression modelling*, April 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02079820

[62] P. LU, O. COLLIOT.*A log-logistic survival model from multimodal data for prediction of Alzheimer's Disease*, April 2019, SAfJR 2019 - Survival Analysis for Junior Researchers, Poster, https://hal.inria.fr/hal-02430943

[63] A. ROUTIER, N. BURGOS, J. GUILLON, J. SAMPER-GONZÁLEZ, J. WEN, S. BOTTANI, A. MARCOUX, M. BACCI, S. FONTANELLA, T. JACQUEMONT, A. WILD, P. GORI, A. GUYOT, P. LU, M. DÍAZ, E. THIBEAU-SUTRE, T. MOREAU, M. TEICHMANN, M.-O. HABERT, S. DURRLEMAN, O. COLLIOT.*Clinica: an open source software platform for reproducible clinical neuroscience studies*, October 2019, working paper or preprint, https://hal.inria.fr/hal-02308126

Project-Team CAGE

Control and Geometry

IN COLLABORATION WITH: Laboratoire Jacques-Louis Lions (LJLL)

IN PARTNERSHIP WITH:
CNRS

Sorbonne Université

RESEARCH CENTER
Paris

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team CAGE

Creation of the Team: 2017 July 01, updated into Project-Team: 2018 August 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)
A6.4. - Automatic control
A6.4.1. - Deterministic control
A6.4.3. - Observability and Controlability
A6.4.4. - Stability and Stabilization
A6.4.5. - Control of distributed parameter systems
A6.4.6. - Optimal control

Other Research Topics and Application Domains:

B1.2. - Neuroscience and cognitive science
B2.6. - Biological and medical imaging
B5.11. - Quantum systems
B7.1.3. - Air traffic

1. Team, Visitors, External Collaborators

Research Scientists

Mario Sigalotti [Team leader, Inria, Senior Researcher, HDR]
Pierre Martinon [Inria, Researcher]
Ugo Boscain [CNRS, Senior Researcher, HDR]
Barbara Gris [CNRS, Researcher, from Nov 2019]

Faculty Members

Jean Michel Coron [Sorbonne Université, Professor, HDR]
Emmanuel Trélat [Sorbonne Université, Professor, HDR]

PhD Students

Nicolas Augier [École polytechnique, PhD Student, until Sep 2019]
Amaury Hayat [École Nationale des Ponts et Chaussées, PhD Student, until Aug 2019]
Mathieu Kohli [École polytechnique, PhD Student, until Sep 2019]
Gontran Lance [École Nationale Supérieure de Techniques Avancées, PhD Student]
Cyril Letrouit [École Normale Supérieure de Paris, PhD Student]
Emilio Molina [Universidad de Chile, PhD Student, from Nov 2019]
Jakub Orłowski [Centrale-Supélec, PhD Student]
Eugenio Pozzoli [Inria, PhD Student]
Remi Robin [École polytechnique, PhD Student, from Oct 2019]
Shengquan Xiang [Univ Pierre et Marie Curie, PhD Student, until Oct 2019]
Christophe Zhang [Univ Pierre et Marie Curie, PhD Student, until Oct 2019]

Post-Doctoral Fellows

Ivan Beschastnyi [Inria, Post-Doctoral Fellow]
Valentina Franceschi [Sorbonne Université, Post-Doctoral Fellow, from Sep 2019]

Karen Habermann [Sorbonne Université, Post-Doctoral Fellow, from Nov 2019]

Administrative Assistant

Mathieu Mourey [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

CAGE's activities take place in the field of mathematical control theory, with applications in three main directions: geometric models for vision, control of quantum mechanical systems, and control of systems with uncertain dynamics.

The relations between control theory and geometry of vision rely on the notion of sub-Riemannian structure, a geometric framework which is used to measure distances in nonholonomic contexts and which has a natural and powerful control theoretical interpretation. We recall that nonholonomicity refers to the property of a velocity constraint that cannot be recast as a state constraint. In the language of differential geometry, a sub-Riemannian structure is a (possibly rank-varying) Lie bracket generating distribution endowed with a smoothly varying norm.

Sub-Riemannian geometry, and in particular the theory of associated (hypoelliptic) diffusive processes, plays a crucial role in the neurogeometrical model of the primary visual cortex due to Petitot, Citti and Sarti, based on the functional architecture first described by Hubel and Wiesel. Such a model can be used as a powerful paradigm for bio-inspired image processing, as already illustrated in the recent literature (including by members of our team). Our contributions to this field are based not only on this approach, but also on another geometric and sub-Riemannian framework for vision, based on pattern matching in the group of diffeomorphisms. In this case admissible diffeomorphisms correspond to deformations which are generated by vector fields satisfying a set of nonholonomic constraints. A sub-Riemannian metric on the infinite-dimensional group of diffeomorphisms is induced by a length on the tangent distribution of admissible velocities. Nonholonomic constraints can be especially useful to describe distortions of sets of interconnected objects (e.g., motions of organs in medical imaging).

Control theory is one of the components of the forthcoming quantum revolution [0], since manipulation of quantum mechanical systems is ubiquitous in applications such as quantum computation, quantum cryptography, and quantum sensing (in particular, imaging by nuclear magnetic resonance). The efficiency of the control action has a dramatic impact on the quality of the coherence and the robustness of the required manipulation. Minimal time constraints and interaction of time scales are important factors for characterizing the efficiency of a quantum control strategy. Time scales analysis is important for evaluation approaches based on adiabatic approximation theory, which is well-known to improve the robustness of the control strategy. CAGE works for the improvement of evaluation and design tools for efficient quantum control paradigms, especially for what concerns quantum systems evolving in infinite-dimensional Hilbert spaces.

Simultaneous control of a continuum of systems with slightly different dynamics is a typical problem in quantum mechanics and also a special case of the third applicative axis to which CAGE is contributing: control of systems with uncertain dynamics. The slightly different dynamics can indeed be seen as uncertainties in the system to be controlled, and simultaneous control rephrased in terms of a robustness task. Robustification, i.e., offsetting uncertainties by suitably designing the control strategy, is a widespread task in automatic control theory, showing up in many applicative domains such as electric circuits or aerospace motion planning. If dynamics are not only subject to static uncertainty, but may also change as time goes, the problem of controlling the system can be recast within the theory of switched and hybrid systems, both in a deterministic and in a probabilistic setting. Our contributions to this research field concern both stabilization (either asymptotic or in finite time) and optimal control, where redundancies and probabilistic tools can be introduced to offset uncertainties.

[0] As anticipated by the recent launch of the FET Flagship on Quantum Technologies

3. Research Program

3.1. Research domain

The activities of CAGE are part of the research in the wide area of control theory. This nowadays mature discipline is still the subject of intensive research because of its crucial role in a vast array of applications.

More specifically, our contributions are in the area of **mathematical control theory**, which is to say that we are interested in the analytical and geometrical aspects of control applications. In this approach, a control system is modeled by a system of equations (of many possible types: ordinary differential equations, partial differential equations, stochastic differential equations, difference equations,...), possibly not explicitly known in all its components, which are studied in order to establish qualitative and quantitative properties concerning the actuation of the system through the control.

Motion planning is, in this respect, a cornerstone property: it denotes the design and validation of algorithms for identifying a control law steering the system from a given initial state to (or close to) a target one. Initial and target positions can be replaced by sets of admissible initial and final states as, for instance, in the motion planning task towards a desired periodic solution. Many specifications can be added to the pure motion planning task, such as robustness to external or endogenous disturbances, obstacle avoidance or penalization criteria. A more abstract notion is that of **controllability**, which denotes the property of a system for which any two states can be connected by a trajectory corresponding to an admissible control law. In mathematical terms, this translates into the surjectivity of the so-called **end-point map**, which associates with a control and an initial state the final point of the corresponding trajectory. The analytical and topological properties of endpoint maps are therefore crucial in analyzing the properties of control systems.

One of the most important additional objective which can be associated with a motion planning task is **optimal control**, which corresponds to the minimization of a cost (or, equivalently, the maximization of a gain) [156]. Optimal control theory is clearly deeply interconnected with calculus of variations, even if the non-interchangeable nature of the time-variable results in some important specific features, such as the occurrence of **abnormal extremals** [120]. Research in optimal control encompasses different aspects, from numerical methods to dynamic programming and non-smooth analysis, from regularity of minimizers to high order optimality conditions and curvature-like invariants.

Another domain of control theory with countless applications is **stabilization**. The goal in this case is to make the system converge towards an equilibrium or some more general safety region. The main difference with respect to motion planning is that here the control law is constructed in feedback form. One of the most important properties in this context is that of **robustness**, i.e., the performance of the stabilization protocol in presence of disturbances or modeling uncertainties. A powerful framework which has been developed to take into account uncertainties and exogenous non-autonomous disturbances is that of hybrid and switched systems [159], [119], [147]. The central tool in the stability analysis of control systems is that of **control Lyapunov function**. Other relevant techniques are based on algebraic criteria or dynamical systems. One of the most important stability property which is studied in the context of control system is **input-to-state stability** [143], which measures how sensitive the system is to an external excitation.

One of the areas where control applications have nowadays the most impressive developments is in the field of **biomedicine and neurosciences**. Improvements both in modeling and in the capability of finely actuating biological systems have concurred in increasing the popularity of these subjects. Notable advances concern, in particular, identification and control for biochemical networks [137] and models for neural activity [106]. Therapy analysis from the point of view of optimal control has also attracted a great attention [140].

Biological models are not the only one in which stochastic processes play an important role. Stock-markets and energy grids are two major examples where optimal control techniques are applied in the non-deterministic setting. Sophisticated mathematical tools have been developed since several decades to allow for such extensions. Many theoretical advances have also been required for dealing with complex systems whose description is based on **distributed parameters** representation and **partial differential equations**. Functional analysis, in particular, is a crucial tool to tackle the control of such systems [153].

Let us conclude this section by mentioning another challenging application domain for control theory: the decision by the European Union to fund a flagship devoted to the development of quantum technologies is a symptom of the role that quantum applications are going to play in tomorrow's society. **Quantum control** is one of the bricks of quantum engineering, and presents many peculiarities with respect to standard control theory, as a consequence of the specific properties of the systems described by the laws of quantum physics. Particularly important for technological applications is the capability of inducing and reproducing coherent state superpositions and entanglement in a fast, reliable, and efficient way [107].

3.2. Scientific foundations

At the core of the scientific activity of the team is the **geometric control** approach, that is, a distinctive viewpoint issued in particular from (elementary) differential geometry, to tackle questions of controllability, observability, optimal control... [70], [111]. The emphasis of such a geometric approach to control theory is put on intrinsic properties of the systems and it is particularly well adapted to study nonlinear and nonholonomic phenomena.

One of the features of the geometric control approach is its capability of exploiting **symmetries and intrinsic structures** of control systems. Symmetries and intrinsic structures can be used to characterize minimizing trajectories, prove regularity properties and describe invariants. An egregious example is given by mechanical systems, which inherently exhibit Lagrangian/Hamiltonian structures which are naturally expressed using the language of symplectic geometry [93]. The geometric theory of quantum control, in particular, exploits the rich geometric structure encoded in the Schrödinger equation to engineer adapted control schemes and to characterize their qualitative properties. The Lie–Galerkin technique that we proposed starting from 2009 [96] builds on this premises in order to provide powerful tests for the controllability of quantum systems defined on infinite-dimensional Hilbert spaces.

Although the focus of geometric control theory is on qualitative properties, its impact can also be disruptive when it is used in combination with quantitative analytical tools, in which case it can dramatically improve the computational efficiency. This is the case in particular in optimal control. Classical optimal control techniques (in particular, Pontryagin Maximum Principle, conjugate point theory, associated numerical methods) can be significantly improved by combining them with powerful modern techniques of geometric optimal control, of the theory of numerical continuation, or of dynamical system theory [152], [139]. Geometric optimal control allows the development of general techniques, applying to wide classes of nonlinear optimal control problems, that can be used to characterize the behavior of optimal trajectories and in particular to establish regularity properties for them and for the cost function. Hence, geometric optimal control can be used to obtain powerful optimal syntheses results and to provide deep geometric insights into many applied problems. Numerical optimal control methods with geometric insight are in particular important to handle subtle situations such as rigid optimal paths and, more generally, optimal syntheses exhibiting abnormal minimizers.

Optimal control is not the only area where the geometric approach has a great impact. Let us mention, for instance, motion planning, where different geometric approaches have been developed: those based on the **Lie algebra** associated with the control system [132], [122], those based on the differentiation of nonlinear flows such as the **return method** [101], [100], and those exploiting the **differential flatness** of the system [105].

Geometric control theory is not only a powerful framework to investigate control systems, but also a useful tool to model and study phenomena that are not *a priori* control-related. Two occurrences of this property play an important role in the activities of CAGE:

- geometric control theory as a tool to investigate properties of mathematical structures;
- geometric control theory as a modeling tool for neurophysical phenomena and for synthesizing biomimetic algorithms based on such models.

Examples of the first type, concern, for instance, hypoelliptic heat kernels [68] or shape optimization [76]. Examples of the second type are inactivation principles in human motricity [79] or neurogeometrical models for image representation of the primary visual cortex in mammals [90].

A particularly relevant class of control systems, both from the point of view of theory and applications, is characterized by the linearity of the controlled vector field with respect to the control parameters. When the controls are unconstrained in norm, this means that the admissible velocities form a distribution in the tangent bundle to the state manifold. If the distribution is equipped with a point-dependent quadratic form (encoding the cost of the control), the resulting geometrical structure is said to be **sub-Riemannian**. Sub-Riemannian geometry appears as the underlying geometry of nonlinear control systems: in a similar way as the linearization of a control system provides local informations which are readable using the Euclidean metric scale, sub-Riemannian geometry provides an adapted non-isotropic class of lenses which are often much more informative. As such, its study is fundamental for control design. The importance of sub-Riemannian geometry goes beyond control theory and it is an active field of research both in differential geometry [130], geometric measure theory [72] and hypoelliptic operator theory [82].

The geometric control approach has historically been related to the development of finite-dimensional control theory. However, its impact in the analysis of distributed parameter control systems and in particular systems of controlled partial differential equations has been growing in the last decades, complementing analytical and numerical approaches, providing dynamical, qualitative and intrinsic insight [99]. CAGE's ambition is to be at the core of this development in the years to come.

4. Application Domains

4.1. First axis: Geometry of vision

A suggestive application of sub-Riemannian geometry and in particular of hypoelliptic diffusion comes from a model of geometry of vision describing the functional architecture of the primary visual cortex V1. In 1958, Hubel and Wiesel (Nobel in 1981) observed that the visual cortex V1 is endowed with the so-called **pinwheel structure**, characterized by neurons grouped into orientation columns, that are sensible both to positions and directions [110]. The mathematical rephrasing of this discovery is that the visual cortex lifts an image from $\mathbf{R}^2$ into the bundle of directions of the plane [97], [136], [138], [109].

A simplified version of the model can be described as follows: neurons of V1 are grouped into orientation columns, each of them being sensitive to visual stimuli at a given point of the retina and for a given direction on it. The retina is modeled by the real plane, i.e., each point is represented by a pair $(x, y) \in \mathbb{R}^2$, while the directions at a given point are modeled by the projective line, i.e. an element θ of the projective line P^1. Hence, the primary visual cortex V1 is modeled by the so called projective tangent bundle $\mathrm{PTR}^2 = \mathbf{R}^2 \times \mathbf{P}^1$. From a neurological point of view, orientation columns are in turn grouped into hypercolumns, each of them being sensitive to stimuli at a given point (x, y) with any direction.

Orientation columns are connected between them in two different ways. The first kind of connections are the vertical (inhibitory) ones, which connect orientation columns belonging to the same hypercolumn and sensible to similar directions. The second kind of connections are the horizontal (excitatory) connections, which connect neurons belonging to different (but not too far) hypercolumns and sensible to the same directions. The resulting metric structure is sub-Riemannian and the model obtained in this way provides a convincing explanation in terms of sub-Riemannian geodesics of gestalt phenomena such as Kanizsa illusory contours.

The sub-Riemannian model for image representation of V1 has a great potential of yielding powerful bio-inspired image processing algorithms [104], [90]. Image inpainting, for instance, can be implemented by reconstructing an incomplete image by activating orientation columns in the missing regions in accordance with sub-Riemannian non-isotropic constraints. The process intrinsically defines an hypoelliptic heat equation on PTR^2 which can be integrated numerically using non-commutative Fourier analysis on a suitable semidiscretization of the group of roto-translations of the plane [88].

We have been working on the model and its software implementation since 2012. This work has been supported by several project, as the ERC starting grant GeCoMethods and the ERC Proof of Concept ARTIV1 of U. Boscain, and the ANR GCM.

A parallel approach that we will pursue and combine with this first one is based on **pattern matching in the group of diffeomorphisms**. We want to extend this approach, already explored in the Riemannian setting [151], [127], to the general sub-Riemannian framework. The paradigm of the approach is the following: consider a distortable object, more or less rigid, discretized into a certain number of points. One may track its distortion by considering the paths drawn by these points. One would however like to know how the object itself (and not its discretized version) has been distorted. The study in [151], [127] shed light on the importance of Riemannian geometry in this kind of problem. In particular, they study the Riemannian submersion obtained by making the group of diffeomorphisms act transitively on the manifold formed by the points of the discretization, minimizing a certain energy so as to take into account the whole object. Settled as such, the problem is Riemannian, but if one considers objects involving connections, or submitted to nonholonomic constraints, like in medical imaging where one tracks the motions of organs, then one comes up with a sub-Riemannian problem. The transitive group is then far bigger, and the aim is to lift curves submitted to these nonholonomic constraints into curves in the set of diffeomorphisms satisfying the corresponding constraints, in a unique way and minimizing an energy (giving rise to a sub-Riemannian structure).

4.2. Second axis: Quantum control

The goal of quantum control is to design efficient protocols for tuning the occupation probabilities of the energy levels of a system. This task is crucial in atomic and molecular physics, with applications ranging from photochemistry to nuclear magnetic resonance and quantum computing. A quantum system may be controlled by exciting it with one or several external fields, such as magnetic or electric fields. The goal of quantum control theory is to adapt the tools originally developed by control theory and to develop new specific strategies that tackle and exploit the features of quantum dynamics (probabilistic nature of wavefunctions and density operators, measure and wavefunction collapse, decoherence, ...). A rich variety of relevant models for controlled quantum dynamics exist, encompassing low-dimensional models (e.g., single-spin systems) and PDEs alike, with deterministic and stochastic components, making it a rich and exciting area of research in control theory.

The controllability of quantum system is a well-established topic when the state space is finite-dimensional [102], thanks to general controllability methods for left-invariant control systems on compact Lie groups [92], [112]. When the state space is infinite-dimensional, it is known that in general the bilinear Schrödinger equation is not exactly controllable [154]. Nevertheless, weaker controllability properties, such as approximate controllability or controllability between eigenstates of the internal Hamiltonian (which are the most relevant physical states), may hold. In certain cases, when the state space is a function space on a 1D manifold, some rather precise description of the set of reachable states has been provided [77]. A similar description for higher-dimensional manifolds seems intractable and at the moment only approximate controllability results are available [128], [134], [113]. The most widely applicable tests for controllability of quantum systems in infinite-dimensional Hilbert spaces are based on the **Lie–Galerkin technique** [96], [85], [86]. They allow, in particular, to show that the controllability property is generic among this class of systems [125].

A family of algorithms which are specific to quantum systems are those based on adiabatic evolution [158], [157], [116]. The basic principle of adiabatic control is that the flow of a slowly varying Hamiltonian can be approximated (up to a phase factor) by a quasi-static evolution, with a precision proportional to the velocity of variation of the Hamiltonian. The advantage of the **adiabatic approach** is that it is constructive and produces control laws which are both smooth and robust to parameter uncertainty. The paradigm is based on the adiabatic perturbation theory developed in mathematical physics [83], [133], [150], where it plays an important role for understanding molecular dynamics. Approximation theory by adiabatic perturbation can be used to describe the evolution of the occupation probabilities of the energy levels of a slowly varying Hamiltonian. Results from the last 15 years, including those by members of our team [64], [89], have highlighted the effectiveness of control techniques based on adiabatic path following.

4.3. Third axis: Stability and uncertain dynamics

Switched and hybrid systems constitute a broad framework for the description of the heterogeneous aspects of systems in which continuous dynamics (typically pertaining to physical quantities) interact with discrete/logical components. The development of the switched and hybrid paradigm has been motivated by a broad range of applications, including automotive and transportation industry [142], energy management [135] and congestion control [126].

Even if both controllability [146] and observability [114] of switched and hybrid systems have attracted much research efforts, the central role in their study is played by the problem of stability and stabilizability. The goal is to determine whether a dynamical or a control system whose evolution is influenced by a time-dependent signal is uniformly stable or can be uniformly stabilized [119], [147]. Uniformity is considered with respect to all signals in a given class. Stability of switched systems lead to several interesting phenomena. For example, even when all the subsystems corresponding to a constant switching law are exponentially stable, the switched systems may have divergent trajectories for certain switching signals [118]. This fact illustrates the fact that stability of switched systems depends not only on the dynamics of each subsystem but also on the properties of the class of switching signals which is considered.

The most common class of switching signals which has been considered in the literature is made of all piecewise constant signals. In this case uniform stability of the system is equivalent to the existence of a common quadratic Lyapunov function [129]. Moreover, provided that the system has finitely many modes, the Lyapunov function can be taken polyhedral or polynomial [80], [81], [103]. A special role in the switched control literature has been played by common quadratic Lyapunov functions, since their existence can be tested rather efficiently (see the surveys [121], [141] and the references therein). It is known, however, that the existence of a common quadratic Lyapunov function is not necessary for the global uniform exponential stability of a linear switched system with finitely many modes. Moreover, there exists no uniform upper bound on the minimal degree of a common polynomial Lyapunov function [124]. More refined tools rely on multiple and non-monotone Lyapunov functions [91]. Let us also mention linear switched systems technics based on the analysis of the Lie algebra generated by the matrices corresponding to the modes of the system [67].

For systems evolving in the plane, more geometrical tests apply, and yield a complete characterization of the stability [84], [73]. Such a geometric approach also yields sufficient conditions for uniform stability in the linear planar case [87].

In many situations, it is interesting for modeling purposes to specify the features of the switched system by introducing **constrained switching rules**. A typical constraint is that each mode is activated for at least a fixed minimal amount of time, called the dwell-time. Switching rules can also be imposed, for instance, by a timed automata. When constraints apply, the common Lyapunov function approach becomes conservative and new tools have to be developed to give more detailed characterizations of stable and unstable systems.

Our approach to constrained switching is based on the idea of relating the analytical properties of the classes of constrained switching laws (shift-invariance, compactness, closure under concatenation, ...) to the stability behavior of the corresponding switched systems. One can introduce **probabilistic uncertainties** by endowing the classes of admissible signals with suitable probability measures. One then looks at the corresponding Lyapunov exponents, whose existence is established by the multiplicative ergodic theorem. The interest of this approach is that probabilistic stability analysis filters out highly ‘exceptional’ worst-case trajectories. Although less explicitly characterized from a dynamical viewpoint than its deterministic counterpart, the probabilistic notion of uniform exponential stability can be studied using several reformulations of Lyapunov exponents proposed in the literature [78], [98], [155].

4.4. Joint theoretical core

The theoretical questions raised by the different applicative area will be pooled in a research axis on the transversal aspects of geometric control theory and sub-Riemannian structures.

We recall that sub-Riemannian geometry is a generalization of Riemannian geometry, whose birth dates back to Carathéodory’s seminal paper on the foundations of Carnot thermodynamics [94], followed by E. Cartan’s address at the International Congress of Mathematicians in Bologna [95]. In the last twenty years, sub-Riemannian geometry has emerged as an independent research domain, with a variety of motivations and ramifications in several parts of pure and applied mathematics. Let us mention geometric analysis, geometric measure theory, stochastic calculus and evolution equations together with applications in mechanics and optimal control (motion planning, robotics, nonholonomic mechanics, quantum control) [62], [63].

One of the main open problems in sub-Riemannian geometry concerns the regularity of length-minimizers [65], [131]. Length-minimizers are solutions to a variational problem with constraints and satisfy a first-order necessary condition resulting from the Pontryagin Maximum Principle (PMP). Solutions of the PMP are either *normal* or *abnormal*. Normal length-minimizer are well-known to be smooth, i.e., C^∞, as it follows by the Hamiltonian nature of the PMP. The question of regularity is then reduced to abnormal length-minimizers. If the sub-Riemannian structure has step 2, then abnormal length-minimizers can be excluded and thus every length-minimizer is smooth. For step 3 structures, the situation is already more complicated and smoothness of length-minimizers is known only for Carnot groups [115], [149]. The question of regularity of length-minimizers is not restricted to the smoothness in the C^∞ sense. A recent result prove that length-minimizers, for sub-Riemannian structures of any step, cannot have corner-like singularities [108]. When the sub-Riemannian structure is analytic, more is known on the size of the set of points where a length-minimizer can lose analyticity [148], regardless of the rank and of the step of the distribution.

An interesting set of recent results in sub-Riemannian geometry concerns the extension to such a setting of the Riemannian notion of sectional curvature. The curvature operator can be introduced in terms of the symplectic invariants of the Jacobi curve [69], [117], [66], a curve in the Lagrange Grassmannian related to the linearization of the Hamiltonian flow. Alternative approaches to curvatures in metric spaces are based either on the associated heat equation and the generalization of the curvature-dimension inequality [74], [75] or on optimal transport and the generalization of Ricci curvature [145], [144], [123], [71].

5. New Software and Platforms

5.1. 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 Andréani, Jinyan Liu, Joseph Frédéric Bonnans and Pierre Martinon
- Contact: Pierre Martinon
- URL: http://bocop.org

5.2. 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. New Results

6.1. Geometry of vision and sub-Riemannian geometry: new results

Let us list here our new results in the geometry of vision axis and, more generally, on hypoelliptic diffusion and sub-Riemannian geometry.

- In [12] we propose a variational model for joint image reconstruction and motion estimation applicable to spatiotemporal imaging. This model consists of two parts, one that conducts image reconstruction in a static setting and another that estimates the motion by solving a sequence of coupled indirect image registration problems, each formulated within the large deformation diffeomorphic metric mapping framework. The proposed model is compared against alternative approaches (optical flow based model and diffeomorphic motion models). Next, we derive efficient algorithms for a time-discretized setting and show that the optimal solution of the time-discretized formulation is consistent with that of the time-continuous one. The complexity of the algorithm is characterized and we conclude by giving some numerical examples in 2D space + time tomography with very sparse and/or highly noisy data.
- The article [16] presents a method to incorporate a deformation prior in image reconstruction via the formalism of deformation modules. The framework of deformation modules allows to build diffeomorphic deformations that satisfy a given structure. The idea is to register a template image against the indirectly observed data via a modular deformation, incorporating this way the deformation prior in the reconstruction method. We show that this is a well-defined regularization method (proving existence, stability and convergence) and present numerical examples of reconstruction from 2-D tomographic simulations and partially-observed images.
- The article [28] adapts the framework of metamorphosis to the resolution of inverse problems with shape prior. The metamorphosis framework allows to transform an image via a balance between geometrical deformations and changes in intensities (that can for instance correspond to the appearance of a new structure). The idea developed here is to reconstruct an image from noisy and indirect observations by registering, via metamorphosis, a template to the observed data. Unlike a registration with only geometrical changes, this framework gives good results when intensities of the template are poorly chosen. We show that this method is a well-defined regularization method (proving existence, stability and convergence) and present several numerical examples.
- In [8] we prove the C^1 regularity for a class of abnormal length-minimizers in rank 2 sub-Riemannian structures. As a consequence of our result, all length-minimizers for rank 2 sub-Riemannian structures of step up to 4 are of class C^1
- In [33] we show that, for a sub-Laplacian Δ on a 3-dimensional manifold M, no point interaction centered at a point $q_0 \in M$ exists.
- In [39] we consider a one-parameter family of Grushin-type singularities on surfaces, and discuss the possible diffusions that extend Brownian motion to the singularity. This gives a quick proof and clear intuition for the fact that heat can only cross the singularity for an intermediate range of the parameter. When crossing is possible and the singularity consists of one point, we give a complete description of these diffusions, and we describe a “best" extension, which respects the isometry group of the surface and also realizes the unique symmetric one-point extension of the Brownian motion, in the sense of Chen-Fukushima. This extension, however, does not correspond to

the bridging extension, which was introduced by Boscain-Prandi, when they previously considered self-adjoint extensions of the Laplace-Beltrami operator on the Riemannian part for these surfaces. We clarify that several of the extensions they considered induce diffusions that are carried by the Marin compactification at the singularity, which is much larger than the (one-point) metric completion. In the case when the singularity is more than one-point, a complete classification of diffusions extending Brownian motion would be unwieldy. Nonetheless, we again describe a "best" extension which respects the isometry group, and in this case, this diffusion corresponds to the bridging extension. A prominent role is played by Bessel processes (of every real dimension) and the classical theory of one-dimensional diffusions and their boundary conditions.

- In [50] we study the notion of geodesic curvature of smooth horizontal curves parametrized by arc length in the Heisenberg group, that is the simplest sub-Riemannian structure. Our goal is to give a metric interpretation of this notion of geodesic curvature as the first corrective term in the Taylor expansion of the distance between two close points of the curve.

We would also like to mention the monograph [30] and the PhD thesis of Mathieu Kohli [3].

6.2. Quantum control: new results

Let us list here our new results in quantum control theory.

- In [29], we discuss the compatibility between the rotating-wave and the adiabatic approximations for controlled quantum systems. Although the paper focuses on applications to two-level quantum systems, the main results apply in higher dimension. Under some suitable hypotheses on the time scales, the two approximations can be combined. As a natural consequence of this, it is possible to design control laws achieving transitions of states between two energy levels of the Hamiltonian that are robust with respect to inhomogeneities of the amplitude of the control input.
- In [34] we study one-parametric perturbations of finite dimensional real Hamiltonians depending on two controls, and we show that generically in the space of Hamiltonians, conical intersections of eigenvalues can degenerate into semi-conical intersections of eigenvalues. Then, through the use of normal forms, we study the problem of ensemble controllability between the eigenstates of a generic Hamiltonian.
- In [35] we discuss which controllability properties of classical Hamiltonian systems are preserved after quantization. We discuss some necessary and some sufficient conditions for small-time controllability of classical systems and quantum systems using the WKB method. In particular, we investigate the conjecture that if the classical system is not small-time controllable, then the corresponding quantum system is not small-time controllable either.
- In [40] we study the controllability problem for a symmetric-top molecule, both for its classical and quantum rotational dynamics. As controlled fields we consider three orthogonally polarized electric fields which interact with the electric dipole of the molecule. We characterize the controllability in terms of the dipole position: when it lies along the symmetry axis of the molecule nor the classical neither the quantum dynamics are controllable, due to the presence of a conserved quantity, the third component of the total angular momentum; when it lies in the orthogonal plane to the symmetry axis, a quantum symmetry arises, due to the superposition of symmetric states, which as no classical counterpart. If the dipole is neither along the symmetry axis nor orthogonal to it, controllability for the classical dynamics and approximate controllability for the quantum dynamics is proved to hold.

We would also like to mention the defense of the PhD thesis of Nicolas Augier (not yet on TEL) on the subject.

6.3. Stability and uncertain dynamics: new results

Let us list here our new results about stability and stabilization of control systems, on the properties of systems with uncertain dynamics.

- In an open channel, a hydraulic jump is an abrupt transition between a torrential (super-critical) flow and a fluvial (subcritical) flow. In [9] hydraulic jumps are represented by discontinuous shock solutions of hyperbolic Saint-Venant equations. Using a Lyapunov approach, we prove that we can stabilize the state of the system in H^2-norm as well as the hydraulic jump location, with simple feedback boundary controls and an arbitrary decay rate, by appropriately choosing the gains of the feedback boundary controls.
- In [10], we study the exponential stabilization of a shock steady state for the inviscid Burgers equation on a bounded interval. Our analysis relies on the construction of an explicit strict control Lyapunov function. We prove that by appropriately choosing the feedback boundary conditions, we can stabilize the state as well as the shock location to the desired steady state in H^2-norm, with an arbitrary decay rate.
- We develop in [19] a method ensuring robustness properties to bang-bang strategies , for general nonlinear control systems. Our main idea is to add bang arcs in the form of needle-like variations of the control. With such bang-bang controls having additional degrees of freedom, steering the control system to some given target amounts to solving an overdetermined nonlinear shooting problem, what we do by developing a least-square approach. In turn, we design a criterion to measure the quality of robustness of the bang-bang strategy, based on the singular values of the end-point mapping, and which we optimize. Our approach thus shows that redundancy implies robustness, and we show how to achieve some compromises in practice, by applying it to the attitude control of a 3d rigid body.
- Partial stability characterizes dynamical systems for which only a part of the state variables exhibits a stable behavior. In his book on partial stability, Vorotnikov proposed a sufficient condition to establish this property through a Lyapunov-like function whose total derivative is upper-bounded by a negative definite function involving only the sub-state of interest. In [20], we show with a simple two-dimensional system that this statement is wrong in general. More precisely, we show that the convergence rate of the relevant state variables may not be uniform in the initial state. We also discuss the impact of this lack of uniformity on the connected issue of robustness with respect to exogenous disturbances.
- The paper [21] elaborates control strategies to prevent clustering effects in opinion formation models. This is the exact opposite of numerous situations encountered in the literature where, on the contrary, one seeks controls promoting consensus. In order to promote declustering, instead of using the classical variance that does not capture well the phenomenon of dispersion, we introduce an entropy-type functional that is adapted to measuring pairwise distances between agents. We then focus on a Hegselmann-Krause-type system and design declustering sparse controls both in finite-dimensional and kinetic models. We provide general conditions characterizing whether clustering can be avoided as function of the initial data. Such results include the description of black holes (where complete collapse to consensus is not avoidable), safety zones (where the control can keep the system far from clustering), basins of attraction (attractive zones around the clustering set) and collapse prevention (when convergence to the clustering set can be avoided).
- The goal of [23] is to compute a boundary control of reaction-diffusion partial differential equation. The boundary control is subject to a constant delay, whereas the equation may be unstable without any control. For this system equivalent to a parabolic equation coupled with a transport equation, a prediction-based control is explicitly computed. To do that we decompose the infinite-dimensional system into two parts: one finite-dimensional unstable part, and one stable infinite-dimensional part. An finite-dimensional delay controller is computed for the unstable part, and it is shown that this controller succeeds in stabilizing the whole partial differential equation. The proof is based on a an explicit form of the classical Artstein transformation, and an appropriate Lyapunov function. A numerical simulation illustrate the constructive design method.
- Given a linear control system in a Hilbert space with a bounded control operator, we establish in [26] a characterization of exponential stabilizability in terms of an observability inequality. Such dual characterizations are well known for exact (null) controllability. Our approach exploits

classical Fenchel duality arguments and, in turn, leads to characterizations in terms of observability inequalities of approximately null controllability and of α-null controllability. We comment on the relationships between those various concepts, at the light of the observability inequalities that characterize them.

- In [37] we propose an extension of the theory of control sets to the case of inputs satisfying a dwell-time constraint. Although the class of such inputs is not closed under concatenation, we propose a suitably modified definition of control sets that allows to recover some important properties known in the concatenable case. In particular we apply the control set construction to dwell-time linear switched systems, characterizing their maximal Lyapunov exponent looking only at trajectories whose angular component is periodic. We also use such a construction to characterize supports of invariant measures for random switched systems with dwell-time constraints.
- In [41] we study asymptotic stability of continuous-time systems with mode-dependent guaranteed dwell time. These systems are reformulated as special cases of a general class of mixed (discrete-continuous) linear switching systems on graphs, in which some modes correspond to discrete actions and some others correspond to continuous-time evolutions. Each discrete action has its own positive weight which accounts for its time-duration. We develop a theory of stability for the mixed systems; in particular, we prove the existence of an invariant Lyapunov norm for mixed systems on graphs and study its structure in various cases, including discrete-time systems for which discrete actions have inhomogeneous time durations. This allows us to adapt recent methods for the joint spectral radius computation (Gripenberg's algorithm and the Invariant Polytope Algorithm) to compute the Lyapunov exponent of mixed systems on graphs.
- Given a discrete-time linear switched system associated with a finite set of matrices, we consider the measures of its asymptotic behavior given by, on the one hand, its deterministic joint spectral radius and, on the other hand, its probabilistic joint spectral radius for Markov random switching signals with given transition matrix and corresponding invariant probability. In [42], we investigate the cases of equality between the two measures.
- In [45] we address the question of the exponential stability for the C^1 norm of general 1-D quasilinear systems with source terms under boundary conditions. To reach this aim, we introduce the notion of basic C^1 Lyapunov functions, a generic kind of exponentially decreasing function whose existence ensures the exponential stability of the system for the C^1 norm. We show that the existence of a basic C^1 Lyapunov function is subject to two conditions: an interior condition, intrinsic to the system, and a condition on the boundary controls. We give explicit sufficient interior and boundary conditions such that the system is exponentially stable for the C^1 norm and we show that the interior condition is also necessary to the existence of a basic C^1 Lyapunov function. Finally, we show that the results conducted in this article are also true under the same conditions for the exponential stability in the C^p norm, for any $p \geq 1$.
- In [46] we study the exponential stability for the C^1 norm of general 2×2 1-D quasilinear hyperbolic systems with source terms and boundary controls. When the eigenvalues of the system have the same sign, any nonuniform steady-state can be stabilized using boundary feedbacks that only depend on measurements at the boundaries and we give explicit conditions on the gain of the feedback. In other cases, we exhibit a simple numerical criterion for the existence of basic C^1 Lyapunov function, a natural candidate for a Lyapunov function to ensure exponential stability for the C^1 norm.
- In [47] we study the exponential stability in the H^2 norm of the nonlinear Saint-Venant (or shallow water) equations with arbitrary friction and slope using a single Proportional-Integral (PI) control at one end of the channel. Using a local dissipative entropy we find a simple and explicit condition on the gain the PI control to ensure the exponential stability of any steady-states. This condition is independent of the slope, the friction, the length of the river, the inflow disturbance and, more surprisingly, the steady-state considered. When the inflow disturbance is time-dependent and no steady-state exist, we still have the Input-to-State stability of the system, and we show that changing slightly the PI control enables to recover the exponential stability of slowly varying trajectories.

- In [48], we address the problem of the exponential stability of density-velocity systems with boundary conditions. Density-velocity systems are omnipresent in physics as they encompass all systems that consist in a flux conservation and a momentum equation. In this paper we show that any such system can be stabilized exponentially quickly in the H^2 norm using simple local feedbacks, provided a condition on the source term which holds for most physical systems, even when it is not dissipative. Besides, the feedback laws obtained only depends on the target values at the boundaries, which implies that they do not depend on the expression of the source term or the force applied on the system and makes them very easy to implement in practice and robust to model errors. For instance, for a river modeled by Saint-Venant equations this means that the feedback laws do not require any information on the friction model, the slope or the shape of the channel considered. This feat is obtained by showing the existence of a basic H^2 Lyapunov functions and we apply it to numerous systems: the general Saint-Venant equations, the isentropic Euler equations, the motion of water in rigid-pipe, the osmosis phenomenon, etc.
- The general context of [56] is the feedback control of an infinite-dimensional system so that the closed-loop system satisfies a fading-memory property and achieves the setpoint tracking of a given reference signal. More specifically , this paper is concerned with the Proportional Integral (PI) regulation control of the left Neumann trace of a one-dimensional reaction-diffusion equation with a delayed right Dirichlet boundary control. In this setting, the studied reaction-diffusion equation might be either open-loop stable or unstable. The proposed control strategy goes as follows. First, a finite-dimensional truncated model that captures the unstable dynamics of the original infinite-dimensional system is obtained via spectral decomposition. The truncated model is then augmented by an integral component on the tracking error of the left Neumann trace. After resorting to the Artstein transformation to handle the control input delay, the PI controller is designed by pole shifting. Stability of the resulting closed-loop infinite-dimensional system, consisting of the original reaction-diffusion equation with the PI controller, is then established thanks to an adequate Lyapunov function. In the case of a time-varying reference input and a time-varying distributed disturbance, our stability result takes the form of an exponential Input-to-State Stability (ISS) estimate with fading memory. Finally, another exponential ISS estimate with fading memory is established for the tracking performance of the reference signal by the system output. In particular, these results assess the setpoint regulation of the left Neumann trace in the presence of distributed perturbations that converge to a steady-state value and with a time-derivative that converges to zero. Numerical simulations are carried out to illustrate the efficiency of our control strategy.
- There exist many ways to stabilize an infinite-dimensional linear autonomous control systems when it is possible. Anyway, finding an exponentially stabilizing feedback control that is as simple as possible may be a challenge. The Riccati theory provides a nice feedback control but may be computationally demanding when considering a discretization scheme. Proper Orthogonal Decomposition (POD) offers a popular way to reduce large-dimensional systems. In [59], we establish that, under appropriate spectral assumptions, an exponentially stabilizing feedback Riccati control designed from a POD finite-dimensional approximation of the system stabilizes as well the infinite-dimensional control system.
- In [60] we consider a 1-D linear transport equation on the interval $(0, L)$, with an internal scalar control. We prove that if the system is controllable in a periodic Sobolev space of order greater than 1, then the system can be stabilized in finite time, and we give an explicit feedback law.
- In [61] we use the backstepping method to study the stabilization of a 1-D linear transport equation on the interval $(0, L)$, by controlling the scalar amplitude of a piecewise regular function of the space variable in the source term. We prove that if the system is controllable in a periodic Sobolev space of order greater than 1, then the system can be stabilized exponentially in that space and, for any given decay rate, we give an explicit feedback law that achieves that decay rate.

6.4. Controllability: new results

Let us list here our new results on controllability beyond the quantum control framework.

- In [13], we study approximate and exact controllability of linear difference equations using as a basic tool a representation formula for its solution in terms of the initial condition, the control, and some suitable matrix coefficients. When the delays are commensurable, approximate and exact controllability are equivalent and can be characterized by a Kalman criterion. The paper focuses on providing characterizations of approximate and exact controllability without the commensurability assumption. In the case of two-dimensional systems with two delays, we obtain an explicit characterization of approximate and exact controllability in terms of the parameters of the problem. In the general setting, we prove that approximate controllability from zero to constant states is equivalent to approximate controllability in L^2. The corresponding result for exact controllability is true at least for two-dimensional systems with two delays.
- In [14] we consider the 2D incompressible Navier-Stokes equation in a rectangle with the usual no-slip boundary condition prescribed on the upper and lower boundaries. We prove that for any positive time, for any finite energy initial data, there exist controls on the left and right boundaries and a distributed force, which can be chosen arbitrarily small in any Sobolev norm in space, such that the corresponding solution is at rest at the given final time. Our work improves earlier results where the distributed force is small only in a negative Sobolev space. It is a further step towards an answer to Jacques-Louis Lions' question about the small-time global exact boundary controllability of the Navier-Stokes equation with the no-slip boundary condition, for which no distributed force is allowed. Our analysis relies on the well-prepared dissipation method already used for Burgers and for Navier-Stokes in the case of the Navier slip-with-friction boundary condition. In order to handle the larger boundary layers associated with the no-slip boundary condition, we perform a preliminary regularization into analytic functions with arbitrarily large analytic radius and prove a long-time nonlinear Cauchy-Kovalevskaya estimate relying only on horizontal analyticity.
- We consider the wave equation on a closed Riemannian manifold. We observe the restriction of the solutions to a measurable subset ω along a time interval $[0, T]$ with $T > 0$. It is well known that, if ω is open and if the pair (ω, T) satisfies the Geometric Control Condition then an observability inequality is satisfied, comparing the total energy of solutions to their energy localized in $\omega \times (0, T)$. The observability constant $C_T(\omega)$ is then defined as the infimum over the set of all nontrivial solutions of the wave equation of the ratio of localized energy of solutions over their total energy. In [17], we provide estimates of the observability constant based on a low/high frequency splitting procedure allowing us to derive general geometric conditions guaranteeing that the wave equation is observable on a measurable subset ω. We also establish that, as $T \to +\infty$, the ratio $C_T(\omega)/T$ converges to the minimum of two quantities: the first one is of a spectral nature and involves the Laplacian eigenfunctions, the second one is of a geometric nature and involves the average time spent in ω by Riemannian geodesics.
- In [22] we consider the problem of controlling parabolic semilinear equations arising in population dynamics, either in finite time or infinite time. These are the monostable and bistable equations on $(0, L)$ for a density of individuals $0 \leq y(t, x) \leq 1$, with Dirichlet controls taking their values in $[0, 1]$. We prove that the system can never be steered to extinction (steady state 0) or invasion (steady state 1) in finite time, but is asymptotically controllable to 1 independently of the size L, and to 0 if the length L of the interval domain is less than some threshold value L^*, which can be computed from transcendental integrals. In the bistable case, controlling to the other homogeneous steady state $0 < \theta < 1$ is much more intricate. We rely on a staircase control strategy to prove that θ can be reached in finite time if and only if $L < L^*$. The phase plane analysis of those equations is instrumental in the whole process. It allows us to read obstacles to controllability, compute the threshold value for domain size as well as design the path of steady states for the control strategy.
- The paper [27] deals with the controllability problem of a linearized Korteweg-de Vries equation on bounded interval. The system has a homogeneous Dirichlet boundary condition and a homogeneous Neumann boundary condition at the right end-points of the interval, a non homogeneous Dirichlet boundary condition at the left end-point which is the control. We prove the null controllability by using a backstepping approach, a method usually used to handle stabilization problems.

- The paper [44] is devoted to the controllability of a general linear hyperbolic system in one space dimension using boundary controls on one side. Under precise and generic assumptions on the boundary conditions on the other side, we previously established the optimal time for the null and the exact controllability for this system for a generic source term. In this work, we prove the null-controllability for any time greater than the optimal time and for any source term. Similar results for the exact controllability are also discussed.
- Given any measurable subset ω of a closed Riemannian manifold and given any $T > 0$, we study in [49] the smallest average time over $[0, T]$ spent by all geodesic rays in ω. This quantity appears naturally when studying observability properties for the wave equation on M, with ω as an observation subset.
- Our goal is to study controllability and observability properties of the 1D heat equation with internal control (or observation) set an interval of size $\epsilon \to 0$. For any ϵ fixed, the heat equation is controllable in any time $T > 0$. It is known that depending on arithmetic properties of the center of the interval, there may exist a minimal time of pointwise control of the heat equation. We relate these two phenomena in [54].
- Our goal in [55] is to relate the observation (or control) of the wave equation on observation domains which evolve in time with some dynamical properties of the geodesic flow. In comparison to the case of static domains of observation, we show that the observability of the wave equation in any dimension of space can be improved by allowing the domain of observation to move.
- In [57] we consider the controllability problem for finite-dimensional linear autonomous control systems with nonnegative controls. Despite the Kalman condition, the unilateral nonnegativity control constraint may cause a positive minimal controllability time. When this happens, we prove that, if the matrix of the system has a real eigenvalue, then there is a minimal time control in the space of Radon measures, which consists of a finite sum of Dirac impulses. When all eigenvalues are real, this control is unique and the number of impulses is less than half the dimension of the space. We also focus on the control system corresponding to a finite-difference spatial discretization of the one-dimensional heat equation with Dirichlet boundary controls, and we provide numerical simulations.

Let us also mention t the book chapter [31], which has been published this year.

6.5. Optimal control: new results

Let us list here our new results in optimal control theory beyond the sub-Riemannian framework.

- In order to determine the optimal strategy to run a race on a curved track according to the lane number, we introduce in [7] 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 types of track since different designs of tracks lead to straights of different lengths. In particular, we find that the tracks with shorter straights lead to better performances, while the double bend track with the longest straight leads to the worst performances and the biggest difference between lanes. Then for a race with two runners, we introduce a psychological interaction: there is an attraction to follow someone just ahead, but after being overtaken, there is a delay before any benefit from this interaction occurs. We provide numerical simulations in different cases. Overall, the results agree with the IAAF rules for lane draws in competition, where the highest ranked athletes get the center lanes, the next ones the outside lanes, while the lowest ranked athletes get the inside lanes.
- Consider a general nonlinear optimal control problem in finite dimension, with constant state and/or control delays. By the Pontryagin Maximum Principle, any optimal trajectory is the projection of a Pontryagin extremal. In [11] we establish that, under appropriate assumptions, Pontryagin extremals depend continuously on the parameter delays, for adequate topologies. The proof of the continuity of the trajectory and of the control is quite easy, however, for the adjoint vector, the proof requires

a much finer analysis. The continuity property of the adjoint with respect to the parameter delay opens a new perspective for the numerical implementation of indirect methods, such as the shooting method. We also discuss the sharpness of our assumptions.

- In [15] we are concerned about the controllability of a general linear hyperbolic system in one space dimension using boundary controls on one side. More precisely, we establish the optimal time for the null and exact controllability of the hyperbolic system under some is a generic setting. We also present examples which yield that the generic requirement is necessary. Our approach is based on the backstepping method paying a special attention on the construction of the kernel and the selection of controls.
- A new approach to estimate traffic energy consumption via traffic data aggregation in (speed,acceleration) probability distributions is proposed in [18]. 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.
- In [24] we consider a spectral optimal design problem involving the Neumann traces of the Dirichlet-Laplacian eigenfunctions on a smooth bounded open subset Ω of $\mathbf{R}^n$. The cost functional measures the amount of energy that Dirichlet eigenfunctions concentrate on the boundary and that can be recovered with a bounded density function. We first prove that, assuming a L^1 constraint on densities, the so-called *Rellich functions* maximize this functional.Motivated by several issues in shape optimization or observation theory where it is relevant to deal with bounded densities, and noticing that the L^∞-norm of *Rellich functions* may be large, depending on the shape of Ω, we analyze the effect of adding pointwise constraints when maximizing the same functional. We investigate the optimality of *bang-bang* functions and *Rellich densities* for this problem. We also deal with similar issues for a close problem, where the cost functional is replaced by a spectral approximation. Finally, this study is completed by the investigation of particular geometries and is illustrated by several numerical simulations.
- In [25] we consider the task of solving an aircraft trajectory optimization problem where the system dynamics have been estimated from recorded data. Additionally, we want to avoid optimized trajectories that go too far away from the domain occupied by the data, since the model validity is not guaranteed outside this region. This motivates the need for a proximity indicator between a given trajectory and a set of reference trajectories. In this presentation, we propose such an indicator based on a parametric estimator of the training set density. We then introduce it as a penalty term in the optimal control problem. Our approach is illustrated with an aircraft minimal consumption problem and recorded data from real flights. We observe in our numerical results the expected trade-off between the consumption and the penalty term.
- In [36] we study how bad can be the singularities of a time-optimal trajectory of a generic control affine system. In the case where the control is scalar and belongs to a closed interval it was recently shown that singularities cannot be, generically, worse than finite order accumulations of Fuller points, with order of accumulation lower than a bound depending only on the dimension of the manifold where the system is set. We extend here such a result to the case where the control has an even number of scalar components and belongs to a closed ball.
- In [38] we develop a geometric analysis and a numerical algorithm, based on indirect methods, to solve optimal guidance of endo-atmospheric launch vehicle systems under mixed control-state

constraints. Two main difficulties are addressed. First, we tackle the presence of Euler singularities by introducing a representation of the configuration manifold in appropriate local charts. In these local coordinates, not only the problem is free from Euler singularities but also it can be recast as an optimal control problem with only pure control constraints. The second issue concerns the initialization of the shooting method. We introduce a strategy which combines indirect methods with homotopies, thus providing high accuracy. We illustrate the efficiency of our approach by numerical simulations on missile interception problems under challenging scenarios.

- We introduce and study in [51] the turnpike property for time-varying shapes, within the viewpoint of optimal control. We focus here on second-order linear parabolic equations where the shape acts as a source term and we seek the optimal time-varying shape that minimizes a quadratic criterion. We first establish existence of optimal solutions under some appropriate sufficient conditions. We then provide necessary conditions for optimality in terms of adjoint equations and, using the concept of strict dissipativity, we prove that state and adjoint satisfy the measure-turnpike property, meaning that the extremal time-varying solution remains essentially close to the optimal solution of an associated static problem. We show that the optimal shape enjoys the exponential turnpike property in term of Hausdorff distance for a Mayer quadratic cost. We illustrate the turnpike phenomenon in optimal shape design with several numerical simulations.
- The work [52] proposes a new approach to optimize the consumption of a hybrid electric vehicle taking into account the traffic conditions. The method is based on a bi-level decomposition in order to make the implementation suitable for online use. The offline lower level computes cost maps thanks to a stochastic optimization that considers the influence of traffic, in terms of speed/acceleration probability distributions. At the online upper level, a deterministic optimization computes the ideal state of charge at the end of each road segment, using the computed cost maps. Since the high computational cost due to the uncertainty of traffic conditions has been managed at the lower level, the upper level is fast enough to be used online in the vehicle. Errors due to discretization and computation in the proposed algorithm have been studied. Finally, we present numerical simulations using actual traffic data, and compare the proposed bi-level method to a deterministic optimization with perfect information about traffic conditions. The solutions show a reasonable over-consumption compared with deterministic optimization, and manageable computational times for both the offline and online parts.
- An extension of the bi-level optimization for the energy management of hybrid electric vehicles (HEVs) proposed in [52] to the eco-routing problem is presented in [53]. Using the knowledge of traffic conditions over the entire road network, we search both the optimal path and state of charge trajectory. This problem results in finding the shortest path on a weighted graph whose nodes are (position, state of charge) pairs for the vehicle, the edge cost being evaluated thanks to the cost maps from optimization at the 'micro' level of a bi-level decomposition. The error due to the discretization of the state of charge is proven to be linear if the cost maps are Lipschitz. The classical A^* algorithm is used to solve the problem, with a heuristic based on a lower bound of the energy needed to complete the travel. The eco-routing method is validated by numerical simulations and compared to the fastest path on a synthetic road network.
- In [58] we study a driftless system on a three-dimensional manifold driven by two scalar controls. We assume that each scalar control has an independent bound on its modulus and we prove that, locally around every point where the controlled vector fields satisfy some suitable nondegeneracy Lie bracket condition, every time-optimal trajectory has at most five bang or singular arcs. The result is obtained using first-and second-order necessary conditions for optimality.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

Contract CIFRE with ArianeGroup (les Mureaux), 2019–2021, funding the thesis of A. Nayet. Participants : M. Cerf (ArianeGroup), E. Trélat (coordinator).

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

- ANR SRGI, for *Sub-Riemannian Geometry and Interactions*, coordinated by **Emmanuel Trélat**, started in 2015 and runs until 2020. Other partners: Toulon University and Grenoble University. SRGI deals with sub-Riemannian geometry, hypoelliptic diffusion and geometric control.
- ANR Finite4SoS, for *Commande et estimation en temps fini pour les Systèmes de Systèmes*, coordinated by Wilfrid Perruquetti, started in 2015 and run up to this year. Other partners: Inria Lille, CAOR - ARMINES. Finite4SoS aims at developing a new promising framework to address control and estimation issues of Systems of Systems subject to model diversity, while achieving robustness as well as severe time response constraints.
- ANR QUACO, for *QUAntum COntrol: PDE systems and MRI applications*, coordinated by Thomas Chambrion, started in 2017 and runs until 2021. Other partners: Lorraine University. QUACO aims at contributing to quantum control theory in two directions: improving the comprehension of the dynamical properties of controlled quantum systems in infinite-dimensional state spaces, and improve the efficiency of control algorithms for MRI.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

Program: H2020-EU.1.3.1. - Fostering new skills by means of excellent initial training of researchers

Call for proposal: MSCA-ITN-2017 - Innovative Training Networks

Project acronym: QUSCO

Project title: Quantum-enhanced Sensing via Quantum Control

Duration: From November 2017 to October 2021.

Coordinator: Christiane Koch

Coordinator for the participant Inria: Ugo Boscain

Abstract: Quantum technologies aim to exploit quantum coherence and entanglement, the two essential elements of quantum physics. Successful implementation of quantum technologies faces the challenge to preserve the relevant nonclassical features at the level of device operation. It is thus deeply linked to the ability to control open quantum systems. The currently closest to market quantum technologies are quantum communication and quantum sensing. The latter holds the promise of reaching unprecedented sensitivity, with the potential to revolutionize medical imaging or structure determination in biology or the controlled construction of novel quantum materials. Quantum control manipulates dynamical processes at the atomic or molecular scale by means of specially tailored external electromagnetic fields. The purpose of QuSCo is to demonstrate the enabling capability of quantum control for quantum sensing and quantum measurement, advancing this field by systematic use of quantum control methods. QuSCo will establish quantum control as a vital part for progress in quantum technologies. QuSCo will expose its students, at the same time, to fundamental questions of quantum mechanics and practical issues of specific applications. Albeit challenging, this reflects our view of the best possible training that the field of quantum technologies can offer. Training in scientific skills is based on the demonstrated tradition of excellence in research of the consortium. It will be complemented by training in communication and commercialization. The latter builds on strong industry participation whereas the former existing expertise on visualization and gamification and combines it with more traditional means of outreach to realize target audience specific public engagement strategies.

8.3. International Research Visitors

8.3.1. Internships

Rosa Kowalewski made an internship under the supervision of Barbara Gris from January to May 2019.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. Member of the Organizing Committees

- Ugo Boscain, Jean-Michel Coron, and Mario Sigalotti organized the workshop “Quantum day: analysis and control”, December 16, Paris
- Ugo Boscain and Emmanuel Trélat organized the minisymposium “Degenerate diffusion processes and their control” at Equadiff 2019, July 8-12, Leiden, The Netherlands
- Ugo Boscain organized (with Aleksey Kostenko and Konstantin Pankrashkin) the workshop “Self-adjoint Extensions in New Settings”, October 6–12, Mathematisches Forschungsinstitut Oberwolfach, Germany
- Barbara Gris organized the minisymposium “Analyse des formes anatomiques” at Smai, May 13–17, Guidel Plages (Morbihan)
- Barbara Gris organized the workshop “Shape analysis in biology”, November 21–22, Paris
- Emmanuel Trélat organized (with Roland Herzog) the workshop “New Trends in PDE-constrained Optimization”, October 14–18, RICAM, Linz, Austria

9.1.2. Scientific Events: Selection

9.1.2.1. Member of the Conference Program Committees

Emmanuel Trélat was member of the Conference Program Committee of the SIAM Conference on Control and its Applications (CT 2019), Chengdu, China.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Ugo Boscain is Associate editor of SIAM Journal of Control and Optimization
- Ugo Boscain is Managing editor of Journal of Dynamical and Control Systems
- Jean-Michel Coron is Editor-in-chief of Comptes Rendus Mathématique
- Jean-Michel Coron is Member of the editorial board of Journal of Evolution Equations
- Jean-Michel Coron is Member of the editorial board of Asymptotic Analysis
- Jean-Michel Coron is Member of the editorial board of ESAIM : Control, Optimisation and Calculus of Variations
- Jean-Michel Coron is Member of the editorial board of Applied Mathematics Research Express
- Jean-Michel Coron is Member of the editorial board of Advances in Differential Equations
- Jean-Michel Coron is Member of the editorial board of Math. Control Signals Systems
- Jean-Michel Coron is Member of the editorial board of Annales de l’IHP, Analyse non linéaire
- Mario Sigalotti is Associate editor of ESAIM : Control, Optimisation and Calculus of Variations
- Mario Sigalotti is Associate editor of Journal on Dynamical and Control Systems
- Emmanuel Trélat is Editor-in-chief of ESAIM : Control, Optimisation and Calculus of Variations

- Emmanuel Trélat is Associate editor of SIAM Review
- Emmanuel Trélat is Associate editor of Syst. Cont. Letters
- Emmanuel Trélat is Associate editor of J. Dynam. Cont. Syst.
- Emmanuel Trélat is Associate editor of Bollettino dell'Unione Matematica Italiana
- Emmanuel Trélat is Associate editor of ESAIM Math. Modelling Num. Analysis
- Emmanuel Trélat is Editor of BCAM Springer Briefs
- Emmanuel Trélat is Associate editor of J. Optim. Theory Appl.
- Emmanuel Trélat is Associate editor of Math. Control Related fields

9.1.4. Invited Talks

- Ugo Boscain was invited speaker at the School and Workshop "Random Matrix Theory and Point Processes", Trieste, Italy, September.
- Barbara Gris was invited speaker at "Young researchers Imaging Seminars", IHP, Paris, February.
- Barbara Gris was invited speaker at GDR MAMOVI, Tours, February.
- Barbara Gris was invited speaker at "Information Geometry", Toulouse, October.
- Emmanuel Trélat was plenary speaker at "Equadiff 2019", Leiden, The Netherlands, July.
- Emmanuel Trélat was invited speaker at "Colloquium du Laboratoire de Mathématiques d'Avignon", Mars.
- Emmanuel Trélat was invited speaker at "DEA 2019", Krakow, Poland, September.
- Emmanuel Trélat was invited speaker at "Quantization in Symplectic Geometry", Cologne, Germany, July.
- Emmanuel Trélat was invited speaker at "Mathematical Models and Methods in Earth and Space Science", Rome, Italy, March.
- Emmanuel Trélat was invited speaker at "Sub-Riemannian Geometry and beyond", Jyväskylä, Finland, February.

9.1.5. Leadership within the Scientific Community

Emmanuel Trélat was director of the Fondation Sciences Mathématiques de Paris (FSMP) until June.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Ugo Boscain thought "Controllability of Closed Quantum Systems" to PhD students at the first QUSCO School, Saarbrucken, Germany.
- Ugo Boscain thought "Automatic Control" (with Mazyar Mirrahimi) at Ecole Polytechnique
- Ugo Boscain thought "MODAL of applied mathematics. Contrôle de modèles dynamiques" at Ecole Polytechnique
- Ugo Boscain thought "Control theory and applications in Quantum Mechanics" to PhD students at SISSA, Trieste, Italy
- Mario Sigalotti thought "Geometric Control Theory" to PhD students at "29th Jyväskylä Summer School", Jyväskylä, Finland
- Mario Sigalotti thought "Équations d'évolution, stabilité et contrôle" to M1 students at Sorbonne Université
- Emmanuel Trélat thought "Control in finite and infinite dimension" to M2 students at Sorbonne Université

9.2.2. Supervision

PhD: Nicolas Augier, “Contrôle adiabatique des systèmes quantiques", September 2019, supervisors: Ugo Boscain, Mario Sigalotti.

PhD: Amaury Hayat, “ Stabilisation de systèmes hyperboliques non-linéaires en dimension un d’espace", May 2019, supervisors: Jean-Michel Coron and Sébastien Boyaval

PhD: Mathieu Kohli, “On the notion of geodesic curvature in sub-Riemannian geometry", September 2019, supervisors: Davide Barilari, Ugo Boscain.

PhD: Jakub Orłowski, “Adaptive control of time-delay systems to counteract pathological brain oscillations", December 2019, supervisors: Antoine Chaillet, and Mario Sigalotti.

PhD: Shengquan Xiang, “Stabilisation rapide d’équations de Burgers et de Korteweg-de Vrie”, June 2019, supervisor: Jean-Michel Coron.

PhD: Christophe Zhang, “Contrôle et stabilisation internes de systèmes hyperboliques 1-D”, October 2019, supervisor: Jean-Michel Coron.

PhD in progress: Gontran Lance, started in September 2018, supervisors: Emmanuel Trélat and Enrique Zuazua.

PhD in progress: Cyril Letrouit, “Équation des ondes sous-riemanniennes", started in September 2019, supervisor Emmanuel Trélat.

PhD in progress: Emilio Molina, “Application of optimal control techniques to natural resources management”, started in September 2018, supervisors: Pierre Martinon, Héctor Ramírez, and Mario Sigalotti.

PhD in progress: Eugenio Pozzoli, “Adiabatic Control of Open Quantum Systems", started in September 2018, supervisors: Ugo Boscain and Mario Sigalotti.

PhD in progress: Rémi Robin, “Orbit spaces of Lie groups and applications to quantum control", started in September 2019, supervisors: Ugo Boscain and Mario Sigalotti.

9.2.3. Juries

- Ugo Boscain and Mario Sigalotti were supervisors and members of the jury of the PhD thesis of Nicolas Augier, Université Paris Saclay
- Ugo Boscain was co-supervisor and member of the jury of the PhD thesis of Mathieu Kohli, Université Paris Saclay
- Jean-Michel Coron was supervisor and member of the jury of the PhD thesis of Christophe Zhang, Sorbonne Université.
- Jean-Michel Coron was supervisor and member of the jury of the PhD thesis of Shengquan Xiang, Sorbonne Université.
- Mario Sigalotti was opponent of the PhD thesis of Eero Hakavuori, Jyväskylä University, Finland
- Mario Sigalotti was president and member of the jury of the PhD thesis of Gerardo Cardona, École des Mines, Paris.
- Emmanuel Trélat was referee and member of the jury of the HDR of Patrick Martinez, Université de Toulouse.
- Emmanuel Trélat was member of the jury of the HDR of Max Cerf, Sorbonne Université.
- Emmanuel Trélat was referee and member of the jury of the PhD thesis of Mathieu Granzotto, Université de Lorraine.
- Emmanuel Trélat was member of the jury of the PhD thesis of Christophe Zhang, Sorbonne Université.
- Emmanuel Trélat was president and member of the jury of the PhD thesis of Armand Koenig, Université de Nice.
- Emmanuel Trélat was member of the jury of the PhD thesis of Shengquan Xiang, Sorbonne Université.

- Emmanuel Trélat was member of the jury of the PhD thesis of Nicolas Hegoburu, Université de Bordeaux.

9.3. Popularization

9.3.1. Internal or external Inria responsibilities

Emmanuel Trélat is member of the Comité d'Honneur du Comité International des Jeux Mathématiques

9.3.2. Articles and contents

- The work [7] by Pierre Martinon has been popularized in the article "Athlétisme, affaire Dreyfus, itinéraire des éboueurs... Comment les maths irriguent notre monde" by Soline Roy, appeared in *Le Figaro*, 1/1/2019.
- Emmanuel Trélat has been interviewed, as director of the Fondation Sciences Mathématiques de Paris, in *Le Figaro*, 2/1/2019.

9.3.3. Interventions

- Emmanuel Trélat gave a cycle of lectures at "Sciences et Société", Nancy
- Emmanuel Trélat gave a lecture at Bibliothèque Nationale de France (BNF) (https://vimeo.com/313123529)
- Emmanuel Trélat gave a general public lecture at "Mois de l'optimisation", Limoges
- Emmanuel Trélat gave a general public lecture in Padua, Italy

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] N. AUGIER.*Contrôle adiabatique des systèmes quantiques*, Université Paris-Saclay (ComUE), September 2019

[2] A. HAYAT.*Stabilization of 1D nonlinear hyperbolic systems by boundary controls*, Sorbonne Université , UPMC, May 2019, https://tel.archives-ouvertes.fr/tel-02274457

[3] M. KOHLI.*On the notion of geodesic curvature in sub-Riemannian geometry*, Université Paris-Saclay, September 2019, https://tel.archives-ouvertes.fr/tel-02325081

[4] J. ORŁOWSKI.*Adaptive control of time-delay systems to counteract pathological brain oscillations*, Université Paris-Saclay (ComUE), December 2019

[5] S. XIANG.*Stabilisation rapide d'équations de Burgers et de Korteweg-de Vrie*, Sorbonne Université, June 2019

[6] C. ZHANG.*Internal control and stabilization of some 1-D hyperbolic systems*, Sorbonne Université, October 2019, https://tel.archives-ouvertes.fr/tel-02464011

Articles in International Peer-Reviewed Journal

[7] A. AFTALION, P. MARTINON.*Optimizing running a race on a curved track*, in "PLoS ONE", September 2019, vol. 14, n^o 9, 23, https://arxiv.org/abs/1811.12321 [*DOI :* 10.1371/JOURNAL.PONE.0221572], https://hal.inria.fr/hal-01936993

[8] D. BARILARI, Y. CHITOUR, F. JEAN, D. PRANDI, M. SIGALOTTI.*On the regularity of abnormal minimizers for rank 2 sub-Riemannian structures*, in "Journal de Mathématiques Pures et Appliquées", 2019, https://arxiv.org/abs/1804.00971 , forthcoming [*DOI :* 10.1016/J.MATPUR.2019.04.008], https://hal.archives-ouvertes.fr/hal-01757343

[9] G. BASTIN, J.-M. CORON, A. HAYAT, P. SHANG.*Boundary feedback stabilization of hydraulic jumps*, in "IFAC Journal of Systems and Control", 2019, https://hal.archives-ouvertes.fr/hal-02004457

[10] G. BASTIN, J.-M. CORON, A. HAYAT, P. SHANG.*Exponential boundary feedback stabilization of a shock steady state for the inviscid Burgers equation*, in "Mathematical Models and Methods in Applied Sciences", 2019, vol. 29, n^{o} 2, p. 271-316, https://hal.archives-ouvertes.fr/hal-01723361

[11] R. BONALLI, B. HÉRISSÉ, E. TRÉLAT.*Solving nonlinear optimal control problems with state and control delays by shooting methods combined with numerical continuation on the delays*, in "SIAM Journal on Control and Optimization", 2019, vol. 57, n^{o} 2, p. 1440–1466, https://arxiv.org/abs/1805.11990 [*DOI :* 10.1137/18M119121X], https://hal.archives-ouvertes.fr/hal-01802752

[12] C. CHEN, B. GRIS, O. ÖKTEM.*Joint Image Reconstruction and Motion Estimation for Spatiotemporal Imaging*, in "SIAM Journal on Imaging Sciences", October 2019, https://hal.archives-ouvertes.fr/hal-02386215

[13] Y. CHITOUR, G. MAZANTI, M. SIGALOTTI.*Approximate and exact controllability of linear difference equations*, in "Journal de l'École polytechnique — Mathématiques", 2020, vol. 7, p. 93–142, https://arxiv.org/abs/1708.06175 [*DOI :* 10.5802/JEP.112], https://hal.inria.fr/hal-01575576

[14] J.-M. CORON, F. MARBACH, F. SUEUR, P. ZHANG.*Controllability of the Navier-Stokes equation in a rectangle with a little help of a distributed phantom force*, in "Annals of PDE", 2019, vol. 5, n^{o} 17, https://arxiv.org/abs/1801.01860 , forthcoming [*DOI :* 10.1007/S40818-019-0073-4], https://hal.archives-ouvertes.fr/hal-01676663

[15] J.-M. CORON, H.-M. NGUYEN.*Optimal time for the controllability of linear hyperbolic systems in one dimensional space*, in "SIAM Journal on Control and Optimization", April 2019, vol. 57, n^{o} 2, p. 1127-1156 [*DOI :* 10.1137/18M1185600], https://hal.archives-ouvertes.fr/hal-01952134

[16] B. GRIS.*Incorporation of a deformation prior in image reconstruction*, in "Journal of Mathematical Imaging and Vision", June 2019, https://hal.archives-ouvertes.fr/hal-01810443

[17] E. HUMBERT, Y. PRIVAT, E. TRÉLAT.*Observability properties of the homogeneous wave equation on a closed manifold*, in "Communications in Partial Differential Equations", 2019, vol. 44, n^{o} 9, p. 749–772, https://arxiv.org/abs/1607.01535 [*DOI :* 10.1080/03605302.2019.1581799], https://hal.archives-ouvertes.fr/hal-01338016

[18] A. LE RHUN, F. BONNANS, G. DE NUNZIO, T. LEROY, P. MARTINON.*A stochastic data-based traffic model applied to vehicles energy consumption estimation*, in "IEEE Transactions on Intelligent Transportation Systems", 2019, forthcoming [*DOI :* 10.1109/TITS.2019.2923292], https://hal.inria.fr/hal-01774621

[19] A. OLIVIER, T. HABERKORN, E. TRÉLAT, E. BOURGEOIS, D.-A. HANDSCHUH.*Redundancy implies robustness for bang-bang strategies*, in "Optimal Control Applications and Methods", 2019, vol. 40, n^{o} 1, p. 85–104, https://arxiv.org/abs/1707.02053 , https://hal.archives-ouvertes.fr/hal-01557937

[20] J. ORLOWSKI, A. CHAILLET, M. SIGALOTTI. *Counterexample to a Lyapunov Condition for Uniform Asymptotic Partial Stability*, in "IEEE Control Systems Letters", April 2020, vol. 4, n^{o} 2, p. 397-401 [*DOI :* 10.1109/LCSYS.2019.2939717], https://hal.archives-ouvertes.fr/hal-02291650

[21] B. PICCOLI, N. POURADIER DUTEIL, E. TRÉLAT. *Sparse control of Hegselmann-Krause models: Black hole and declustering*, in "SIAM Journal on Control and Optimization", 2019, vol. 57, n^{o} 4, p. 2628–2659 [*DOI :* 10.1137/18M1168911], https://hal.archives-ouvertes.fr/hal-01699261

[22] C. POUCHOL, E. TRÉLAT, E. ZUAZUA. *Phase portrait control for 1D monostable and bistable reaction-diffusion equations*, in "Nonlinearity", 2019, vol. 32, n^{o} 3, p. 884–909, https://arxiv.org/abs/1805.10786 [*DOI :* 10.1088/1361-6544/AAF07E], https://hal.archives-ouvertes.fr/hal-01800382

[23] C. PRIEUR, E. TRÉLAT. *Feedback stabilization of a 1D linear reaction-diffusion equation with delay boundary control*, in "IEEE Transactions on Automatic Control", 2019, vol. 64, n^{o} 4, p. 1415-1425, https://arxiv.org/abs/1709.02735 [*DOI :* 10.1109/TAC.2018.2849560], https://hal.archives-ouvertes.fr/hal-01583199

[24] Y. PRIVAT, E. TRÉLAT, E. ZUAZUA. *Spectral shape optimization for the Neumann traces of the Dirichlet-Laplacian eigenfunctions*, in "Calculus of Variations and Partial Differential Equations", 2019, vol. 58, n^{o} 2, 64, https://arxiv.org/abs/1809.05316 [*DOI :* 10.1007/S00526-019-1522-3], https://hal.archives-ouvertes.fr/hal-01872896

[25] C. ROMMEL, F. BONNANS, P. MARTINON, B. GREGORUTTI. *Gaussian Mixture Penalty for Trajectory Optimization Problems*, in "Journal of Guidance, Control, and Dynamics", August 2019, vol. 42, n^{o} 8, p. 1857–1862 [*DOI :* 10.2514/1.G003996], https://hal.inria.fr/hal-01819749

[26] E. TRÉLAT, G. WANG, Y. XU. *Characterization by observability inequalities of controllability and stabilization properties*, in "Pure and Applied Analysis", 2020, vol. 2, n^{o} 1, p. 93–122, https://arxiv.org/abs/1811.01543 [*DOI :* 10.2140/PAA.2020.2.93], https://hal.archives-ouvertes.fr/hal-01911941

[27] S. XIANG. *Null controllability of a linearized Korteweg-de Vries equation by backstepping approach*, in "SIAM Journal on Control and Optimization", April 2019 [*DOI :* 10.1137/17M1115253], https://hal.archives-ouvertes.fr/hal-01468750

[28] O. ÖKTEM, B. GRIS, C. CHEN. *Image reconstruction through metamorphosis*, in "Inverse Problems", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-01773633

International Conferences with Proceedings

[29] N. AUGIER, U. BOSCAIN, M. SIGALOTTI. *On the compatibility between the adiabatic and the rotating wave approximations in quantum control*, in "CDC 2019 - 58th Conference on Decision and Control", Nice, France, December 2019, https://arxiv.org/abs/1909.02226 , https://hal.inria.fr/hal-02277852

Scientific Books (or Scientific Book chapters)

[30] A. AGRACHEV, D. BARILARI, U. BOSCAIN. *A Comprehensive Introduction to sub-Riemannian Geometry*, in "A Comprehensive Introduction to sub-Riemannian Geometry", 2019, Introduction de l'ouvrage, forthcoming, https://hal.archives-ouvertes.fr/hal-02019181

[31] U. BOSCAIN, M. SIGALOTTI.*Introduction to controllability of non-linear systems*, in "Contemporary Research in Elliptic PDEs and Related Topics", S. DIPIERRO (editor), Springer, 2019 [*DOI :* 10.1007/978-3-030-18921-1_4], https://hal.inria.fr/hal-02421207

[32] L. YOUNES, B. GRIS, A. TROUVÉ.*Sub-Riemannian Methods in Shape Analysis*, in "Handbook of Variational Methods for Nonlinear Geometric Data", Springer, 2020, forthcoming, https://hal.archives-ouvertes.fr/hal-02386227

Other Publications

[33] R. ADAMI, U. BOSCAIN, V. FRANCESCHI, D. PRANDI.*Point interactions for 3D sub-Laplacians*, February 2019, https://arxiv.org/abs/1902.05475 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02020844

[34] N. AUGIER, U. BOSCAIN, M. SIGALOTTI.*Semi-conical eigenvalue intersections and the ensemble controllability problem for quantum systems*, September 2019, https://arxiv.org/abs/1909.02271 - working paper or preprint, https://hal.inria.fr/hal-02277653

[35] I. BESCHASTNYI, U. BOSCAIN, M. SIGALOTTI.*An obstruction to small-time controllability of the bilinear Schrödinger equation*, December 2019, https://arxiv.org/abs/1911.12994 - working paper or preprint, https://hal.inria.fr/hal-02385244

[36] F. BOAROTTO, Y. CHITOUR, M. SIGALOTTI.*Fuller singularities for generic control-affine systems with an even number of controls*, September 2019, https://arxiv.org/abs/1909.01061 - working paper or preprint, https://hal.inria.fr/hal-02276960

[37] F. BOAROTTO, M. SIGALOTTI.*Dwell-time control sets and applications to the stability analysis of linear switched systems*, February 2019, https://arxiv.org/abs/1902.03757 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02012606

[38] R. BONALLI, B. HÉRISSÉ, E. TRÉLAT.*Optimal Control of Endo-Atmospheric Launch Vehicle Systems: Geometric and Computational Issues*, March 2019, https://arxiv.org/abs/1710.11501 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01626869

[39] U. BOSCAIN, R. W. NEEL.*Extensions of Brownian motion to a family of Grushin-type singularities*, December 2019, https://arxiv.org/abs/1910.02256 - 11 pages [*DOI :* 10.02256], https://hal.archives-ouvertes.fr/hal-02394958

[40] U. BOSCAIN, E. POZZOLI, M. SIGALOTTI.*Classical and quantum controllability of a rotating 3D symmetric molecule*, December 2019, https://arxiv.org/abs/1910.01924 - working paper or preprint [*DOI :* 10.01924], https://hal.inria.fr/hal-02421593

[41] Y. CHITOUR, N. GUGLIELMI, V. Y. PROTASOV, M. SIGALOTTI.*Switching systems with dwell time: computation of the maximal Lyapunov exponent*, December 2019, https://arxiv.org/abs/1912.10214 - working paper or preprint, https://hal.inria.fr/hal-02423619

[42] Y. CHITOUR, G. MAZANTI, M. SIGALOTTI.*On the gap between deterministic and probabilistic joint spectral radii for discrete-time linear systems*, November 2019, https://arxiv.org/abs/1812.08399 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01961003

[43] J.-M. CORON, L. GRÜNE, K. WORTHMANN.*Model Predictive Control, Cost Controllability, and Homogeneity* *, December 2019, https://arxiv.org/abs/1906.05112 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02151402

[44] J.-M. CORON, H.-M. NGUYEN.*Null-controllability of linear hyperbolic systems in one dimensional space*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02334761

[45] A. HAYAT.*Exponential stability of general 1-D quasilinear systems with source terms for the C 1 norm under boundary conditions*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01613139

[46] A. HAYAT.*On boundary stability of inhomogeneous 2 × 2 1-D hyperbolic systems for the C^1 norm.*, January 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01790104

[47] A. HAYAT.*PI controller for the general Saint-Venant equations*, January 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01827988

[48] A. HAYAT, P. SHANG.*Exponential stability of density-velocity systems with boundary conditions and source term for the H^2 norm*, July 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02190778

[49] E. HUMBERT, Y. PRIVAT, E. TRÉLAT.*Geometric and probabilistic results for the observability of the wave equation*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01652890

[50] M. KOHLI.*A metric interpretation of the geodesic curvature in the heisenberg group*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01916425

[51] G. LANCE, E. TRÉLAT, E. ZUAZUA.*Turnpike in optimal shape design*, December 2019, https://arxiv.org/abs/1912.02621 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02388972

[52] A. LE RHUN, F. BONNANS, G. DE NUNZIO, T. LEROY, P. MARTINON.*A bi-level energy management strategy for HEVs under probabilistic traffic conditions*, September 2019, working paper or preprint, https://hal.inria.fr/hal-02278359

[53] A. LE RHUN, F. BONNANS, G. DE NUNZIO, T. LEROY, P. MARTINON.*An Eco-routing algorithm for HEVs under traffic conditions*, November 2019, working paper or preprint, https://hal.inria.fr/hal-02356277

[54] C. LETROUIT.*From internal to pointwise control for the 1D heat equation and minimal control time*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02013076

[55] C. LETROUIT.*Infinite-time observability of the wave equation with time-varying observation domains under a geodesic recurrence condition*, April 2019, https://arxiv.org/abs/1904.11354 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02102135

[56] H. LHACHEMI, C. PRIEUR, E. TRÉLAT.*PI Regulation of a Reaction-Diffusion Equation with Delayed Boundary Control*, September 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02294321

[57] J. LOHÉAC, E. TRÉLAT, E. ZUAZUA.*Nonnegative control of finite-dimensional linear systems*, 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02335968

[58] M. SIGALOTTI.*Bounds on time-optimal concatenations of arcs for two-input driftless 3D systems*, November 2019, https://arxiv.org/abs/1911.10811 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02378196

[59] E. TRÉLAT, G. WANG, Y. XU.*Stabilization of infinite-dimensional linear control systems by POD reduced-order Riccati feedback*, June 2019, https://arxiv.org/abs/1906.10339 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02163854

[60] C. ZHANG.*Finite-time internal stabilization of a linear 1-D transport equation*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01980349

[61] C. ZHANG.*Internal rapid stabilization of a 1-D linear transport equation with a scalar feedback*, November 2019, https://arxiv.org/abs/1810.11214 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01905098

References in notes

[62] D. BARILARI, U. BOSCAIN, M. SIGALOTTI (editors). *Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. 1*, EMS Series of Lectures in Mathematics, European Mathematical Society (EMS), Zürich, 2016, vi+324, Lecture notes from the IHP Trimester held at the Institut Henri Poincaré, Paris and from the CIRM Summer School “Sub-Riemannian Manifolds: From Geodesics to Hypoelliptic Diffusion” held in Luminy, Fall 2014, https://doi.org/10.4171/163

[63] D. BARILARI, U. BOSCAIN, M. SIGALOTTI (editors). *Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. II*, EMS Series of Lectures in Mathematics, European Mathematical Society (EMS), Zürich, 2016, viii+299, Lecture notes from the IHP Trimester held at the Institut Henri Poincaré, Paris and from the CIRM Summer School “Sub-Riemannian Manifolds: From Geodesics to Hypoelliptic Diffusion” held in Luminy, Fall 2014, https://doi.org/10.4171/163

[64] R. ADAMI, U. BOSCAIN.*Controllability of the Schrödinger Equation via Intersection of Eigenvalues*, in "Proceedings of the 44th IEEE Conference on Decision and Control", 2005, p. 1080–1085

[65] A. A. AGRACHEV.*Some open problems*, in "Geometric control theory and sub-Riemannian geometry", Springer INdAM Ser., Springer, Cham, 2014, vol. 5, p. 1–13

[66] A. AGRACHEV, D. BARILARI, L. RIZZI.*Curvature: a variational approach*, in "Mem. Amer. Math. Soc.", 2018, vol. 256, n^{o} 1225, v+142

[67] A. A. AGRACHEV, Y. BARYSHNIKOV, D. LIBERZON.*On robust Lie-algebraic stability conditions for switched linear systems*, in "Systems Control Lett.", 2012, vol. 61, n^{o} 2, p. 347–353, https://doi.org/10.1016/j.sysconle.2011.11.016

[68] A. AGRACHEV, U. BOSCAIN, J.-P. GAUTHIER, F. ROSSI.*The intrinsic hypoelliptic Laplacian and its heat kernel on unimodular Lie groups*, in "J. Funct. Anal.", 2009, vol. 256, n^{o} 8, p. 2621–2655, https://doi.org/10.1016/j.jfa.2009.01.006

[69] A. AGRACHEV, P. W. Y. LEE.*Generalized Ricci curvature bounds for three dimensional contact subriemannian manifolds*, in "Math. Ann.", 2014, vol. 360, n^{o} 1-2, p. 209–253, https://doi.org/10.1007/s00208-014-1034-6

[70] A. A. AGRACHEV, Y. L. SACHKOV.*Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences, Springer-Verlag, Berlin, 2004, vol. 87, xiv+412, Control Theory and Optimization, II, https://doi.org/10.1007/978-3-662-06404-7

[71] L. AMBROSIO, N. GIGLI, G. SAVARÉ.*Metric measure spaces with Riemannian Ricci curvature bounded from below*, in "Duke Math. J.", 2014, vol. 163, n^o 7, p. 1405–1490, https://doi.org/10.1215/00127094-2681605

[72] L. AMBROSIO, P. TILLI.*Topics on analysis in metric spaces*, Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, 2004, vol. 25, viii+133

[73] M. BALDE, U. BOSCAIN, P. MASON.*A note on stability conditions for planar switched systems*, in "Internat. J. Control", 2009, vol. 82, n^o 10, p. 1882–1888, https://doi.org/10.1080/00207170902802992

[74] F. BAUDOIN, N. GAROFALO.*Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries*, in "J. Eur. Math. Soc. (JEMS)", 2017, vol. 19, n^o 1, p. 151–219, https://doi.org/10.4171/JEMS/663

[75] F. BAUDOIN, J. WANG.*Curvature dimension inequalities and subelliptic heat kernel gradient bounds on contact manifolds*, in "Potential Anal.", 2014, vol. 40, n^o 2, p. 163–193, https://doi.org/10.1007/s11118-013-9345-x

[76] T. BAYEN.*Analytical parameterization of rotors and proof of a Goldberg conjecture by optimal control theory*, in "SIAM J. Control Optim.", 2008, vol. 47, n^o 6, p. 3007–3036 [*DOI :* 10.1137/070705325]

[77] K. BEAUCHARD, J.-M. CORON.*Controllability of a quantum particle in a moving potential well*, in "J. Funct. Anal.", 2006, vol. 232, n^o 2, p. 328–389

[78] M. BENAÏM, S. LE BORGNE, F. MALRIEU, P.-A. ZITT.*Qualitative properties of certain piecewise deterministic Markov processes*, in "Ann. Inst. Henri Poincaré Probab. Stat.", 2015, vol. 51, n^o 3, p. 1040–1075, https://doi.org/10.1214/14-AIHP619

[79] B. BERRET, C. DARLOT, F. JEAN, T. POZZO, C. PAPAXANTHIS, J. P. GAUTHIER.*The inactivation principle: mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements*, in "PLoS Comput. Biol.", 2008, vol. 4, n^o 10, e1000194, 25, https://doi.org/10.1371/journal.pcbi.1000194

[80] F. BLANCHINI.*Nonquadratic Lyapunov functions for robust control*, in "Automatica J. IFAC", 1995, vol. 31, n^o 3, p. 451–461, http://dx.doi.org/10.1016/0005-1098(94)00133-4

[81] F. BLANCHINI, S. MIANI.*A new class of universal Lyapunov functions for the control of uncertain linear systems*, in "IEEE Trans. Automat. Control", 1999, vol. 44, n^o 3, p. 641–647, http://dx.doi.org/10.1109/9.751368

[82] A. BONFIGLIOLI, E. LANCONELLI, F. UGUZZONI.*Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007, xxvi+800

[83] M. BORN, V. FOCK.*Beweis des adiabatensatzes*, in "Zeitschrift für Physik A Hadrons and Nuclei", 1928, vol. 51, n^o 3–4, p. 165–180

[84] U. BOSCAIN.*Stability of planar switched systems: the linear single input case*, in "SIAM J. Control Optim.", 2002, vol. 41, n^{o} 1, p. 89–112 [*DOI :* 10.1137/S0363012900382837]

[85] U. BOSCAIN, M. CAPONIGRO, T. CHAMBRION, M. SIGALOTTI.*A weak spectral condition for the controllability of the bilinear Schrödinger equation with application to the control of a rotating planar molecule*, in "Comm. Math. Phys.", 2012, vol. 311, n^{o} 2, p. 423–455, https://doi.org/10.1007/s00220-012-1441-z

[86] U. BOSCAIN, M. CAPONIGRO, M. SIGALOTTI.*Multi-input Schrödinger equation: controllability, tracking, and application to the quantum angular momentum*, in "J. Differential Equations", 2014, vol. 256, n^{o} 11, p. 3524–3551, https://doi.org/10.1016/j.jde.2014.02.004

[87] U. BOSCAIN, G. CHARLOT, M. SIGALOTTI.*Stability of planar nonlinear switched systems*, in "Discrete Contin. Dyn. Syst.", 2006, vol. 15, n^{o} 2, p. 415–432, https://doi.org/10.3934/dcds.2006.15.415

[88] U. BOSCAIN, R. A. CHERTOVSKIH, J. P. GAUTHIER, A. O. REMIZOV.*Hypoelliptic diffusion and human vision: a semidiscrete new twist*, in "SIAM J. Imaging Sci.", 2014, vol. 7, n^{o} 2, p. 669–695 [*DOI :* 10.1137/130924731]

[89] U. BOSCAIN, F. CHITTARO, P. MASON, M. SIGALOTTI.*Adiabatic control of the Schroedinger equation via conical intersections of the eigenvalues*, in "IEEE Trans. Automat. Control", 2012, vol. 57, n^{o} 8, p. 1970–1983

[90] U. BOSCAIN, J. DUPLAIX, J.-P. GAUTHIER, F. ROSSI.*Anthropomorphic image reconstruction via hypoelliptic diffusion*, in "SIAM J. Control Optim.", 2012, vol. 50, n^{o} 3, p. 1309–1336 [*DOI :* 10.1137/11082405X]

[91] M. S. BRANICKY.*Multiple Lyapunov functions and other analysis tools for switched and hybrid systems*, in "IEEE Trans. Automat. Control", 1998, vol. 43, n^{o} 4, p. 475–482, Hybrid control systems, https://doi.org/10.1109/9.664150

[92] R. W. BROCKETT.*System theory on group manifolds and coset spaces*, in "SIAM J. Control", 1972, vol. 10, p. 265–284

[93] F. BULLO, A. D. LEWIS.*Geometric control of mechanical systems*, Texts in Applied Mathematics, Springer-Verlag, New York, 2005, vol. 49, xxiv+726, Modeling, analysis, and design for simple mechanical control systems [*DOI :* 10.1007/978-1-4899-7276-7]

[94] C. CARATHÉODORY.*Untersuchungen über die Grundlagen der Thermodynamik*, in "Math. Ann.", 1909, vol. 67, n^{o} 3, p. 355–386, https://doi.org/10.1007/BF01450409

[95] E. CARTAN.*Sur la represéntation géométrique des systèmes matériels non holonomes*, in "Proceedings of the International Congress of Mathematicians. Volume 4.", 1928, p. 253–261

[96] T. CHAMBRION, P. MASON, M. SIGALOTTI, U. BOSCAIN.*Controllability of the discrete-spectrum Schrödinger equation driven by an external field*, in "Ann. Inst. H. Poincaré Anal. Non Linéaire", 2009, vol. 26, n^{o} 1, p. 329–349, https://doi.org/10.1016/j.anihpc.2008.05.001

[97] G. CITTI, A. SARTI.*A cortical based model of perceptual completion in the roto-translation space*, in "J. Math. Imaging Vision", 2006, vol. 24, n^{o} 3, p. 307–326, http://dx.doi.org/10.1007/s10851-005-3630-2

[98] F. COLONIUS, G. MAZANTI.*Decay rates for stabilization of linear continuous-time systems with random switching*, in "Math. Control Relat. Fields", 2019

[99] J.-M. CORON.*Control and nonlinearity*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2007, vol. 136, xiv+426

[100] J.-M. CORON.*On the controllability of nonlinear partial differential equations*, in "Proceedings of the International Congress of Mathematicians. Volume I", Hindustan Book Agency, New Delhi, 2010, p. 238–264

[101] J.-M. CORON.*Global asymptotic stabilization for controllable systems without drift*, in "Math. Control Signals Systems", 1992, vol. 5, n^o 3, p. 295–312, https://doi.org/10.1007/BF01211563

[102] D. D'ALESSANDRO.*Introduction to quantum control and dynamics*, Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series, Chapman & Hall/CRC, Boca Raton, FL, 2008, xiv+343

[103] W. P. DAYAWANSA, C. F. MARTIN.*A converse Lyapunov theorem for a class of dynamical systems which undergo switching*, in "IEEE Trans. Automat. Control", 1999, vol. 44, n^o 4, p. 751–760, http://dx.doi.org/10.1109/9.754812

[104] R. DUITS, E. FRANKEN.*Left-invariant diffusions on the space of positions and orientations and their application to crossing-preserving smoothing of HARDI images*, in "Int. J. Comput. Vis.", 2011, vol. 92, n^o 3, p. 231–264, https://doi.org/10.1007/s11263-010-0332-z

[105] M. FLIESS, J. LÉVINE, P. MARTIN, P. ROUCHON.*Flatness and defect of non-linear systems: introductory theory and examples*, in "Internat. J. Control", 1995, vol. 61, n^o 6, p. 1327–1361, https://doi.org/10.1080/00207179508921959

[106] A. FRANCI, R. SEPULCHRE.*A three-scale model of spatio-temporal bursting*, in "SIAM J. Appl. Dyn. Syst.", 2016, vol. 15, n^o 4, p. 2143–2175 [*DOI :* 10.1137/15M1046101]

[107] S. J. GLASER, U. BOSCAIN, T. CALARCO, C. P. KOCH, W. KÖCKENBERGER, R. KOSLOFF, I. KUPROV, B. LUY, S. SCHIRMER, T. SCHULTE-HERBRÜGGEN, D. SUGNY, F. K. WILHELM.*Training Schrödinger's cat: quantum optimal control. Strategic report on current status, visions and goals for research in Europe*, in "European Physical Journal D", 2015, vol. 69, 279 [*DOI :* 10.1140/EPJD/E2015-60464-1]

[108] E. HAKAVUORI, E. LE DONNE.*Non-minimality of corners in subriemannian geometry*, in "Invent. Math.", 2016, p. 1–12 [*DOI :* 10.1007/S00222-016-0661-9]

[109] R. K. HLADKY, S. D. PAULS.*Minimal surfaces in the roto-translation group with applications to a neuro-biological image completion model*, in "J. Math. Imaging Vision", 2010, vol. 36, n^o 1, p. 1–27, https://doi.org/10.1007/s10851-009-0167-9

[110] D. HUBEL, T. WIESEL.*Brain and Visual Perception: The Story of a 25-Year Collaboration*, Oxford University Press, Oxford, 2004

[111] V. JURDJEVIC.*Geometric control theory*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1997, vol. 52, xviii+492

[112] V. JURDJEVIC, H. J. SUSSMANN.*Control systems on Lie groups*, in "J. Differential Equations", 1972, vol. 12, p. 313–329, https://doi.org/10.1016/0022-0396(72)90035-6

[113] M. KEYL, R. ZEIER, T. SCHULTE-HERBRUEGGEN.*Controlling Several Atoms in a Cavity*, in "New J. Phys.", 2014, vol. 16, 065010

[114] F. KÜSTERS, S. TRENN.*Switch observability for switched linear systems*, in "Automatica J. IFAC", 2018, vol. 87, p. 121–127, https://doi.org/10.1016/j.automatica.2017.09.024

[115] E. LE DONNE, G. P. LEONARDI, R. MONTI, D. VITTONE.*Extremal Curves in Nilpotent Lie Groups*, in "Geom. Funct. Anal.", jul 2013, vol. 23, n^{o} 4, p. 1371–1401 [*DOI :* 10.1007/s00039-013-0226-7], http://arxiv.org/abs/1207.3985

[116] Z. LEGHTAS, A. SARLETTE, P. ROUCHON.*Adiabatic passage and ensemble control of quantum systems*, in "Journal of Physics B", 2011, vol. 44, n^{o} 15

[117] C. LI, I. ZELENKO.*Jacobi equations and comparison theorems for corank 1 sub-Riemannian structures with symmetries*, in "J. Geom. Phys.", 2011, vol. 61, n^{o} 4, p. 781–807, https://doi.org/10.1016/j.geomphys.2010.12.009

[118] D. LIBERZON, J. P. HESPANHA, A. S. MORSE.*Stability of switched systems: a Lie-algebraic condition*, in "Systems Control Lett.", 1999, vol. 37, n^{o} 3, p. 117–122, https://doi.org/10.1016/S0167-6911(99)00012-2

[119] D. LIBERZON.*Switching in systems and control*, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, 2003, xiv+233, https://doi.org/10.1007/978-1-4612-0017-8

[120] D. LIBERZON.*Calculus of variations and optimal control theory*, Princeton University Press, Princeton, NJ, 2012, xviii+235, A concise introduction

[121] H. LIN, P. J. ANTSAKLIS.*Stability and stabilizability of switched linear systems: a survey of recent results*, in "IEEE Trans. Automat. Control", 2009, vol. 54, n^{o} 2, p. 308–322, http://dx.doi.org/10.1109/TAC.2008.2012009

[122] W. LIU.*Averaging theorems for highly oscillatory differential equations and iterated Lie brackets*, in "SIAM J. Control Optim.", 1997, vol. 35, n^{o} 6, p. 1989–2020 [*DOI :* 10.1137/S0363012994268667]

[123] J. LOTT, C. VILLANI.*Ricci curvature for metric-measure spaces via optimal transport*, in "Ann. of Math. (2)", 2009, vol. 169, n^{o} 3, p. 903–991, https://doi.org/10.4007/annals.2009.169.903

[124] P. MASON, U. BOSCAIN, Y. CHITOUR.*Common polynomial Lyapunov functions for linear switched systems*, in "SIAM J. Control Optim.", 2006, vol. 45, n^{o} 1, p. 226–245 (electronic) [*DOI :* 10.1137/040613147]

[125] P. MASON, M. SIGALOTTI.*Generic controllability properties for the bilinear Schrödinger equation*, in "Comm. Partial Differential Equations", 2010, vol. 35, n^{o} 4, p. 685–706, https://doi.org/10.1080/03605300903540919

[126] L. MASSOULIÉ.*Stability of distributed congestion control with heterogeneous feedback delays*, in "IEEE Trans. Automat. Control", 2002, vol. 47, n^{o} 6, p. 895–902, Special issue on systems and control methods for communication networks, https://doi.org/10.1109/TAC.2002.1008356

[127] M. I. MILLER, A. TROUVÉ, L. YOUNES.*Geodesic shooting for computational anatomy*, in "J. Math. Imaging Vision", 2006, vol. 24, n^{o} 2, p. 209–228, https://doi.org/10.1007/s10851-005-3624-0

[128] M. MIRRAHIMI.*Lyapunov control of a quantum particle in a decaying potential*, in "Ann. Inst. H. Poincaré Anal. Non Linéaire", 2009, vol. 26, n^{o} 5, p. 1743–1765, https://doi.org/10.1016/j.anihpc.2008.09.006

[129] A. P. MOLCHANOV, Y. S. PYATNITSKIY.*Lyapunov functions that specify necessary and sufficient conditions for absolute stability of nonlinear nonstationary control systems, I, II, III*, in "Automat. Remote Control", 1986, vol. 47, p. 344–354, 443–451, 620–630

[130] R. MONTGOMERY.*A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2002, vol. 91, xx+259

[131] R. MONTI.*The regularity problem for sub-Riemannian geodesics*, in "Geometric control theory and sub-Riemannian geometry", Springer INdAM Ser., Springer, Cham, 2014, vol. 5, p. 313–332, https://doi.org/10.1007/978-3-319-02132-4_18

[132] R. M. MURRAY, S. S. SASTRY.*Nonholonomic motion planning: steering using sinusoids*, in "IEEE Trans. Automat. Control", 1993, vol. 38, n^{o} 5, p. 700–716, https://doi.org/10.1109/9.277235

[133] G. NENCIU.*On the adiabatic theorem of quantum mechanics*, in "J. Phys. A", 1980, vol. 13, n^{o} 2, p. L15–L18, http://stacks.iop.org/0305-4470/13/L15

[134] V. NERSESYAN.*Growth of Sobolev norms and controllability of the Schrödinger equation*, in "Comm. Math. Phys.", 2009, vol. 290, n^{o} 1, p. 371–387

[135] D. PATINO, M. BÂJA, P. RIEDINGER, H. CORMERAIS, J. BUISSON, C. IUNG.*Alternative control methods for DC-DC converters: an application to a four-level three-cell DC-DC converter*, in "Internat. J. Robust Nonlinear Control", 2011, vol. 21, n^{o} 10, p. 1112–1133, https://doi.org/10.1002/rnc.1651

[136] J. PETITOT.*Neurogéomètrie de la vision. Modèles mathématiques et physiques des architectures fonctionnelles*, Les Éditions de l'École Polythechnique, 2008

[137] J. RUESS, J. LYGEROS.*Moment-based methods for parameter inference and experiment design for stochastic biochemical reaction networks*, in "ACM Trans. Model. Comput. Simul.", 2015, vol. 25, n^{o} 2, Art. 8, 25, https://doi.org/10.1145/2688906

[138] A. SARTI, G. CITTI, J. PETITOT.*The symplectic structure of the primary visual cortex*, in "Biol. Cybernet.", 2008, vol. 98, n^{o} 1, p. 33–48, http://dx.doi.org/10.1007/s00422-007-0194-9

[139] H. SCHÄTTLER, U. LEDZEWICZ.*Geometric optimal control*, Interdisciplinary Applied Mathematics, Springer, New York, 2012, vol. 38, xx+640, Theory, methods and examples, https://doi.org/10.1007/978-1-4614-3834-2

[140] H. SCHÄTTLER, U. LEDZEWICZ.*Optimal control for mathematical models of cancer therapies*, Interdisciplinary Applied Mathematics, Springer, New York, 2015, vol. 42, xix+496, An application of geometric methods, https://doi.org/10.1007/978-1-4939-2972-6

[141] R. SHORTEN, F. WIRTH, O. MASON, K. WULFF, C. KING.*Stability criteria for switched and hybrid systems*, in "SIAM Rev.", 2007, vol. 49, n^o 4, p. 545–592 [*DOI :* 10.1137/05063516X]

[142] S. SOLMAZ, R. SHORTEN, K. WULFF, F. Ó CAIRBRE.*A design methodology for switched discrete time linear systems with applications to automotive roll dynamics control*, in "Automatica J. IFAC", 2008, vol. 44, n^o 9, p. 2358–2363, https://doi.org/10.1016/j.automatica.2008.01.014

[143] E. D. SONTAG.*Input to state stability: basic concepts and results*, in "Nonlinear and optimal control theory", Lecture Notes in Math., Springer, Berlin, 2008, vol. 1932, p. 163–220, https://doi.org/10.1007/978-3-540-77653-6_3

[144] K.-T. STURM.*On the geometry of metric measure spaces. I*, in "Acta Math.", 2006, vol. 196, n^o 1, p. 65–131, https://doi.org/10.1007/s11511-006-0002-8

[145] K.-T. STURM.*On the geometry of metric measure spaces. II*, in "Acta Math.", 2006, vol. 196, n^o 1, p. 133–177, https://doi.org/10.1007/s11511-006-0003-7

[146] Z. SUN, S. S. GE, T. H. LEE.*Controllability and reachability criteria for switched linear systems*, in "Automatica J. IFAC", 2002, vol. 38, n^o 5, p. 775–786, https://doi.org/10.1016/S0005-1098(01)00267-9

[147] Z. SUN, S. S. GE.*Stability theory of switched dynamical systems*, Communications and Control Engineering Series, Springer, London, 2011, xx+253, https://doi.org/10.1007/978-0-85729-256-8

[148] H. J. SUSSMANN.*A regularity theorem for minimizers of real-analytic subriemannian metrics*, in "53rd IEEE Conference on Decision and Control", 2014, p. 4801-4806

[149] K. TAN, X. YANG.*Subriemannian geodesics of Carnot groups of step 3*, in "ESAIM Control Optim. Calc. Var.", 2013, vol. 19, n^o 1, p. 274–287, https://doi.org/10.1051/cocv/2012006

[150] S. TEUFEL.*Adiabatic perturbation theory in quantum dynamics*, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2003, vol. 1821, vi+236

[151] A. TROUVÉ, L. YOUNES.*Metamorphoses through Lie group action*, in "Found. Comput. Math.", 2005, vol. 5, n^o 2, p. 173–198, https://doi.org/10.1007/s10208-004-0128-z

[152] E. TRÉLAT.*Contrôle optimal*, Mathématiques Concrètes. [Concrete Mathematics], Vuibert, Paris, 2005, vi+246, Théorie & applications. [Theory and applications]

[153] M. TUCSNAK, G. WEISS.*Observation and control for operator semigroups*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 2009, xii+483, https://doi.org/10.1007/978-3-7643-8994-9

[154] G. TURINICI.*On the controllability of bilinear quantum systems*, in "Mathematical models and methods for ab initio Quantum Chemistry", M. DEFRANCESCHI, C. LE BRIS (editors), Lecture Notes in Chemistry, Springer, 2000, vol. 74

[155] M. VIANA.*Lectures on Lyapunov exponents*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2014, vol. 145, xiv+202, https://doi.org/10.1017/CBO9781139976602

[156] R. VINTER.*Optimal control*, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, 2000, xviii+507

[157] D. WISNIACKI, G. MURGIDA, P. TAMBORENEA.*Quantum control using diabatic and adiabatic transitions*, in "AIP Conference Proceedings", AIP, 2007, vol. 963, n^{o} 2, p. 840–842

[158] L. YATSENKO, S. GUÉRIN, H. JAUSLIN.*Topology of adiabatic passage*, in "Phys. Rev. A", 2002, vol. 65, 043407, 7

[159] A. VAN DER SCHAFT, H. SCHUMACHER.*An introduction to hybrid dynamical systems*, Lecture Notes in Control and Information Sciences, Springer-Verlag London, Ltd., London, 2000, vol. 251, xiv+174, https://doi.org/10.1007/BFb0109998

Project-Team CAMBIUM

Programming languages: type systems, concurrency, proofs of programs

IN PARTNERSHIP WITH:
Collège de France

RESEARCH CENTER
Paris

THEME
Proofs and Verification

Table of contents

Project-Team CAMBIUM

Creation of the Project-Team: 2019 August 01

Keywords:

Computer Science and Digital Science:

A1.1.1. - Multicore, Manycore
A1.1.3. - Memory models
A2.1. - Programming Languages
A2.1.1. - Semantics of programming languages
A2.1.3. - Object-oriented programming
A2.1.4. - Functional programming
A2.1.6. - Concurrent programming
A2.1.11. - Proof languages
A2.2. - Compilation
A2.2.1. - Static analysis
A2.2.2. - Memory models
A2.2.4. - Parallel architectures
A2.2.5. - Run-time systems
A2.4. - Formal method for verification, reliability, certification
A2.4.1. - Analysis
A2.4.3. - Proofs
A2.5.4. - Software Maintenance & Evolution
A7.1.2. - Parallel algorithms
A7.2. - Logic in Computer Science
A7.2.2. - Automated Theorem Proving
A7.2.3. - Interactive Theorem Proving

Other Research Topics and Application Domains:

B5.2.3. - Aviation
B6.1. - Software industry
B6.6. - Embedded systems
B9.5.1. - Computer science

1. Team, Visitors, External Collaborators

Research Scientists

François Pottier [Team leader, Inria, Senior Researcher, HDR]
Damien Doligez [Inria, Researcher]
Ioannis Filippidis [Inria, Starting Research Position]
Xavier Leroy [Collège de France, Professor]
Jean-Marie Madiot [Inria, Researcher]
Luc Maranget [Inria, Researcher]
Gabriel Radanne [Inria, Starting Research Position, from Oct 2019]
Didier Rémy [Inria, Senior Researcher, HDR]

Technical Staff

Florian Angeletti [Inria, Engineer, from Sep 2019]
Sébastien Hinderer [Inria, Engineer]

PhD Students

Frédéric Bour [Tarides, PhD Student, from Oct 2019, CIFRE]
Basile Clément [Inria, PhD Student, from Dec 2019]
Nathanaël Courant [École Normale Supérieure de Paris, PhD Student, from Sep 2019]
Paulo Emílio de Vilhena [Inria, PhD Student, from Sep 2019]
Armaël Guéneau [Inria, PhD Student]
Quentin Ladeveze [Inria, PhD Student, from Oct 2019]
Glen Mével [Inria, PhD Student]

Visiting Scientist

Jacques Garrigue [Nagoya University, from Sep 2019]

Administrative Assistant

Hélène Milome [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

The research conducted in the Cambium team aims at improving the safety, reliability and security of software through advances in programming languages and in formal program verification. Our work is centered on the design, formalization, and implementation of programming languages, with particular emphasis on type systems and type inference, formal program verification, shared-memory concurrency and weak memory models. We are equally interested in theoretical foundations and in applications to real-world problems. The OCaml programming language and the CompCert C compiler embody many of our research results.

2.2. Software reliability and reusability

Software nowadays plays a pervasive role in our environment: it runs not only on general-purpose computers, as found in homes, offices, and data centers, but also on mobile phones, credit cards, inside transportation systems, factories, and so on. Furthermore, whereas building a single isolated software system was once rightly considered a daunting task, today, tens of millions of developers throughout the world collaborate to develop software components that have complex interdependencies. Does this mean that the "software crisis" of the early 1970s, which Dijsktra described as follows, is over?

> *By now it is generally recognized that the design of any large sophisticated system is going to be a very difficult job, and whenever one meets people responsible for such undertakings, one finds them very much concerned about the reliability issue, and rightly so.* – Edsger W. Dijkstra

To some extent, the crisis is indeed over. In the past five decades, strong emphasis has been put on **modularity** and **reusability**. It is by now well-understood how to build reusable software components, thus avoiding repeated programming effort and reducing costs. The availability of hundreds of thousands of such components, hosted in collaborative repositories, has allowed the software industry to bloom in a manner that was unimaginable a few decades ago.

As pointed out by Dijkstra, however, the problem is not just to build software, but to ensure that it works. Today, the **reliability** of most software leaves a lot to be desired. Consumer-grade software, including desktop, Web, and mobile phone applications, often crashes or exhibits unexpected behavior. This results in loss of time, loss of data, and also can often be exploited for malicious purposes by attackers. Reliability includes **safety**—exhibiting appropriate behavior under normal usage conditions—and **security**—resisting abuse in the hands of an attacker.

Today, achieving very high levels of reliability is possible, albeit at a tremendous cost in time and money. In the aerospace industry, for instance, high reliability is obtained via meticulous development processes, extensive testing efforts, and external reviewing by independent certification authorities. There and elsewhere, **formal verification** is also used, instead of or in addition to the above methods. In the hardware industry, model-checking is used to verify microprocessor components. In the critical software industry, deductive program verification has been used to verify operating system kernels, file systems, compilers, and so on. Unfortunately, these methods are difficult to apply in industries that have strong cost and time-to-market constraints, such as the automotive industry, let alone the general software industry.

Today, thus, we arguably still are experiencing a "reliable-software crisis". Although we have become pretty good at producing and evolving software, we still have difficulty producing cheap reliable software.

How to resolve this crisis remains, to a large extent, an open question. Modularity and reusability seem needed now more than ever, not only in order to avoid repeated programming effort and reduce the likelihood of errors, but also and foremost to avoid repeated specification and verification effort. Still, apparently, the languages that we use to write software are not expressive enough, and the logics and tools that we use to verify software are not mature enough, for this crisis to be behind us.

2.3. Qualities of a programming language

A programming language is the medium through which an intent (software design) is expressed (program development), acted upon (program execution), and reasoned about (verification). It would be a mistake to argue that, with sufficient dedication, effort, time and cleverness, good software can be written in any programming language. Although this may be true in principle, in reality, the choice of an adequate programming language can be the deciding factor between software that works and software that does not, or even cannot be developed at all.

We believe, in particular, that it is crucial for a programming language to be **safe**, **expressive**, to encourage **modularity**, and to have a simple, well-defined **semantics**.

- **Safety**. The execution of a program must not ever be allowed to go wrong in an unpredictable way. Examples of behaviors that must be forbidden include reading or writing data outside of the memory area assigned by the operating system to the process and executing arbitrary data as if it were code. A programming language is safe if every safety violation is gracefully detected either at compile time or at runtime.
- **Expressiveness**. The programming language should allow programmers to think in terms of concise, high-level abstractions—including the concepts and entities of the application domain—as opposed to verbose, low-level representations or encodings of these concepts.
- **Modularity**. The programming language should make it easy to develop a software component in isolation, to describe how it is intended to be composed with other components, and to check at composition time that this intent is respected.
- **Semantics**. The programming language should come with a mathematical definition of the meaning of programs, as opposed to an informal, natural-language description. This definition should ideally be formal, that is, amenable to processing by a machine. A well-defined semantics is a prerequisite for proving that the language is safe (in the above sense) and for proving that a specific program is correct (via model-checking, deductive program verification, or other formal methods).

The safety of a programming language is usually achieved via a combination of design decisions, compile-time type-checking, and runtime checking. As an example design decision, memory deallocation, a dangerous operation, can be placed outside of the programmer's control. As an example of compile-time type-checking, attempting to use an integer as if it were a pointer can be considered a type error; a program that attempts to do this is then rejected by the compiler before it is executed. Finally, as an example of runtime checking, attempting to access an array outside of its bounds can be considered a runtime error: if a program attempts to do this, then its execution is aborted.

Type-checking can be viewed as an automated means of establishing certain correctness properties of programs. Thus, type-checking is a form of “lightweight formal methods” that provides weak guarantees but whose burden seems acceptable to most programmers. However, type-checking is more than just a program analysis that detects a class of programming errors at compile time. Indeed, types offer a language in which the interaction between one program component and the rest of the program can be formally described. Thus, they can be used to express a high-level description of the service provided by this component (i.e., its API), independently of its implementation. At the same time, they protect this component against misuse by other components. In short, “type structure is a syntactic discipline for enforcing levels of abstraction”. In other words, types offer basic support for expressiveness and modularity, as described above.

For this reason, types play a central role in programming language design. They have been and remain a fundamental research topic in our group. More generally, the design of new programming languages and new type systems and the proof of their safety has been and remains an important theme. The continued evolution of OCaml, as well as the design and formalization of Mezzo [2], are examples.

2.4. Design, implementation, and evolution of OCaml

Our group’s expertise in programming language design, formalization and implementation has traditionally been focused mainly on the programming language OCaml [22]. OCaml can be described as a high-level statically-typed general-purpose programming language. Its main features include first-class functions, algebraic data structures and pattern matching, automatic memory management, support for traditional imperative programming (mutable state, exceptions), and support for modularity and encapsulation (abstract types; modules and functors; objects and classes).

OCaml meets most of the key criteria that we have put forth above. Thanks to its static type discipline, which rejects unsafe programs, it is **safe**. Because its type system is equipped with powerful features, such as polymorphism, abstract types, and type inference, it is **expressive**, **modular**, and concise. Although OCaml as a whole does not have a **formal semantics**, many fragments of it have been formally studied in isolation. As a result, we believe that OCaml is a good language in which to develop complex software components and software systems and (possibly) to verify that they are correct.

OCaml has long served a dual role as a vehicle for our programming language research and as a mature real-world programming language. This remains true today, and we wish to preserve this dual role. On the research side, there are many directions in which the language could be extended. On the applied side, OCaml is used within academia (for research and for teaching) and in the industry. It is maintained by a community of active contributors, which extends beyond our team at Inria. It comes with a package manager, **opam**, a rich ecosystem of libraries, and a set of programming tools, including an IDE (Merlin), support for debugging and performance profiling, etc.

OCaml has been used to develop many complex systems, such as proof assistants (Coq, HOL Light), automated theorem provers (Alt-Ergo, Zenon), program verification tools (Why3), static analysis engines (Astrée, Frama-C, Infer, Flow), programming languages and compilers (SCADE, Reason, Hack), Web servers (Ocsigen), operating systems (MirageOS, Docker), financial systems (at companies such as Jane Street, LexiFi, Nomadic Labs), and so on.

2.5. Software verification

We have already mentioned the importance of formal verification to achieve the highest levels of software quality. One of our major contributions to this field has been the verification of programming tools, namely the CompCert optimizing compiler for the C language [8] and the Verasco abstract interpretation-based static analyzer [6]. Technically, this is deductive verification of purely functional programs, using the Coq proof assistant both as the prover and the programming language. Scientifically, CompCert and Verasco are milestones in the area of program proof, due to the complexity and realism of the code generation, optimization, and static analysis techniques that are verified. Practically, these formally-verified tools strengthen the guarantees that can be obtained by formal verification of critical software and reduce the need for other verification activities, attracting the interest of Airbus and other companies that develop critical embedded software.

CompCert is implemented almost entirely in Gallina, the purely functional programming language that lies at the heart of Coq. Extraction, a whole-program translation from Gallina to OCaml, allows Gallina programs to be compiled to native code and efficiently executed. Unfortunately, Gallina is a very restrictive language: it rules out all side effects, including nontermination, mutable state, exceptions, delimited control, nondeterminism, input/output, and concurrency. In comparison, most industrial programming languages, including OCaml, are vastly more expressive and convenient. Thus, there is a clear need for us to also be able to verify software components that are written in OCaml and exploit side effects.

To reason about the behavior of effectful programs, one typically uses a “program logic”, that is, a system of deduction rules that are tailor-made for this purpose, and can be built into a verification tool. Since the late 1960s, program logics for imperative programming languages with global mutable state have been in wide use. A key advance was made in the 2000s with the appearance of Separation Logic, which emphasizes local reasoning and thereby allows reasoning about a callee independently of its caller, about one heap fragment independently of the rest of the heap, about one thread independently of all other threads, and so on. Today, this field is extremely active: the development of powerful program logics for rich effectful programming languages, such as OCaml or Multicore OCaml, is a thriving and challenging research area.

Our team has expertise in this field. For several years, François Pottier has been investigating the theoretical foundations and applications of several features of modern Separation Logics, such as “hidden state” and “monotonic state”. Jean-Marie Madiot has contributed to the Verified Software Toolchain, which includes a version of Concurrent Separation Logic for a subset of C. Arthur Charguéraud [0] has developed CFML, an implementation of Separation Logic for a subset of OCaml. Armaël Guéneau has extended CFML with the ability to simultaneously verify the correctness and the time complexity of an OCaml component. Glen Mével and Paulo de Vilhena are currently investigating the use of Iris, a descendant of Concurrent Separation Logic, to carry out proofs of Multicore OCaml programs.

We envision several ways of using OCaml components that have been verified using a program logic. In the simplest scenario, some key OCaml components, such as the standard library, are verified, and are distributed for use in unverified applications. This increases the general trustworthiness of the OCaml system, but does not yield strong guarantees of correctness. In a second scenario, a fully verified application is built out of verified OCaml components, therefore it comes with an end-to-end correctness guarantee. In a third scenario, while some components are written and verified directly at the level of OCaml, others are first written and verified in Gallina, then translated down to verified OCaml components by an improved version of Coq’s extraction mechanism. In this scenario, it is possible to fully verify an application that combines effectful OCaml code and side-effect-free Gallina code. This scenario represents an improvement over the current state of the art. Today, CompCert includes several OCaml components, which cannot be verified in Coq. As a result, the data produced by these components must be validated by verified checkers.

2.6. Shared-memory concurrency

Concurrent shared-memory programming seems required in order to extract maximum performance out of the multicore general-purpose processors that have been in wide use for more than a decade. (GPUs and other special-purpose processors offer even greater raw computing power, but are not easily exploited in the symbolic computing applications that we are usually interested in.) Unfortunately, concurrent programming is notoriously more difficult than sequential programming. This can be attributed to a “state-space explosion problem”: the number of permitted program executions grows exponentially with the number of concurrent agents involved. Shared memory introduces an additional, less notorious, difficulty: on a modern multicore processor, execution does *not* follow the strong model where the instructions of one thread are interleaved with the instructions of other threads, and where reads and writes to memory instantaneously take effect. To properly understand and analyze a program, one must first formally define the semantics of the programming language, or of the device that is used to execute the program. The aspect of the semantics that governs the interaction of threads through memory is known as a **memory model**. Most modern memory models are **weak** in the sense that they offer fewer guarantees than the strong model sketched above.

[0] Formerly a PhD student in our team, today a researcher at Inria Nancy Grand-Est, team Camus.

Describing a memory model in precise mathematical language, in a manner that is at the same time faithful with respect to real-world machines and exploitable as a basis for reasoning about programs, is a challenging problem and a domain of active research, where thorough testing and verification are required.

Luc Maranget and Jean-Marie Madiot have acquired an expertise in the domain of weak memory models, including so-called axiomatic models and event-structure-based models. Moreover, Luc Maranget develops **diy-herd-litmus**, a unique software suite for defining, simulating and testing memory models. In short, **diy** generates so-called *litmus tests* from concise specifications; **herd** simulates litmus tests with respect to memory models expressed in the domain-specific language CAT; **litmus** executes litmus tests on real hardware. These tools have been instrumental in finding bugs in the deployed processors IBM Power5 and ARM Cortex-A9. Moreover, within industry, some models are now written in CAT, either for internal use, such as the AArch64 model by Will Deacon (ARM), or for publication, such as the RISC-V model by Luc Maranget and the HSA model by Jade Alglave and Luc Maranget.

For a long time, the OCaml language and runtime system have been restricted to sequential execution, that is, execution of a single computation thread on a single processor core. Yet, since 2014 approximately, the Multicore OCaml project at OCaml Labs (Cambridge, UK) is preparing a version of OCaml where multiple threads execute concurrently and communicate with each other via shared memory.

In principle, it seems desirable for Multicore OCaml to become the standard version of OCaml. Integrating Multicore OCaml into mainstream OCaml, however, is a major undertaking. The runtime system is deeply impacted: in particular, OCaml's current high-performance garbage collector must be replaced with an entirely new concurrent collector. The memory model and operational semantics of the language must be clearly defined. At the programming-language level, several major extensions are proposed, including effect handlers (a generalization of exception handlers, introducing a form of delimited control) and a new type-and-effect-discipline that statically detects and rejects unhandled effects.

3. Research Program

3.1. Research Directions

Our research proposal is organized along three main axes, namely **programming language design and implementation**, **concurrency**, and **program verification**. These three areas have strong connections. For instance, the definition and implementation of Multicore OCaml intersects the first two axes, whereas creating verification technology for Multicore OCaml programs intersects the last two.

In short, the "programming language design and implementation" axis includes:

- The search for richer type disciplines, in an effort to make our programming languages safer and more expressive. Two domains, namely modules and effects, appear of particular interest. In addition, we view type inference as an important cross-cutting concern.
- The continued evolution of OCaml. The major evolutions that we envision in the medium term are the integration of Multicore OCaml, the addition of modular implicits, and a redesign of the type-checker.
- Research on refactoring and program transformations.

The "concurrency" axis includes:

- Research on weak memory models, including axiomatic models, operational models, and event-structure models.
- Research on the Multicore OCaml memory model. This might include proving that the axiomatic and operational presentations of the model agree; testing the Multicore OCaml implementation to ensure that it conforms to the model; and extending the model with new features, should the need arise.

The “program verification” axis includes:

- The continued evolution of CompCert.
- Building new verified tools, such as verified compilers for domain-specific languages, verified components for the Coq type-checker, and so on.
- Verifying algorithms and data structures implemented in OCaml and in Multicore OCaml and enriching Separation Logic with new features, if needed, to better support this activity.
- The continued development of tools for TLA+.

4. Application Domains

4.1. Formal methods

We develop techniques and tools for the formal verification of critical software:

- program logics based on CFML and Iris for the deductive verification of software, including concurrency and algorithmic complexity aspects;
- verified development tools such as the CompCert verified C compiler, which extends properties established by formal verification at the source level all the way to the final executable code.

Some of these techniques have already been used in the nuclear industry (MTU Friedrichshafen uses CompCert to develop emergency diesel generators) and are under evaluation in the aerospace industry.

4.2. High-assurance software

Software that is not critical enough to undergo formal verification can still benefit greatly, in terms of reliability and security, from a functional, statically-typed programming language. The OCaml type system offers several advanced tools (generalized algebraic data types, abstract types, extensible variant and object types) to express many data structure invariants and safety properties and have them automatically enforced by the type-checker. This makes OCaml a popular language to develop high-assurance software, in particular in the financial industry. OCaml is the implementation language for the Tezos blockchain and cryptocurrency. It is also used for automated trading at Jane Street and for modeling and pricing of financial contracts at Bloomberg, Lexifi and Simcorp. OCaml is also widely used to implement code verification and generation tools at Facebook, Microsoft, CEA, Esterel Technologies, and many academic research groups, at Inria and elsewhere.

4.3. Design and test of microprocessors

The **diy** tool suite and the underlying methodology is in use at ARM Ltd to design and test the memory model of ARM architectures. In particular, the internal reference memory model of the ARMv8 (or AArch64) architecture has been written “in house” in Cat, our domain-specific language for specifying and simulating memory models. Moreover, our test generators and runtime infrastructure are used routinely at ARM to test various implementations of their architectures.

4.4. Teaching programming

Our work on the OCaml language family has an impact on the teaching of programming. OCaml is one of the programming languages selected by the French Ministry of Education for teaching Computer Science in classes préparatoires scientifiques. OCaml is also widely used for teaching advanced programming in engineering schools, colleges and universities in France, the USA, and Japan. The MOOC “Introduction to Functional Programming in OCaml”, developed at University Paris Diderot, is available on the France Université Numérique platform and comes with an extensive platform for self-training and automatic grading of exercises, developed in OCaml itself.

5. New Software and Platforms

5.1. 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/

5.2. Compcert

The CompCert formally-verified C compiler

KEYWORDS: Compilers - Formal methods - Deductive program verification - C - Coq

FUNCTIONAL DESCRIPTION: CompCert is a compiler for the C programming language. Its intended use is the compilation of life-critical and mission-critical software written in C and meeting high levels of assurance. It accepts most of the ISO C 99 language, with some exceptions and a few extensions. It produces machine code for the ARM, PowerPC, RISC-V, and x86 architectures. What sets CompCert C apart from any other production compiler, is that it is formally verified to be exempt from miscompilation issues, using machine-assisted mathematical proofs (the Coq proof assistant). In other words, the executable code it produces is proved to behave exactly as specified by the semantics of the source C program. This level of confidence in the correctness of the compilation process is unprecedented and contributes to meeting the highest levels of software assurance. In particular, using the CompCert C compiler is a natural complement to applying formal verification techniques (static analysis, program proof, model checking) at the source code level: the correctness proof of CompCert C guarantees that all safety properties verified on the source code automatically hold as well for the generated executable.

RELEASE FUNCTIONAL DESCRIPTION: Novelties include a formally-verified type checker for CompCert C, a more careful modeling of pointer comparisons against the null pointer, algorithmic improvements in the handling of deeply nested struct and union types, much better ABI compatibility for passing composite values, support for GCC-style extended inline asm, and more complete generation of DWARF debugging information (contributed by AbsInt).

- Participants: Xavier Leroy, Sandrine Blazy, Jacques-Henri Jourdan, Sylvie Boldo and Guillaume Melquiond
- Partner: AbsInt Angewandte Informatik GmbH
- Contact: Xavier Leroy
- URL: http://compcert.inria.fr/

5.3. Diy

Do It Yourself

KEYWORD: Parallelism

FUNCTIONAL DESCRIPTION: The diy suite provides a set of tools for testing shared memory models: the litmus tool for running tests on hardware, various generators for producing tests from concise specifications, and herd, a memory model simulator. Tests are small programs written in x86, Power or ARM assembler that can thus be generated from concise specification, run on hardware, or simulated on top of memory models. Test results can be handled and compared using additional tools.

- Participants: Jade Alglave and Luc Maranget
- Partner: University College London UK
- Contact: Luc Maranget
- URL: http://diy.inria.fr/

5.4. Menhir

KEYWORDS: Compilation - Context-free grammars - Parsing

FUNCTIONAL DESCRIPTION: Menhir is a LR(1) parser generator for the OCaml programming language. That is, Menhir compiles LR(1) grammar specifications down to OCaml code. Menhir was designed and implemented by François Pottier and Yann Régis-Gianas.

- Contact: François Pottier
- Publications: A Simple, Possibly Correct LR Parser for C11 - Reachability and Error Diagnosis in LR(1) Parsers

5.5. CFML

Interactive program verification using characteristic formulae

KEYWORDS: Coq - Software Verification - Deductive program verification - Separation Logic

FUNCTIONAL DESCRIPTION: The CFML tool supports the verification of OCaml programs through interactive Coq proofs. CFML proofs establish the full functional correctness of the code with respect to a specification. They may also be used to formally establish bounds on the asymptotic complexity of the code. The tool is made of two parts: on the one hand, a characteristic formula generator implemented as an OCaml program that parses OCaml code and produces Coq formulae, and, on the other hand, a Coq library that provides notations and tactics for manipulating characteristic formulae interactively in Coq.

- Participants: Arthur Charguéraud, Armaël Guéneau and François Pottier
- Contact: Arthur Charguéraud
- URL: http://www.chargueraud.org/softs/cfml/

5.6. TLAPS

TLA+ proof system

KEYWORD: Proof assistant

SCIENTIFIC DESCRIPTION: TLAPS is a platform for developing and mechanically verifying proofs about TLA+ specifications. The TLA+ proof language is hierarchical and explicit, allowing a user to decompose the overall proof into proof steps that can be checked independently. TLAPS consists of a proof manager that interprets the proof language and generates a collection of proof obligations that are sent to backend verifiers. The current backends include the tableau-based prover Zenon for first-order logic, Isabelle/TLA+, an encoding of TLA+ set theory as an object logic in the logical framework Isabelle, an SMT backend designed for use with any SMT-lib compatible solver, and an interface to a decision procedure for propositional temporal logic.

FUNCTIONAL DESCRIPTION: TLAPS is a proof assistant for the TLA+ specification language.

NEWS OF THE YEAR: Work in 2019 focused on providing support for reasoning about TLA+'s ENABLED and action composition constructs. We also prepared a minor release, fixing some issues and switching to Z3 as the default SMT back-end solver.

- Participants: Damien Doligez, Stephan Merz and Ioannis Filippidis
- Contact: Stephan Merz
- URL: https://tla.msr-inria.inria.fr/tlaps/content/Home.html

5.7. ZENON

KEYWORD: Automated theorem proving

FUNCTIONAL DESCRIPTION: Zenon is an automatic theorem prover based on the tableaux method. Given a first-order statement as input, it outputs a fully formal proof in the form of a Coq proof script. It has special rules for efficient handling of equality and arbitrary transitive relations. Although still in the prototype stage, it already gives satisfying results on standard automatic-proving benchmarks.

Zenon is designed to be easy to interface with front-end tools (for example integration in an interactive proof assistant), and also to be retargeted to output scripts for different frameworks (for example, Isabelle and Dedukti).

- Author: Damien Doligez
- Contact: Damien Doligez
- URL: http://zenon-prover.org/

5.8. hevea

hevea is a fast latex to html translator.

KEYWORDS: LaTeX - Web

FUNCTIONAL DESCRIPTION: HEVEA is a LATEX to html translator. The input language is a fairly complete subset of LATEX 2 (old LATEX style is also accepted) and the output language is html that is (hopefully) correct with respect to version 5. HEVEA understands LATEX macro definitions. Simple user style files are understood with little or no modifications. Furthermore, HEVEA customisation is done by writing LATEX code.

HEVEA is written in Objective Caml, as many lexers. It is quite fast and flexible. Using HEVEA it is possible to translate large documents such as manuals, books, etc. very quickly. All documents are translated as one single html file. Then, the output file can be cut into smaller files, using the companion program HACHA. HEVEA can also be instructed to output plain text or info files.

Information on HEVEA is available at http://hevea.inria.fr/.

- Author: Luc Maranget
- Contact: Luc Maranget
- URL: http://hevea.inria.fr/

6. New Results

6.1. Programming language design and implementation

6.1.1. The OCaml system

Participants: Damien Doligez, Armaël Guéneau, Xavier Leroy, Luc Maranget, David Allsop [Cambridge University], Florian Angeletti, Frédéric Bour [Facebook, until Sep 2019], Stephen Dolan [Cambridge University], Alain Frisch [Lexifi], Jacques Garrigue [Nagoya University], Sébastien Hinderer [SED], Nicolás Ojeda

Bär [Lexifi], Gabriel Radanne, Thomas Refis [Jane Street], Gabriel Scherer [Inria team Parsifal], Mark Shinwell [Jane Street], Leo White [Jane Street], Jeremy Yallop [Cambridge University].

This year, we released four versions of the OCaml system: versions 4.08.0, 4.08.1, 4.09.0, and 4.09.1. Versions 4.08.1 and 4.09.1 are minor releases that respectively fix 6 and 5 issues. Versions 4.08.0 and 4.09.0 are major releases that introduce new language features, improve performance and usability, and fix about 50 issues. The main novelties are:

- User-defined binding operators are now supported, with syntax similar to `let*`, `let+`, `and*`. These operators make it much easier to write OCaml code in monadic style or using applicative structures.
- The `open` construct now applies to arbitrary module expressions in structures and to applicative paths in signatures.
- A new notion of user-defined "alerts" generalizes the "deprecated" warning.
- New modules were added to the standard library: `Fun`, `Bool`, `Int`, `Option`, `Result`.
- Many floating-point functions were added, including fused multiply-add, as well as a new `Float.Array` submodule.
- Many error messages were improved, as well as error and warning reporting mechanisms.
- Pattern-matching constructs that correspond to affine functions are now optimized into arithmetic computations.

6.1.2. Evolution of the OCaml type system

Participants: Florian Angeletti, Jacques Garrigue, Thomas Refis [Jane Street], Didier Rémy, Gabriel Radanne, Gabriel Scherer [Inria team Parsifal], Leo White [Jane Street].

In addition to the work done on the above releases, efforts have been done to improve the type system and its implementation. Those include:

- Formalizing the typing of the pattern-matching of generalized algebraic data types (GADTs).
- Fixing some issues related to the incompleteness of the treatment of GADTs.
- Proposing extensions of the type system to reduce this incompleteness in concrete cases, by refining the information on abstract types.
- Exploring practical ways to obtain more polymorphism for functions whose soundness does not rely on the value restriction.
- Improving the readability of the type-checker code.
- Making the module layer of the type-checker more incremental, in order to improve efficiency and to facilitate integration with documentation tools.

6.1.3. Refactoring with ornaments in ML

Participants: Didier Rémy, Thomas Williams [Google Paris].

Thomas Williams, Lucas Baudin, and Didier Rémy have been working on refactoring and other transformations of ML programs based on mixed ornamentation and disornamentation. Ornaments have been introduced as a way of describing changes in data type definitions that can reorganize or add pieces of data. After a new data structure has been described as an ornament of an older one, the functions that operate on the bare structure can be partially or sometimes totally lifted into functions that operate on the ornamented structure.

This year, Williams and Rémy improved the formalization of the lifting framework. In particular, they introduced an intermediate language, in which nonexpansive expressions can be marked on source terms and traced during reduction. This allows to treat the nonexpansive part of expansive expressions as nonexpansive and use equational reasoning on nonexpansive parts of terms that appear in types. This approach significantly simplifies the metatheory of ornaments. This calculus could also have some interest in itself, beyond ornaments, to study languages with side effects.

6.1.4. A better treatment of type abbreviations during type inference

Participants: Didier Rémy, Carine Morel.

During her M2 internship under the supervision of Didier Rémy, Carine Morel revisited the treatment of type abbreviations in type inference for ML-like type systems, using a modern approach based on typing constraints [24]. Instead of expanding type abbreviations prior to unification, both the original abbreviated view and all expanded views are kept during unification, so as to avoid unnecessary expansions and use the least-expanded view whenever possible in the result of unification.

6.2. Software specification and verification

6.2.1. The CompCert formally-verified compiler

Participants: Xavier Leroy, Jacques-Henri Jourdan [CNRS], Michael Schmidt [AbsInt GmbH], Bernhard Schommer [AbsInt GmbH].

In the context of our work on compiler verification, since 2005, we have been developing and formally verifying a moderately-optimizing compiler for a large subset of the C programming language, generating assembly code for the ARM, PowerPC, RISC-V and x86 architectures [8]. This compiler comprises a back-end part, which translates the Cminor intermediate language to PowerPC assembly and which is reusable for source languages other than C [7], and a front-end, which translates the CompCert C subset of C to Cminor. The compiler is mostly written within the specification language of the Coq proof assistant, from which Coq's extraction facility generates executable OCaml code. The compiler comes with a 100000-line machine-checked Coq proof of semantic preservation establishing that the generated assembly code executes exactly as prescribed by the semantics of the source C program.

This year, we added a new optimization to CompCert: "if-conversion", that is, the replacement of conditional statements and expressions by conditional move operations and similar branchless instruction sequences. As a consequence, fewer conditional branch instructions are generated. This replacement usually improves worst-case execution time (WCET), because mispredicted conditional branches tremendously increase execution time. This replacement is also interesting for cryptographic code and other programs that manipulate secret data: conditional branches over secret data take time that depends on the data, leaking some information, while conditional move instructions are constant-time and do not leak. The new if-conversion optimization plays a role in the ongoing work of Inria team Celtique on compilation that preserves constant-time properties. Its proof of semantic preservation is nontrivial and prompted the development of a new kind of simulation diagram.

Other recent improvements to the CompCert C compiler include:

- a new code generator targeting the AArch64 instruction set, that is, the 64-bit mode of the ARMv8 architecture;
- the ability to specify the semantics of certain built-in functions, making them amenable to optimizations such as constant propagation and common subexpression elimination;
- improvements to the verified C parser generated by Menhir, including fewer run-time checks, faster validation, and the removal of all axioms from the proof.

We released two versions of CompCert incorporating these improvements: version 3.5 in February 2019 and version 3.6 in September 2019.

6.2.2. Time credits and time receipts in Iris

Participants: Glen Mével, François Pottier, Jacques-Henri Jourdan [CNRS].

From March to August 2018, Glen Mével did an M2 internship at Gallium, where he was co-advised by Jacques-Henri Jourdan (CNRS) and François Pottier. Glen extended the program logic Iris with time credits and time receipts.

Time credits are a well-understood concept, and have been used in several papers already by Armaël Guéneau, Arthur Charguéraud, and François Pottier. However, because Iris is implemented and proved sound inside Coq, extending Iris with time credits requires a nontrivial proof, which Glen carried out, based on a program transformation which inserts "tick" instructions into the code. As an application of time credits, Glen verified inside Iris the correctness of Okasaki's notion of "debits", which allows reasoning about the time complexity of programs that use thunks.

Time receipts are a new concept, which allows proving that certain undesirable events, such as integer overflows, cannot occur until a very long time has elapsed. Glen extended Iris with time receipts and proved the soundness of this extension. As an application of time credits and receipts together, Jacques-Henri Jourdan updated Charguéraud and Pottier's earlier verification of the Union-Find data structure [12] and proved that integer ranks cannot realistically overflow, even if they are stored using only $\log W$ bits, where W is the number of bits in a machine word.

This work carried out in 2018 has been published at ESOP 2019 [16].

6.2.3. A program logic for Multicore Ocaml

Participants: Glen Mével, François Pottier, Jacques-Henri Jourdan [CNRS].

Glen Mével, who is co-advised by Jacques-Henri Jourdan and François Pottier, has been working on designing a mechanized program logic for Multicore OCaml.

One of the key challenges is to enable deductive reasoning under a weak memory model. In such a model, the behaviors of a program are no longer described by a naive interleaving semantics. Thus, the operational semantics that describes a weak memory model often feels unnatural to the programmer, and is difficult to reason about.

This year, Glen designed and implemented a proof system on top of Iris, a modular separation logic framework whose implementation and soundness proof are both expressed in Coq. This system allows mechanized program verification for a fragment of the Multicore OCaml language. It provides a certain degree of abstraction over the low-level operational semantics, in the hope of simplifying reasoning. This abstraction includes an abstract concept of "local view" of the shared memory; views are exchanged between threads via atomic locations.

A few simple concurrent data structures have been proven correct using the system. They include several variants of locks and mutual exclusion algorithms.

Glen presented preliminary results at the Iris Workshop in October 2019.

6.2.4. Verifying a generic local solver in Iris

Participants: Paulo Emílio de Vilhena, Jacques-Henri Jourdan [CNRS], François Pottier.

From March to August 2019, Paulo Emílio de Vilhena did an M2 internship in our team, where he was advised by François Pottier, with precious help from Jacques-Henri Jourdan (CNRS).

Paulo verified a short but particularly subtle piece of code, namely a "local generic solver", that is, an on-demand, incremental, memoizing least fixed point computation algorithm. This algorithm is a slightly simplified version of `Fix`[0], an OCaml library published by François Pottier in 2009.

The specification of this algorithm is simple: the solver computes the optimal least fixed point of a system of monotone equations. Although the solver relies on mutable internal state for memoization and for "spying", a form of dynamic dependency discovery, no side effects are mentioned in the specification. The challenge is precisely to formally justify why it is permitted to hide these side effects from the user.

The verification is carried out in Iris, a modern breed of concurrent separation logic. Iris is embedded in Coq, so the proof is machine-checked. The proof makes crucial use of prophecy variables, a novel feature of Iris. Auxiliary contributions include a restricted infinitary conjunction rule for Iris and a specification and proof of Longley's "modulus" function, an archetypical example of spying.

[0] https://gitlab.inria.fr/fpottier/fix

This paper [13] has been accepted for presentation at the conference POPL 2020, which will take place in New Orleans in January 2020.

6.2.5. Formal reasoning about asymptotic complexity

Participants: Armaël Guéneau, Arthur Charguéraud [Inria team Camus], François Pottier, Jacques-Henri Jourdan [CNRS].

For several years, Armaël Guéneau, Arthur Charguéraud, François Pottier have been investigating the use of Separation Logic, extended with Time Credits, as an approach to the formal verification of the time complexity of OCaml programs. In 2018 and 2019, in collaboration with Jacques-Henri Jourdan, Armaël has worked on a more ambitious case study, namely a state-of-the-art incremental cycle detection algorithm, whose amortized complexity analysis is nontrivial. Armaël has proposed an improved and simplified algorithm and has carried out a machine-checked proof of its complexity. Furthermore, the verified algorithm has been released and is now used in production inside the Dune build system for OCaml. A paper has been published and presented at the International Conference on Interactive Theorem Proving (ITP 2019) [15]. A more detailed version of these results appears in Armaël Guéneau's dissertation [11], which was defended on December 16, 2019.

6.2.6. TLA+

Participants: Damien Doligez, Leslie Lamport [Microsoft Research], Ioannis Filippidis, Stephan Merz [Inria team VeriDis].

Damien Doligez is the head of the "Tools for Proofs" team in the Microsoft-Inria Joint Centre. The aim of this project is to extend the TLA+ language with a formal language for hierarchical proofs, formalizing Lamport's ideas [25], and to build tools for writing TLA+ specifications and mechanically checking the proofs.

We have made a bug-fix release of TLAPS (version 1.4.4). In parallel, we are working on adding features for dealing with temporal properties, that is, fairness and liveness. We have implemented support for the ENABLED operator and the action composition operator in TLA+ proofs. This support is still experimental, but we hope to release a new version of TLAPS next year with these features.

6.3. Shared-memory concurrency

6.3.1. Instruction fetch in the ARMv8 architecture

Participants: Luc Maranget, Peter Sewell [University of Cambridge], Ben Simmer [University of Cambridge].

Modern multi-core and multi-processor computers do not follow the intuitive "sequential consistency" model that would define a concurrent execution as the interleaving of the executions of its constituent threads and that would command instantaneous writes to the shared memory. This situation is due both to in-core optimisations such as speculative and out-of-order execution of instructions, and to the presence of sophisticated (and cooperating) caching devices between processors and memory. Luc Maranget is taking part in an international research effort to define the semantics of the computers of the multi-core era, and more generally of shared-memory parallel devices or languages, with a clear initial focus on devices.

Luc Maranget participates in project REMS, for *Rigorous Engineering for Mainstream Systems*, an EPSRC project led by Peter Sewell. This year Luc Maranget took part in a research effort that resulted in a paper entitled *ARMv8-A system semantics: instruction fetch in relaxed architectures*. This paper has been accepted for presentation at ESOP 2020. This paper introduces a robust model of instruction fetch and cache maintenance, a central aspect of a processor system's semantics, for ARMv8-A. Luc Maranget specifically extended the **litmus** and **diy** test generators so as to account for self-modifying code. He also performed part of the experiments that support the instruction fetch model.

6.3.2. An ARMv8 mixed-size memory model

Participants: Luc Maranget, Jade Alglave [ARM Ltd & University College London].

Jade Alglave and Luc Maranget have completed their work on a mixed-size version of the ARMv8 memory model. This model builds on the `aarch64.cat` model authored by Will Deacon (ARM Ltd). The model is now ready, and a paper has been written. They hope to work around certain intellectual property restrictions and to submit this paper for publication next year.

6.3.3. Work on diy

Participants: Luc Maranget, Jade Alglave [ARM Ltd & University College London], Antoine Hacquard.

The **diy** suite (for "Do It Yourself") provides a set of tools for testing shared memory models: the **litmus** tool for running tests on hardware, various generators for producing tests from concise specifications, and **herd**, a memory model simulator. Tests are small programs written in x86, Power, ARM, generic (LISA) assembler, or a subset of the C language that can thus be generated from concise specifications, run on hardware, or simulated on top of memory models. Test results can be handled and compared using additional tools. On distinctive feature of our system is Cat, a domain-specific language for memory models.

This year, new synchronisation primitives and instructions were added to various models. Some sizable developments occurred that facilitate the integration of mixed-size models into **herd**: a default definition of the same-instruction relation, which allows using mixed-size models on all tests; an automatic adjustment of the machine's elementary granularity, which facilitates massive testing; and the addition of equivalence classes and relations on them as basic values, which extends the expressiveness of Cat to some abstract mixed-size models.

During a 3-month internship, Antoine Hacquard (an EPITA second-year student) extended the complete tool suite to handle a new target, namely X86_64. The addition of this new target significantly enhances the **diy** tool suite, as X86_64 is a very popular architecture. Moreover, Antoine Hacquard implemented all memory access instructions for all sizes (from byte to quadword), which enabled us to design a mixed-size TSO model for this very popular architecture.

6.3.4. Unifying axiomatic and operational weak memory models

Participants: Quentin Ladeveze, Jean-Marie Madiot, Jade Alglave [ARM Ltd & University College London], Simon Castellan [Imperial College London].

Modern multi-processors optimize the running speed of programs using a variety of techniques, including caching, instruction reordering, and branch speculation. While those techniques are perfectly invisible to sequential programs, such is not the case for concurrent programs that execute several threads and share memory: threads do not share at every point in time a single consistent view of memory. A *weak memory model* offers only weak consistency guarantees when reasoning about the permitted behaviors of a program. Until now, there have been two kinds of such models, based on different mathematical foundations: axiomatic models and operational models.

Axiomatic models explicitly represent the dependencies between the program and memory actions. These models are convenient for causal reasoning about programs. They are also well-suited to the simulation and testing of *hardware* microprocessors.

Operational models represent program states directly, thus can be used to reason on programs: program logics become applicable, and the reasoning behind nondeterministic behavior is much clearer. This makes them preferable for reasoning about *software*.

Jean-Marie Madiot has been collaborating with weak memory model expert Jade Alglave and concurrent game semantics researcher Simon Castellan in order to unify these styles, in a way that attempts to combine the best of both approaches. The first results are a formalisation of TSO-style architectures using partial-order techniques similar to the ones used in game semantics, and a proof of a stronger-than-state-of-art "data-race freedom" theorem: well-synchronised programs can assume a strong memory model.

Since October 2019, Luc Maranget and Jean-Marie Madiot are advising a PhD candidate, Quentin Ladeveze. His goal is to further generalize and formalize weak memory models. This involves reasoning about linearizations of interdependent acyclic relations.

This is a first step towards tractable verification of concurrent programs, combining software verification using concurrent program logics, in the top layer, and hardware testing using weak memory models, in the bottom layer. Our hope is to leave no unverified gap between software and hardware, even (and especially) in the presence of concurrency.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. The Caml Consortium

Participant: Damien Doligez.

The Caml Consortium, is a formal structure where industrial and academic users of OCaml can support the development of the language and associated tools, express their specific needs, and contribute to the long-term stability of OCaml. Membership fees are used to fund specific developments targeted towards industrial users. Members of the Consortium automatically benefit from very liberal licensing conditions on the OCaml system, allowing for instance the OCaml compiler to be embedded within proprietary applications.

Damien Doligez chairs the Caml Consortium.

The Consortium currently has 9 member companies:

- Aesthetic Integration
- Citrix
- Docker
- Esterel Technologies
- Facebook
- Jane Street
- LexiFi
- Microsoft
- SimCorp

The Caml Consortium is being gradually phased out. In the future, we would like to replace it entirely with the OCaml Software Foundation, discussed below.

7.2. Bilateral Grants with Industry

7.2.1. The OCaml Software Foundation

Participants: Damien Doligez, Xavier Leroy.

The OCaml Software Foundation (OCSF),[0] established in 2018 under the umbrella of the Inria Foundation, aims to promote, protect, and advance the OCaml programming language and its ecosystem, and to support and facilitate the growth of a diverse and international community of OCaml users.

Damien Doligez and Xavier Leroy serve as advisors on the foundation's Executive Committee.

We receive substantial basic funding from the OCaml Software Foundation in order to support research activity related to OCaml.

[0] http://ocaml-sf.org/

7.2.2. Funding from Nomadic Labs

Nomadic Labs, a Paris-based company, has implemented the Tezos blockchain and cryptocurrency entirely in OCaml. This year, Nomadic Labs and Inria have signed a framework agreement (“contrat-cadre”) that allows Nomadic Labs to fund multiple research efforts carried out by Inria groups. Within this framework, we have received three 3-year grants:

- “Évolution d’OCaml”. This grant is intended to fund a number of improvements to OCaml, including the addition of new features and a possible re-design of the OCaml type-checker. This grant has allowed us to fund Jacques Garrigue’s visit (10 months) and to hire Gabriel Radanne on a Starting Research Position (3 years).
- “Maintenance d’OCaml”. This grant is intended to fund the day-to-day maintenance of OCaml as well as the considerable work involved in managing the release cycle. This grant has allowed us to hire Florian Angeletti as an engineer for 3 years.
- “Multicore OCaml”. This grant is intended to encourage research work on Multicore OCaml within our team. This grant has allowed us to fund Glen Mével’s PhD thesis (3 years).

7.2.3. Funding from the Microsoft-Inria joint lab

Funding from the Microsoft-Inria joint lab has allowed us to hire Ioannis Filippidis on a Starting Research Position (until March 2020) to work on the TLAPS system.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR projects

8.1.1.1. Vocal

Participants: Armaël Guéneau, Xavier Leroy, François Pottier.

The “Vocal” project (2015–2020) aims at developing the first mechanically verified library of efficient general-purpose data structures and algorithms. It is funded by *Agence Nationale de la Recherche* under its “appel à projets générique 2015”.

A first release of the library has been published in December 2018. It contains a small number of verified data structures, including resizable vectors, hash tables, priority queues, and Union-Find.

In 2019, progress was made on the definition of Gospel, a standard language for annotating OCaml programs with logical specifications, which could be understood and processed by several verification tools, including Why3 and CFML.

8.2. International Research Visitors

8.2.1. Visits of International Scientists

Jacques Garrigue (Nagoya University) is staying with our team in Paris from September 2019 to June 2020. He has long been one of the key designers and implementors of the OCaml type system. We are collaborating on the design of new language features and on a possible re-design of the type-checker implementation.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Selection

9.1.1.1. Chair of Conference Program Committees

François Pottier was the program chair of the International Conference on Functional Programming (ICFP 2019) which took place in Berlin, Germany in August 2019.

9.1.1.2. Member of the Conference Program Committees

Xavier Leroy was on the program committee of PERR 2019, the 3rd Workshop on Program Equivalence and Relational Reasoning, part of the ETAPS 2019 joint conferences.

Xavier Leroy was on the program committee of FOSSACS 2020, the 23rd International Conference on Foundations of Software Science and Computation Structures, part of the ETAPS 2020 joint conferences.

Didier Rémy was a member of the program committee for FLOPS 2020, the 15th International Symposium on Functional and Logic Programming.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

Xavier Leroy is area editor for Journal of the ACM, in charge of the Programming Languages area. He is a member of the editorial board of Journal of Automated Reasoning.

François Pottier is a member of the ICFP steering committee and a member of the editorial boards of the Journal of Functional Programming and the Proceedings of the ACM on Programming Languages.

Until September 2019, Didier Rémy was a member of the ML Family workshop steering committee.

9.1.3. Scientific Expertise

Didier Rémy co-authored the Inria white book on Cybersecurity [17], [18] and a companion document of *Recommendations for Inria's Management*.

9.1.4. Research Administration

Damien Doligez chairs the Caml Consortium.

François Pottier is a member of Inria Paris' *Commission de Développement Technologique* and the president of Inria Paris' *Comité de Suivi Doctoral*.

Didier Rémy is Inria's delegate in the MPRI's pedagogical and management board of the *Master Parisien de Recherche en Informatique* (MPRI).

Didier Rémy set up the Inria-Nomadic Labs partnership and is currently the co-chair of its steering committee.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master (M2): "Proofs of Programs", Jean-Marie Madiot, 18 HETD, MPRI, Université Paris Diderot, France.

Master (M2): "Programming shared memory multicore machines", Luc Maranget, 18 HETD, MPRI, Université Paris Diderot, France. Starting December 2019, Luc Maranget is in charge of this course.

Master (M2): "Functional programming and type systems", François Pottier, 18 HETD, MPRI, Université Paris Diderot, France.

Master (M2): "Functional programming and type systems", Didier Rémy, 18 HETD, MPRI, Université Paris Diderot, France. Didier Rémy is in charge of this course.

Licence (L3): Jean-Marie Madiot, "Introduction à l'informatique", 40 HETD, École Polytechnique, France.

Open lectures: Xavier Leroy, *Sémantiques mécanisées: quand la machine raisonne sur ses langages*, 19 HETD, Collège de France, France.

Summer school: Xavier Leroy, *Proving the correctness of a compiler*, 6 HETD, 2019 EUtypes summer school on Types for Programming and Verification, North Macedonia.

9.2.2. Supervision

PhD in progress: Frédéric Bour, “An interactive, modular proof environment for OCaml”, Université Paris Diderot, since October 2019, advised by François Pottier and Thomas Gazagnaire (Taridès).

PhD in progress: Basile Clément, “Domain-specific language and machine learning compiler for the automatic synthesis of high-performance numerical libraries”, École Normale Supérieure, since September 2018, advised by Xavier Leroy since October 2019.

PhD in progress: Nathanaël Courant, “Towards an efficient, formally-verified proof checker for Coq”, Université Paris Diderot, since September 2019, advised by Xavier Leroy.

PhD: Armaël Guéneau, “Mechanized Verification of the Correctness and Asymptotic Complexity of Programs”, Université Paris Diderot, defended on December 16, 2019 [11], advised by Arthur Charguéraud and François Pottier.

PhD in progress: Quentin Ladeveze, “Generic conditions for DRF-SC in axiomatic memory models”, Université Paris Diderot, since October 2019, advised by Luc Maranget and Jean-Marie Madiot.

PhD in progress: Glen Mével, “Towards a system for proving the correctness of concurrent Multicore OCaml programs”, Université Paris Diderot, since November 2018, advised by Jacques-Henri Jourdan and François Pottier.

PhD in progress: Thomas Williams, “Putting Ornaments into practice”, Université Paris Diderot, since September 2014, advised by Didier Rémy.

9.2.3. Juries

Xavier Leroy participated in the hiring committee for a professor position at the UFR d’Informatique of U. Paris Diderot.

Xavier Leroy chaired the jury for the Habilitation defense of Yann Régis-Gianas (U. Paris Diderot, November 2019).

Xavier Leroy was external reviewer for the PhD of Andrea Condoluci (U. Bologna, to be defended in 2020).

Jean-Marie Madiot served as an examiner at the computer science oral examination for the “second concours” of École Normale Supérieure de Lyon.

Didier Rémy participated in the hiring committee for an Inria Centrale-Supélec Chair in cybersecurity.

9.3. Popularization

9.3.1. Interventions

Xavier Leroy was on the committee of the 2019 Sephora Berrebi scholarship for women in mathematics and computer science. At the award ceremony (Paris, France, February 2019), he gave a popular science talk on Grace Hopper and the birth of the compiler.

Xavier Leroy gave popularization talks on deductive software verification at the Aerospace Lab Conferences of ONERA (Palaiseau, France, June 2019) and at the Summer 2019 BOB conference (Berlin, Germany, August 2019).

10. Bibliography

Major publications by the team in recent years

[1] J. Alglave, L. Maranget, M. Tautschnig. *Herding cats: modelling, simulation, testing, and data-mining for weak memory*, in "ACM Transactions on Programming Languages and Systems", 2014, vol. 36, n^o 2, http://dx.doi.org/10.1145/2627752

[2] T. BALABONSKI, F. POTTIER, J. PROTZENKO.*The design and formalization of Mezzo, a permission-based programming language*, in "ACM Transactions on Programming Languages and Systems", 2016, vol. 38, n^o 4, p. 14:1–14:94, http://doi.acm.org/10.1145/2837022

[3] K. CHAUDHURI, D. DOLIGEZ, L. LAMPORT, S. MERZ.*Verifying Safety Properties With the TLA+ Proof System*, in "Automated Reasoning, 5th International Joint Conference, IJCAR 2010", Lecture Notes in Computer Science, Springer, 2010, vol. 6173, p. 142–148, http://dx.doi.org/10.1007/978-3-642-14203-1_12

[4] J. CRETIN, D. RÉMY.*System F with Coercion Constraints*, in "CSL-LICS 2014: Computer Science Logic / Logic In Computer Science", ACM, 2014, http://dx.doi.org/10.1145/2603088.2603128

[5] A. GUÉNEAU, J.-H. JOURDAN, A. CHARGUÉRAUD, F. POTTIER.*Formal Proof and Analysis of an Incremental Cycle Detection Algorithm*, in "Interactive Theorem Proving", J. HARRISON, J. O'LEARY, A. TOLMACH (editors), Leibniz International Proceedings in Informatics, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, September 2019, vol. 141, https://hal.inria.fr/hal-02167236

[6] J.-H. JOURDAN, V. LAPORTE, S. BLAZY, X. LEROY, D. PICHARDIE.*A Formally-Verified C Static Analyzer*, in "POPL'15: 42nd ACM Symposium on Principles of Programming Languages", ACM Press, January 2015, p. 247-259, http://dx.doi.org/10.1145/2676726.2676966

[7] X. LEROY.*A formally verified compiler back-end*, in "Journal of Automated Reasoning", 2009, vol. 43, n^o 4, p. 363–446, http://dx.doi.org/10.1007/s10817-009-9155-4

[8] X. LEROY.*Formal verification of a realistic compiler*, in "Communications of the ACM", 2009, vol. 52, n^o 7, p. 107–115, http://doi.acm.org/10.1145/1538788.1538814

[9] G. MÉVEL, J.-H. JOURDAN, F. POTTIER.*Time Credits and Time Receipts in Iris*, in "European Symposium on Programming", Lecture Notes in Computer Science, Springer, April 2019, vol. 11423, p. 3-29 [*DOI :* 10.1007/978-3-030-17184-1_1], https://hal.archives-ouvertes.fr/hal-02183311

[10] N. POUILLARD, F. POTTIER.*A unified treatment of syntax with binders*, in "Journal of Functional Programming", 2012, vol. 22, n^o 4–5, p. 614–704, http://dx.doi.org/10.1017/S0956796812000251

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] A. GUÉNEAU.*Mechanized Verification of the Correctness and Asymptotic Complexity of Programs*, Université de Paris, December 2019, https://hal.inria.fr/tel-02437532

Articles in International Peer-Reviewed Journal

[12] 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", March 2019, vol. 62, n^o 3, p. 331–365 [*DOI :* 10.1007/s10817-017-9431-7], https://hal.inria.fr/hal-01652785

[13] P. E. DE VILHENA, F. POTTIER, J.-H. JOURDAN.*Spy Game: Verifying a Local Generic Solver in Iris*, in "Proceedings of the ACM on Programming Languages", January 2020, n^o 4 [*DOI :* 10.1145/3371101], https://hal.archives-ouvertes.fr/hal-02351562

Invited Conferences

[14] X. LEROY.*In Search of Software Perfection : An introduction to deductive software verification*, in "BOB Summer 2019 Konferenz", Berlin, Germany, August 2019, https://hal.inria.fr/hal-02392114

International Conferences with Proceedings

[15] A. GUÉNEAU, J.-H. JOURDAN, A. CHARGUÉRAUD, F. POTTIER.*Formal Proof and Analysis of an Incremental Cycle Detection Algorithm : (extended version)*, in "Interactive Theorem Proving", Portland, United States, J. HARRISON, J. O'LEARY, A. TOLMACH (editors), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, September 2019, n^o 141, https://hal.inria.fr/hal-02167236

[16] G. MÉVEL, J.-H. JOURDAN, F. POTTIER.*Time Credits and Time Receipts in Iris*, in "European Symposium on Programming", Prague, Czech Republic, Lecture Notes in Computer Science, Springer, April 2019, vol. 11423, p. 3-29 [*DOI :* 10.1007/978-3-030-17184-1_1], https://hal.archives-ouvertes.fr/hal-02183311

Scientific Books (or Scientific Book chapters)

[17] S. KREMER, L. MÉ, D. RÉMY, V. ROCA.*Cybersecurity : Current challenges and Inria's research directions*, Inria white book, Inria, January 2019, n^o 3, 172, https://hal.inria.fr/hal-01993308

[18] S. KREMER, L. MÉ, D. RÉMY, V. ROCA.*Cybersécurité : Défis actuels et axes de recherche à l'Inria*, Inria white book, Inria, May 2019, n^o 3, 18, https://hal.inria.fr/hal-02414281

[19] X. LEROY.*Le logiciel, entre l'esprit et la matière*, Leçons inaugurales du Collège de France, Fayard, April 2019, vol. 284, https://hal.inria.fr/hal-02370113

[20] X. LEROY.*Le logiciel, entre l'esprit et la matière : Leçon inaugurale prononcée au Collège de France le jeudi 15 novembre 2018*, OpenEdition Books, December 2019, https://hal.inria.fr/hal-02405754

[21] X. LEROY.*Software, between mind and matter*, Inaugural lecture at Collège de France, Collège de France, November 2019, https://hal.inria.fr/hal-02392159

Research Reports

[22] X. LEROY, D. DOLIGEZ, A. FRISCH, J. GARRIGUE, D. RÉMY, J. VOUILLON.*The OCaml system release 4.09: Documentation and user's manual*, Inria, September 2019, p. 1-789, https://hal.inria.fr/hal-00930213

[23] X. LEROY.*The CompCert C verified compiler: Documentation and user's manual : Version 3.6*, Inria, September 2019, p. 1-78, https://hal.inria.fr/hal-01091802

Other Publications

[24] C. MOREL.*Type inference and modular elaboration with constraints for ML extended with type abbreviations*, Université Paris Diderot-Paris 7, September 2019, https://hal.inria.fr/hal-02361707

References in notes

[25] L. LAMPORT.*How to write a 21st century proof*, in "Journal of Fixed Point Theory and Applications", 2012, vol. 11, p. 43–63, http://dx.doi.org/10.1007/s11784-012-0071-6

Project-Team CASCADE

Construction and Analysis of Systems for Confidentiality and Authenticity of Data and Entities

IN COLLABORATION WITH: Département d'Informatique de l'Ecole Normale Supérieure

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Paris

RESEARCH CENTER
Paris

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team CASCADE

Creation of the Project-Team: 2008 July 01

Keywords:

Computer Science and Digital Science:

A4. - Security and privacy
A4.3. - Cryptography
A4.3.1. - Public key cryptography
A4.3.3. - Cryptographic protocols
A4.8. - Privacy-enhancing technologies
A7. - Theory of computation
A8.5. - Number theory

Other Research Topics and Application Domains:

B6.4. - Internet of things
B9.5.1. - Computer science
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

David Pointcheval [Team leader, CNRS, Senior Researcher, HDR]
Michel Ferreira Abdalla [CNRS, Researcher, HDR]
Georg Fuchsbauer [Inria, Researcher]
Brice Minaud [Inria, Researcher]
Phong-Quang Nguyen [Inria, Senior Researcher, HDR]
Hoeteck Wee [CNRS, Researcher, HDR]

PhD Students

Balthazar Bauer [Inria]
Jérémy Chotard [CNRS]
Baptiste Cottier [Worldline, from Feb 2019]
Romain Gay [Ecole Normale Supérieure Paris, until Mar 2019]
Lenaïck Gouriou [Leanear, from Nov 2019]
Chloé Hébant [CNRS]
Louiza Khati [ANSSI, until Jun 2019]
Michele Orrù [CNRS]
Antoine Plouviez [Inria]
Mélissa Rossi [Thales & ANSSI]
Théo Ryffel [Inria, from Feb 2019]

Post-Doctoral Fellows

Ehsan Ebrahimi [Ecole Normale Supérieure Paris, from Feb 2019]
Pooya Farshim [CNRS, until Sep 2019]
Junqing Gong [CNRS]
Azam Soleimanian [Ecole Normale Supérieure Paris, from Feb 2019]

2. Overall Objectives

2.1. Presentation

Cryptographic algorithms are the equivalent of locks, seals, security stamps and identification documents over the Internet. They are essential to protect our online bank transactions, credit cards, medical and personal information, and to support e-commerce and e-government. They come in different flavors. Encryption algorithms are necessary to protect sensitive information such as medical data, financial information and Personal Identification Numbers (PINs) from prying eyes. Digital signature algorithms (in combination with hash functions) and MAC algorithms replace hand-written signatures in electronic transactions. Identification protocols allow to securely verify the identity of a remote party. As a whole, cryptology is a research area with a high strategic impact in industry, for individuals, and for society as a whole. The research activity of project-team CASCADE addresses the following topics, which cover most of the areas that are currently active in the international cryptographic community, with a focus on public-key algorithms:

1. Implementation of cryptographic algorithms, and applied cryptography;
2. Algorithm and protocol design, and provable security;
3. Theoretical and practical attacks.

2.2. Design of Provably Secure Primitives and Protocols

Since the beginning of public-key cryptography, with the seminal Diffie-Hellman paper, many suitable algorithmic problems for cryptography have been proposed and many cryptographic schemes have been designed, together with more or less heuristic proofs of their security relative to the intractability of the underlying problems. However, many of those schemes have thereafter been broken. The simple fact that a cryptographic algorithm withstood cryptanalytic attacks for several years has often been considered as a kind of validation procedure, but schemes may take a long time before being broken. An example is the Chor-Rivest cryptosystem, based on the knapsack problem, which took more than 10 years to be totally broken by Serge Vaudenay, whereas before this attack it was believed to be strongly secure. As a consequence, the lack of attacks at some time should never be considered as a full security validation of the proposal.

A completely different paradigm is provided by the concept of “provable” security. A significant line of research has tried to provide proofs in the framework of computational complexity theory (a.k.a. “reductionist” security proofs): the proofs provide reductions from a well-studied problem (factoring, RSA or the discrete logarithm) to an attack against a cryptographic protocol.

At the beginning, researchers just tried to define the security notions required by actual cryptographic schemes, and then to design protocols which could achieve these notions. The techniques were directly derived from complexity theory, providing polynomial reductions. However, their aim was essentially theoretical. They were indeed trying to minimize the required assumptions on the primitives (one-way functions or permutations, possibly trapdoor, etc), without considering practicality. Therefore, they just needed to design a scheme with polynomial-time algorithms, and to exhibit polynomial reductions from the basic mathematical assumption on the hardness of the underlying problem to an attack of the security notion, in an asymptotic way. However, such a result has no practical impact on actual security. Indeed, even with a polynomial reduction, one may be able to break the cryptographic protocol within a few hours, whereas the reduction just leads to an algorithm against the underlying problem which requires many years. Therefore, those reductions only prove the security when very huge (and thus maybe unpractical) parameters are in use, under the assumption that no polynomial-time algorithm exists to solve the underlying problem. For many years, more efficient reductions have been expected, under the denomination of either “exact security" or “concrete security”, which provide more practical security results, with concrete efficiency properties.

Unfortunately, in many cases, even just provable security is at the cost of an important loss in terms of efficiency for the cryptographic protocol. Thus, some models have been proposed, trying to deal with the security of efficient schemes: some concrete objects are identified with ideal (or black-box) ones. For example, it is by now usual to identify hash functions with ideal random functions, in the so-called "random-oracle model". Similarly, block ciphers are identified with families of truly random permutations in the "ideal cipher model". Another kind of idealization has also been introduced in cryptography, the black-box group, where the group operation, in any algebraic group, is defined by a black-box: a new element necessarily comes from the addition (or the subtraction) of two already known elements. It is by now called the "generic group model", extended to the bilinear and multi-linear setting. Some works even require several ideal models together to provide some new validations.

But still, such idealization cannot be instantiated in practice, and so one prefers provable security without such idealized assumptions, under new and possibly stronger computational assumptions. As a consequence, a cryptographer has to deal with the following four important steps, which are **all** main goals of ours:

computational assumptions, which are the foundation of the security. We thus need to have a strong evidence that the computational problems are reasonably hard to solve.

security model, which makes precise the security notions one wants to achieve, as well as the means the adversary may be given. We contribute to this point, in several ways:

- by providing security models for many primitives and protocols;
- by enhancing some classical security models;
- by considering new means for the adversary, such as side-channel information.

design of new schemes/protocols, or more efficient ones, with additional features, etc.

security proof, which consists in exhibiting a reduction.

3. Research Program

3.1. Quantum-Safe Cryptography

The security of almost all public-key cryptographic protocols in use today relies on the presumed hardness of problems from number theory such as factoring and computing discrete logarithms. This is problematic because these problems have very similar underlying structure, and its unforeseen exploit can render all currently used public-key cryptography insecure. This structure was in fact exploited by Shor to construct efficient quantum algorithms that break all hardness assumptions from number theory that are currently in use. And so naturally, an important area of research is to build provably secure protocols based on mathematical problems that are unrelated to factoring and discrete log. One of the most promising directions in this line of research is using lattice problems as a source of computational hardness, which also offer features that other alternative public-key cryptosystems (such as MQ-based, code-based or hash-based schemes) cannot provide.

3.2. Advanced Encryption

Fully Homomorphic Encryption (FHE) has become a very active research area since 2009, when IBM announced the discovery of a FHE scheme by Craig Gentry. FHE allows to perform any computation on encrypted data, yielding the result encrypted under the same key. This enables outsourcing computation in the Cloud, on encrypted data, so the Cloud provider does not learn any information. However, FHE does not allow to share the result.

Functional encryption is another recent tool that allows an authority to deliver functional decryption keys, for any function f of his choice, so that when applied to the encryption of a message m, the functional decryption key yields $f(m)$. Since m can be a large vector, f can be an aggregation or statistical function: on encrypted data, one can get the result $f(m)$ in clear.

While this functionality has initially been defined in theory, our team has been very active in designing concrete instantiations for practical purposes.

3.3. Security amidst Concurrency on the Internet

Cryptographic protocols that are secure when executed in isolation can become completely insecure when multiple such instances are executed concurrently (as is unavoidable on the Internet) or when used as a part of a larger protocol. For instance, a man-in-the-middle attacker participating in two simultaneous executions of a cryptographic protocol might use messages from one of the executions in order to compromise the security of the second – Lowe's attack on the Needham-Schroeder authentication protocol and Bleichenbacher's attack on SSL work this way. Our research addresses security amidst concurrent executions in secure computation and key exchange protocols.

Secure computation allows several mutually distrustful parties to collaboratively compute a public function of their inputs, while providing the same security guarantees as if a trusted party had performed the computation. Potential applications for secure computation include anonymous voting, privacy-preserving auctions and data-mining. Our recent contributions on this topic include

1. new protocols for secure computation in a model where each party interacts only once, with a single centralized server; this model captures communication patterns that arise in many practical settings, such as that of Internet users on a website, and
2. efficient constructions of universally composable commitments and oblivious transfer protocols, which are the main building blocks for general secure computation.

In key exchange protocols, we are actively involved in designing new password-authenticated key exchange protocols, as well as the analysis of the widely-used SSL/TLS protocols.

3.4. Electronic Currencies and the Blockchain

Electronic cash (e-cash) was first proposed in the 1980s but has never been deployed on a large scale. Other means of digital payments are instead largely replacing physical cash, but they do not respect the citizens' right to privacy, which includes their right of anonymous payments of moderate sums. Recently, so-called decentralized currencies, such as Bitcoin, have become a third type of payments in addition to physical cash, and card and other (non-anonymous) electronic payments. The continuous growth of popularity and usage of this new kind of currencies, also called "cryptocurrencies", have triggered a renewed interest in cryptographic e-cash.

On the one hand, our group investigates "centralized" e-cash, in keeping with the current economic model that has money be issued by (central) banks (while cryptocurrencies use money distribution as an incentive for participation in the system, on which its stability hinges). Of particular interest among centralized e-cash schemes is transferable e-cash, which allows users to transfer coins between each other without interacting with a third party (or the blockchain). Existing efficient e-cash schemes are not transferable, as they require coins to be deposited at the bank after having been used in a payment. Our goal is to propose efficient transferable e-cash schemes.

Another direction concerns (decentralized) cryptocurrencies, whose adoption has grown tremendously over the last few years. While in Bitcoin all transactions are publicly posted on the so-called "blockchain", other cryptocurrencies such as *Zcash* respect user privacy, whose security guarantees we have analyzed. Apart from privacy, two pressing challenges for cryptocurrencies, and blockchains in general, are sustainability and scalability. Regarding the former, we are addressing the electricity waste caused by the concept of "proof of work" used by all major cryptocurrencies by proposing alternatives; for the latter, we are working on proposals that avoid the need for all data having to be stored on the blockchain forever.

Blockchains have meanwhile found many other applications apart from electronic money. Together with Microsoft Research, our group investigates decentralized means of authentication that uses cryptography to guarantee privacy.

4. Application Domains

4.1. Privacy for the Cloud

Many companies have already started the migration to the Cloud and many individuals share their personal informations on social networks. While some of the data are public information, many of them are personal and even quite sensitive. Unfortunately, the current access mode is purely right-based: the provider first authenticates the client, and grants him access, or not, according to his rights in the access-control list. Therefore, the provider itself not only has total access to the data, but also knows which data are accessed, by whom, and how: privacy, which includes secrecy of data (confidentiality), identities (anonymity), and requests (obliviousness), should be enforced. Moreover, while high availability can easily be controlled, and thus any defect can immediately be detected, failures in privacy protection can remain hidden for a long time. The industry of the Cloud introduces a new implicit trust requirement: nobody has any idea at all of where and how his data are stored and manipulated, but everybody should blindly trust the providers. The providers will definitely do their best, but this is not enough. Privacy-compliant procedures cannot be left to the responsibility of the provider: however strong the trustfulness of the provider may be, any system or human vulnerability can be exploited against privacy. This presents too huge a threat to tolerate. *The distribution of the data and the secrecy of the actions must be given back to the users. It requires promoting privacy as a global security notion.*

In order to protect the data, one needs to encrypt it. Unfortunately, traditional encryption systems are inadequate for most applications involving big, complex data. Recall that in traditional public key encryption, a party encrypts data to a single known user, which lacks the expressiveness needed for more advanced data sharing. In enterprise settings, a party will want to share data with groups of users based on their credentials. Similarly, individuals want to selectively grant access to their personal data on social networks as well as documents and spreadsheets on Google Docs. Moreover, the access policy may even refer to users who do not exist in the system at the time the data is encrypted. Solving this problem requires an entirely new way of encrypting data.

A first natural approach would be **fully homomorphic encryption** (FHE, see above), but a second one is also **functional encryption**, that is an emerging paradigm for public-key encryption: it enables more fine-grained access control to encrypted data, for instance, the ability to specify a decryption policy in the ciphertext so that only individuals who satisfy the policy can decrypt, or the ability to associate keywords to a secret key so that it can only decrypt documents containing the keyword. Our work on functional encryption centers around two goals:

1. to obtain more efficient pairings-based functional encryption;
2. and to realize new functionalities and more expressive functional encryption schemes.

Another approach is **secure multi-party computation protocols**, where interactivity might provide privacy in a more efficient way. Recent implicit interactive proofs of knowledge can be a starting point. But stronger properties are first expected for improving privacy. They can also be integrated into new ad-hoc broadcast systems, in order to distribute the management among several parties, and eventually remove any trust requirements.

Strong privacy for the Cloud would have a huge societal impact since it would revolutionize the trust model: users would be able to make safe use of outsourced storage, namely for personal, financial and medical data, without having to worry about failures or attacks of the server.

5. Highlights of the Year

5.1. Awards

- The paper “The Gap-Problems: A New Class of Problems for the Security of Cryptographic Schemes”, by Tatsuaki Okamoto and David Pointcheval from PKC 2001, has received the Test-of-Time award at PKC 2019
- Raphaël Bost received the “2019 PhD Thesis Award” from the GDR Sécurité Informatique
- Mélissa Rossi received the L’Oréal-UNESCO For Women in Science Rising Talent Award Scholarship 2019.

6. New Results

6.1. Results

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

- Advanced primitives for privacy in the cloud
- Efficient functional encryption
- Attribute and predicate encryption schemes
- New primitives for efficient anonymous authentication
- Applications to machine learning
- Blockchain protocols
- Searchable Encryption

7. Partnerships and Cooperations

7.1. National Initiatives with Industry

7.1.1. ANBLIC: Analysis in Blind Clouds

Program: FUI

Duration: January 2018 – December 2020

Coordinator: Wallix

Partners: UPEC, CEA, Atos, SOGETI, CoeSSI

Local coordinator: David Pointcheval

The main goal is to industrialize for the first time several privacy enhancing technologies that are on the edge of theory and practice.

Fully Homomorphic Encryption let cloud providers compute arbitrary functions on their client’s encrypted data, ensuring at the same time full privacy and functionality. Functional Encryption is a refinement of classical encryption, which allows data owners to delegate fine-grained access to their data. Thus it is possible to enable the computation of aggregated statistics over your personal data, while cryptographically ensuring its confidentiality.

However both these technologies still suffer from prohibitive inefficiencies for business applications. ANBLIC’s academic partners will create new cryptographic schemes and performance models, tailored for industrial use cases, and create the first real-life scenario of encrypted queries on encrypted data and on open data.

7.1.2. RISQ: Regroupement de l’Industrie française pour la Sécurité Post-Quantique

Program: GDN

Duration: February 2017 – September 2020

Coordinator: Secure-IC

Partners: ANSSI, AIRBUS, C-S, CEA LIST, CryptoExperts, Inria/ENS/CASCADE, GEMALTO, Inria POLSYS, Inria AriC, IRISA, Orange Labs, THALES, UVSQ, PCQC

Local coordinator: Michel Abdalla and Phong Nguyen since September 2019

The main goal of RISQ is to help the French Industry and Academia become a significant international player in the transition to post-quantum cryptography.

7.2. National Collaborations with Academics

7.2.1. EnBiD: Encryption for Big Data

Program: ANR JCJC

Duration: October 2014 – September 2019

PI: Hoeteck Wee

Partners: Université Paris 2, Université Limoges

The main objective of this project is to study techniques for efficient and expressive functional encryption schemes. Functional encryption is a novel paradigm for public-key encryption that enables both fine-grained access control and selective computation on encrypted data, as is necessary to protect big, complex data in the cloud.

7.2.2. EfTrEC: Efficient Transferable E-Cash

Program: ANR JCJC

Duration: October 2016 – December 2019

PI: Georg Fuchsbauer

Partners: Université Paris 2

This project deals with e-cash systems which let users transfer electronic coins between them offline. The main objectives of this project are:

- establish a clean formal model for the primitive;
- construct schemes which are practically efficient;
- develop schemes that are resistant to attacks on quantum computers.

7.2.3. SaFED: Safe and Functional Encrypted Databases

Program: ANR JCJC

Duration: October 2019 – Septembre 2023

PI: Brice Minaud

Partners: ENS, DGA

This project addresses the security of encrypted databases, with the proposal of new searchable encryption techniques and deeper security analysis.

7.2.4. ALAMBIC: AppLicAtions of MalleaBIlity in Cryptography

Program: ANR PRC

Duration: October 2016 – September 2020

PI: Damien Vergnaud

Partners: ENS Lyon, Université Limoges

The main objectives of the proposal are the following:

- Define theoretical models for "malleable" cryptographic primitives that capture strong practical attacks (in particular, in the settings of secure computation outsourcing, server-aided cryptography, cloud computing and cryptographic proof systems);
- Analyze the security and efficiency of primitives and constructions that rely on malleability;
- Conceive novel cryptographic primitives and constructions (for secure computation outsourcing, server-aided cryptography, multi-party computation, homomorphic encryption and their applications);
- Implement these new constructions in order to validate their efficiency and effective security.

7.3. European Initiatives

7.3.1. CryptoCloud: Cryptography for the Cloud

Program: FP7 ERC Advanced Grant

Duration: June 2014 – May 2020

PI: David Pointcheval

The goal of the CryptoCloud project is to develop new interactive tools to provide privacy in the Cloud.

7.3.2. SAFEcrypto: Secure Architectures of Future Emerging Cryptography

Program: H2020

Duration: January 2015 – January 2019

Coordinator: The Queen's University of Belfast

Partners: Inria/ENS (France), Emc Information Systems International (Ireland), Hw Communications (United Kingdom), The Queen's University of Belfast (United Kingdom), Ruhr-Universitaet Bochum (Germany), Thales Uk (United Kingdom), Universita della Svizzera italiana (Switzerland), IBM Research Zurich (Switzerland)

Local coordinator: Michel Abdalla

SAFEcrypto will provide a new generation of practical, robust and physically secure post quantum cryptographic solutions that ensure long-term security for future ICT systems, services and applications. Novel public-key cryptographic schemes (digital signatures, authentication, public-key encryption, identity-based encryption) will be developed using lattice problems as the source of computational hardness. The project will involve algorithmic and design optimisations, and implementations of the lattice-based cryptographic schemes addressing the cost, energy consumption, performance and physical robustness needs of resource-constrained applications, such as mobile, battery-operated devices, and of real-time applications such as network security, satellite communications and cloud. Currently a significant threat to cryptographic applications is that the devices on which they are implemented leak information, which can be used to mount attacks to recover secret information. In SAFEcrypto the first analysis and development of physical-attack resistant methodologies for lattice-based cryptographic implementations will be undertaken. Effective models for the management, storage and distribution of the keys utilised in the proposed schemes (key sizes may be in the order of kilobytes or megabytes) will also be provided. This project will deliver proof-of-concept demonstrators of the novel lattice-based public-key cryptographic schemes for three practical real-word case studies with real-time performance and low power consumption requirements. In comparison to current state-of-the-art implementations of conventional public-key cryptosystems (RSA and Elliptic Curve Cryptography (ECC)), SAFEcrypto's objective is to achieve a range of lattice-based architectures that provide comparable area costs, a 10-fold speed-up in throughput for real-time application scenarios, and a 5-fold reduction in energy consumption for low-power and embedded and mobile applications.

7.3.3. ECRYPT-NET: Advanced Cryptographic Technologies for the Internet of Things and the Cloud

Program: H2020 ITN

Duration: March 2015 – February 2019

Coordinator: KU Leuven (Belgium)

Partners: KU Leuven (Belgium), Inria/ENS (France), Ruhr-Universität Bochum (Germany), Royal Holloway, University of London (UK), University of Bristol (UK), CryptoExperts (France), NXP Semiconductors (Belgium), Technische Universiteit Eindhoven (the Netherlands)

Local coordinator: Michel Abdalla

ECRYPT-NET is a research network of six universities and two companies, as well as 7 associated companies, that intends to develop advanced cryptographic techniques for the Internet of Things and the Cloud and to create efficient and secure implementations of those techniques on a broad range of platforms.

7.3.4. aSCEND: Secure Computation on Encrypted Data

Program: H2020 ERC Starting Grant

Duration: June 2015 – May 2021

PI: Hoeteck Wee

The goals of the aSCEND project are (i) to design pairing- and lattice-based functional encryption that are more efficient and ultimately viable in practice; and (ii) to obtain a richer understanding of expressive functional encryption schemes and to push the boundaries from encrypting data to encrypting software.

7.3.5. FENTEC: Functional Encryption Technologies

Program: H2020

Duration: January 2018 – December 2020

Coordinator: ATOS Spain SA

Scientific coordinator: Michel Abdalla

Partners: Inria/ENS (France), Flensburg University (Germany), KU Leuven (Belgium), University of Helsinki (Finland), Nagra (Switzerland), XLAB (Switzerland), University of Edinburgh (United Kingdom), WALLIX (France)

Local coordinator: Michel Abdalla

Functional encryption (FE) has recently been introduced as a new paradigm of encryption systems to overcome all-or-nothing limitations of classical encryption. In an FE system the decryptor deciphers a function over the message plaintext: such functional decryptability makes it feasible to process encrypted data (e.g. on the Internet) and obtain a partial view of the message plaintext. This extra flexibility over classical encryption is a powerful enabler for many emerging security technologies (i.e. controlled access, searching and computing on encrypted data, program obfuscation...). FENTEC's mission is to make the functional encryption paradigm ready for wide-range applications, integrating it in ICT technologies as naturally as classical encryption. The primary objective is the efficient and application-oriented development of functional encryption systems. FENTEC's team of cryptographers, software and hardware experts and information technology industry partners will document functional encryption needs of specific applications and subsequently design, develop, implement and demonstrate applied use of functional cryptography. Ultimately, a functional encryption library for both SW and HW-oriented application will be documented and made public so that it may be used by European ICT entities. With it, the FENTEC team will build emerging security technologies that increase the trustworthiness of the European ICT services and products. Concretely, the FENTEC team will showcase the expressiveness and versatility of the functional encryption paradigm in 3 use cases:

- Privacy-preserving digital currency, enforcing flexible auditing models
- Anonymous data analytics enabling computation of statistics over encrypted data, protecting European Fundamental Rights of Data Protection and Privacy
- Key and content distribution with improved performance & efficiency as foundational technology for establishing secure communication among a vast number of IOT devices.

7.4. International Initiatives with Industry

7.4.1. CryptBloC: Cryptography for the Blockchain

Partners: MSR Redmond (USA), MSR Cambridge (UK), Inria

Duration: October 2017 – October 2021

PI: Georg Fuchsbauer

The goal of this Microsoft-Inria joint project on privacy and decentralization is to use cryptography to improve privacy on the blockchain and decentralized systems more generally. We will investigate means of privacy-preserving authentication, such as electronic currencies, and other applications of blockchain and distributed transparency mechanisms.

7.5. International Research Visitors

7.5.1. Professors

- Sep 1 - Oct 31, 2019: Manuel Barbosa (University of Porto)
- Jun 20 - 21, 2019: Jean Paul Degabriele (TU Darmstadt)
- Jun 20 - 30, 2019: Joël Alwen (Wickr)
- Jul 4-5, 2019: David Wu (University of Virginia)

7.5.2. PhD students

- Jun 18 - 25, 2019: Ward Beullens (KU Leuven)
- Jun 15 - Jul 1, 2019: Rotem Tsabary (Weizmann)
- June 1 - 30, 2019: Hendrik Waldner (Edinburgh)
- Jun 23 - Jul 3, 2019: Naty Peter (Ben-Gurion University)

7.6. Internships

- Apr-Sep 2019: Hugo Marival (Ecole Polytechnique) - Michel Abdalla and David Pointcheval
- Apr-Sep 2019: Thibaut Bagory (ENS Paris-Saclay - UVSQ) - Brice Minaud
- Oct-Dec 2019: Marie Euler (X - DGA) - Brice Minaud

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. Events and Activities

- Quarterly Paris Crypto Days (https://pariscryptoday.github.io) supported by CryptoCloud and aSCEND
- Seminars are organized: https://crypto.di.ens.fr/web2py/index/seminars

- OPACity – Obfuscation: Proofs, Attacks & Candidates, (http://crypto-events.di.ens.fr/opacity/) Eurocrypt affiliated event. May 19, 2019, Darmstadt, supported by aSCEND
- BibTeX database of papers related to Cryptography, open and widely used by the community (https://cryptobib.di.ens.fr)

8.1.1.2. Steering Committees of International Conferences

- Steering committee of CANS: David Pointcheval
- Steering committee of PKC: David Pointcheval
- Steering committee of LATINCRYPT: Michel Abdalla (chair)
- Steering committee of Information-Theoretic Cryptography Conference: Hoeteck Wee

8.1.1.3. Board of International Organisations

- Board of the *International Association for Cryptologic Research* (IACR): Michel Abdalla (2013 – 2021)
- President-Elect of the *International Association for Cryptologic Research* (IACR): Michel Abdalla (2019 –)

8.1.2. Scientific Events Selection

8.1.2.1. Program Committee Member

- ACNS '19 (Bogotá, Colombia): Michel Abdalla
- ASIACRYPT '19 (Kobe, Japan): Brice Minaud, Hoeteck Wee
- CT-RSA '19 (San Francisco, California, USA): David Pointcheval
- EUROCRYPT '19 (Darmstadt, Germany): Michel Abdalla
- LATINCRYPT '19 (Santiago, Chile): Michel Abdalla
- PKC '20 (Edinburgh, Scotland): Georg Fuchsbauer

8.1.2.2. Editorial Boards of Journals

Editor-in-Chief

- of the *International Journal of Applied Cryptography (IJACT)* – Inderscience Publishers: David Pointcheval

Associate Editor

- of *ETRI Journal*: Michel Abdalla
- of *Journal of Mathematical Cryptology*: Phong Nguyen
- of *Applicable Algebra in Engineering, Communication and Computing*: David Pointcheval

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

- Master: Michel Abdalla, Brice Minaud, Cryptography, M2, MPRI
- Master: Phong Nguyen, Cryptography, M2, ESIEA
- Bachelor: Brice Minaud, Georg Fuchsbauer, David Pointcheval, Introduction to Cryptology, L3/M1, ENS
- Bachelor: Michel Abdalla, Formal Languages, Computability, and Complexity, L3/M1,
- Bachelor: Georg Fuchsbauer, Cryptography, École supérieure d'ingénieurs Léonard-de-Vinci

8.2.2. Defenses

- PhD: Romain Gay, Public-key Encryption, Revisited: Tight Security and Richer Functionalities, ENS, March 18th, 2019 (Supervisors: Michel Abdalla & Hoeteck Wee)

- PhD: Anca Nitulescu, A Tale of SNARKs: Quantum Resilience, Knowledge Extractability and Data Privacy, ENS, April 1st, 2019 (Supervisors: Michel Abdalla, Dario Fiore & David Pointcheval)
- PhD: Razvan Rosie, On the Achievability of White-Box Cryptography, ENS, May 28th, 2019 (Supervisor: Michel Abdalla)
- PhD: Louiza Khati, Full Disk Encryption and Beyond, ENS, July 15th, 2019 (Supervisor: Damien Vergnaud)
- PhD: Jérémy Chotard, Delegation in Functional Encryption, ENS, December 2nd, 2019 (Supervisor: David Pointcheval, with Duong Hieu Phan, at Limoges)

8.2.3. Supervision

- PhD in progress: Michele Orrù, Functional Encryption, from 2016, Georg Fuchsbauer & Hoeteck Wee
- PhD in progress: Balthazar Bauer, Transferable e-Cash, from 2017, Georg Fuchsbauer
- PhD in progress: Chloé Hébant, Big Data and Privacy, from 2017, David Pointcheval (with Duong Hieu Phan, at Limoges)
- PhD in progress: Mélissa Rossi, Post-Quantum Cryptography, from 2017, Michel Abdalla (with Henri Gilbert at ANSSI)
- PhD in progress: Antoine Plouviez, Privacy and Decentralization, from 2018, Georg Fuchsbauer
- PhD in progress: Baptiste Cottier, Privacy-preserving anomaly detection, from 2019, David Pointcheval (with Olivier Blazy, at Limoges)
- PhD in progress: Théo Ryffel, Privacy-preserving federated learning, from 2019, Francis Bach and David Pointcheval
- PhD in progress: Lénaïck Gouriou, Advanced encryption with post-quantum security, from 2019, David Pointcheval (with Cécile Delerablée at Leanear)

8.2.4. Committees

- PhD Marc Beunardeau. *Cryptographie Appliquée à la Sécurité des Systèmes d'Information* – Ecole Normale Supérieure – France – January 15th, 2019: David Pointcheval (Chair)
- HdR Catalin Hritcu. *The Quest for Formally Secure Compartmentalizing Compilation* – Ecole Normale Supérieure – France – January 29th, 2019: David Pointcheval
- PhD Romain Gay. *Public-key Encryption, Revisited: Tight Security and Richer Functionalities* – Ecole Normale Supérieure – France – March 18th, 2019: Michel Abdalla and Hoeteck Wee (Co-supervisors)
- PhD Mary Maller. *Practical zero-knowledge arguments from structured reference strings* – University College London – United Kingdom – March 26th, 2019: Georg Fuchsbauer (Reviewer)
- PhD Anca Nitulescu. *A Tale of SNARKs: Quantum Resilience, Knowledge Extractability and Data Privacy* – Ecole Normale Supérieure – France – April 1st, 2019: Michel Abdalla and David Pointcheval (Co-supervisors)
- PhD José Miguel López Becerra. *Provable Security Analysis for the PAKE Problem* – University of Luxembourg – Luxembourg – May 14th, 2019: Michel Abdalla
- PhD Razvan Rosie. *On the Achievability of White-Box Cryptography* – Ecole Normale Supérieure – France – May 28th, 2019: Michel Abdalla (Supervisor)
- HdR Olivier Blazy. *Hash Proofs Systems and Applications to Implicit Cryptography* – Université de Limoges – France – June 12th, 2019: David Pointcheval
- PhD Aurélien Dupin. *Secure Multi-Party Computation and Privacy* – Université Bretagne Loire – France – June 13rd, 2019: David Pointcheval (Co-supervisor)

- PhD Louiza Khati. *Full Disk Encryption and Beyond* – Ecole Normale Supérieure – France – July 15th, 2019: David Pointcheval
- PhD Julian Loss. *New Techniques for the Modular Analysis of Digital Signature Schemes* – Ruhr Universität Bochum – Germany – July 20th, 2019; David Pointcheval (Reviewer)
- PhD Aisling Connolly. *New Notions of Security, Broken Assumptions, and Increased Efficiency* – Ecole Normale Supérieure – France – September 13th, 2019: Michel Abdalla
- PhD Chen Qian. *Lossy Trapdoor Primitives, Zero-Knowledge Proofs and Applications* – Université de Rennes 1 – France – October 4th, 2019: Michel Abdalla
- PhD Alice Pellet–Mary. *On ideal lattices and the GGH13 multilinear map* – Ecole Normale Supérieure de Lyon – France – October 16th, 2019: David Pointcheval (Chair and Reviewer)
- HdR Guilhem Castagnos. *Cryptography based on Quadratic Fields: Cryptanalyses, Primitives and Protocols* – Université de Bordeaux – France – November 8th, 2019: David Pointcheval (Reviewer)
- PhD Joseph Lallemand. *Vote électronique : définitions et techniques d'analyse* – Université de Lorraine – France – November 8th, 2019: David Pointcheval (Reviewer)
- PhD Loïc Ferreira. *Secure Tunnels for Constrained Environments* – Université Bretagne Loire – France – November 18th, 2019: David Pointcheval
- PhD Pauline Bert. *Lattice-based signatures: from construction to implementation* – Université Bretagne Loire – France – November 29th, 2019: David Pointcheval
- PhD Jérémy Chotard. *Delegation in functional encryption* – Université de Limoges - France – Décember 2nd, 2019: David Pointcheval (Co-supervisor) and Michel Abdalla (Chair)

9. Bibliography

Major publications by the team in recent years

[1] M. ABDALLA, D. CATALANO, D. FIORE.*Verifiable Random Functions: Relations to Identity-Based Key Encapsulation and New Constructions*, in "Journal of Cryptology", 2014, vol. 27, n^{o} 3, p. 544-593

[2] M. ABE, G. FUCHSBAUER, J. GROTH, K. HARALAMBIEV, M. OHKUBO.*Structure-Preserving Signatures and Commitments to Group Elements*, in "Journal of Cryptology", 2016, vol. 29, n^{o} 2, p. 363–421

[3] F. BENHAMOUDA, O. BLAZY, C. CHEVALIER, D. POINTCHEVAL, D. VERGNAUD.*New Techniques for SPHFs and Efficient One-Round PAKE Protocols*, in "Advances in Cryptology – Proceedings of CRYPTO '13 (1)", R. CANETTI, J. A. GARAY (editors), Lecture Notes in Computer Science, Springer, 2013, vol. 8042, p. 449-475

[4] P. CHAIDOS, V. CORTIER, G. FUCHSBAUER, D. GALINDO.*BeleniosRF: A Non-interactive Receipt-Free Electronic Voting Scheme*, in "Proceedings of the 23rd ACM Conference on Computer and Communications Security (CCS '16)", E. R. WEIPPL, S. KATZENBEISSER, C. KRUEGEL, A. C. MYERS, S. HALEVI (editors), ACM Press, 2016, p. 1614–1625

[5] Y. DODIS, D. POINTCHEVAL, S. RUHAULT, D. VERGNAUD, D. WICHS.*Security Analysis of Pseudo-Random Number Generators with Input: /dev/random is not Robust*, in "Proceedings of the 20th ACM Conference on Computer and Communications Security (CCS '13)", Berlin, Germany, V. D. GLIGOR, M. YUNG (editors), ACM Press, 2013, p. 647–658

[6] R. Gay, D. Hofheinz, E. Kiltz, H. Wee.*Tightly CCA-Secure Encryption Without Pairings*, in "Advances in Cryptology – Proceedings of Eurocrypt '16 (2)", M. Fischlin, J.-S. Coron (editors), Lecture Notes in Computer Science, Springer, 2016, vol. 9665, p. 1–27

[7] S. Gorbunov, V. Vaikuntanathan, H. Wee.*Predicate Encryption for Circuits from LWE*, in "Advances in Cryptology – Proceedings of CRYPTO '15 (2)", R. Gennaro, M. Robshaw (editors), Lecture Notes in Computer Science, Springer, 2015, vol. 9216, p. 503-523

[8] V. Lyubashevsky, C. Peikert, O. Regev.*On Ideal Lattices and Learning with Errors over Rings*, in "Journal of the ACM", 2013, vol. 60, n^o 6, p. 43:1–43:35

[9] W. Quach, H. Wee, D. Wichs.*Laconic Function Evaluation and Applications*, in "59th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2018)", M. Thorup (editor), IEEE, 2018

Publications of the year

Doctoral Dissertations and Habilitation Theses

[10] J. Chotard.*Delegation in functional encryption*, Université de Limoges, France, December 2019, https://hal.archives-ouvertes.fr/tel-02394349

[11] R. Gay.*Public-Key Encryption, Revisited: Tight Security and Richer Functionalities*, PSL Research University, March 2019, https://tel.archives-ouvertes.fr/tel-02137987

[12] L. Khati.*Full Disk Encryption and Beyond*, Université PSL ; ENS Paris - Ecole Normale Supérieure de Paris, July 2019, Equipe cascade, ENS, Inria, https://tel.archives-ouvertes.fr/tel-02318449

[13] A. Nitulescu.*A tale of SNARKs: quantum resilience, knowledge extractability and data privacy*, École Normale Supérieure (Paris), April 2019, https://hal.archives-ouvertes.fr/tel-02129544

[14] R. Roşie.*On the achievability of white-box cryptography*, PSL Research University, May 2019, https://tel.archives-ouvertes.fr/tel-02332996

Articles in International Peer-Reviewed Journal

[15] M. Abdalla, F. Benhamouda, D. Pointcheval.*On the Tightness of Forward-Secure Signature Reductions*, in "Journal of Cryptology", January 2019, vol. 32, n^o 1, p. 84-150 [*DOI :* 10.1007/s00145-018-9283-2], https://hal.inria.fr/hal-01722996

[16] A. Connolly, P. Farshim, G. Fuchsbauer.*Security of Symmetric Primitives against Key-Correlated Attacks*, in "IACR Transactions on Symmetric Cryptology", September 2019, https://hal.inria.fr/hal-02396311

[17] M. Seo, M. Abdalla, D. H. Lee, J. H. Park.*New technique for chosen-ciphertext security based on non-interactive zero-knowledge*, in "Information Sciences", July 2019, vol. 490, p. 18-35 [*DOI :* 10.1016/j.ins.2019.03.063], https://hal.inria.fr/hal-02135837

International Conferences with Proceedings

[18] M. Abdalla, F. Benhamouda, R. Gay.*From Single-Input to Multi-client Inner-Product Functional Encryption*, in "Advances in Cryptology – ASIACRYPT 2019", Kobe, Japan, S. D. Galbraith, S. Moriai

(editors), Lecture Notes in Computer Science, November 2019, vol. 11923, p. 552-582 [*DOI :* 10.1007/978-3-030-34618-8_19], https://hal.inria.fr/hal-02375577

[19] M. ABDALLA, F. BENHAMOUDA, M. KOHLWEISS, H. WALDNER.*Decentralizing Inner-Product Functional Encryption*, in "Public-Key Cryptography – PKC 2019", Beijing, China, D. LIN, K. SAKO (editors), Lecture Notes in Computer Science, April 2019, vol. 11443, p. 128-157 [*DOI :* 10.1007/978-3-030-17259-6_5], https://hal.inria.fr/hal-02135871

[20] M. ABDALLA, F. BENHAMOUDA, A. PASSELÈGUE.*Algebraic XOR-RKA-Secure Pseudorandom Functions from Post-Zeroizing Multilinear Maps*, in "Advances in Cryptology – ASIACRYPT 2019", Kobe, Japan, S. D. GALBRAITH, S. MORIAI (editors), Lecture Notes in Computer Science, November 2019, vol. 11922, p. 386-412 [*DOI :* 10.1007/978-3-030-34621-8_14], https://hal.inria.fr/hal-02375594

[21] M. BARBOSA, D. CATALANO, A. SOLEIMANIAN, B. WARINSCHI.*Efficient Function-Hiding Functional Encryption: From Inner-Products to Orthogonality*, in "Topics in Cryptology – CT-RSA 2019", San Francisco, United States, M. MATSUI (editor), Lecture Notes in Computer Science, February 2019, vol. 11405, p. 127-148 [*DOI :* 10.1007/978-3-030-12612-4_7], https://hal.inria.fr/hal-02363215

[22] A. BAUER, H. GILBERT, G. RENAULT, M. ROSSI.*Assessment of the Key-Reuse Resilience of NewHope*, in "CT-RSA 2019 - The Cryptographers' Track at the RSA Conference", San Francisco, United States, M. MATSUI (editor), Lecture Notes in Computer Science, Springer, February 2019, vol. 11405, p. 272-292 [*DOI :* 10.1007/978-3-030-12612-4_14], https://hal.archives-ouvertes.fr/hal-02139910

[23] W. BEULLENS, H. WEE.*Obfuscating Simple Functionalities from Knowledge Assumptions*, in "PKC 2019 - International Conference on Practice and Theory of Public Key Cryptography", Beijing, China, April 2019, p. 254-283 [*DOI :* 10.1007/978-3-030-17259-6_9], https://hal.inria.fr/hal-02358436

[24] F. BOURSE, D. POINTCHEVAL, O. SANDERS.*Divisible E-Cash from Constrained Pseudo-Random Functions*, in "ASIACRYPT 2019 - 25th Annual International Conference on the Theory and Application of Cryptology and Information Security", Kobe, Japan, Advances in Cryptology, December 2019, vol. LNCS, n^{o} 11922, https://hal.inria.fr/hal-02357173

[25] E. DUFOUR-SANS, D. POINTCHEVAL.*Unbounded Inner-Product Functional Encryption with Succinct Keys*, in "ACNS 2019 - 17th International Conference on Applied Cryptography and Network Security", Bogota, Colombia, Springer, May 2019, vol. LNCS, n^{o} 11464, p. 426-441 [*DOI :* 10.1007/978-3-030-21568-2_21], https://hal.inria.fr/hal-02357165

[26] G. FUCHSBAUER.*WI Is Not Enough: Zero-Knowledge Contingent (Service) Payments Revisited*, in "ACM CCS 2019 - 26th ACM Conference on Computer and Communications Security", London, United Kingdom, ACM Press, November 2019, p. 49-62 [*DOI :* 10.1145/3319535.3354234], https://hal.inria.fr/hal-02396308

[27] G. FUCHSBAUER, C. KAMATH, K. KLEIN, K. PIETRZAK.*Adaptively Secure Proxy Re-encryption*, in "Public-Key Cryptography – PKC 2019", Beijing, China, D. LIN, K. SAKO (editors), LNCS - Lecture Notes in Computer Science, Springer, April 2019, vol. 11443 [*DOI :* 10.1007/978-3-030-17259-6_11], https://hal.inria.fr/hal-02396301

[28] G. FUCHSBAUER, M. ORRÙ, Y. SEURIN.*Aggregate Cash Systems: A Cryptographic Investigation of Mimblewimble*, in "Advances in Cryptology - EUROCRYPT 2019", Darmstadt, Germany, Y. ISHAI, V. RIJMEN

(editors), LNCS - Lecture Notes in Computer Science, Springer, May 2019, vol. 11476 [*DOI :* 10.1007/978-3-030-17653-2_22], https://hal.inria.fr/hal-02396305

[29] J. GONG, B. WATERS, H. WEE.*ABE for DFA from k-Lin*, in "CRYPTO 2019 - 39th Annual International Cryptology Conference", Santa Barbara, United States, August 2019, p. 732-764 [*DOI :* 10.1007/978-3-030-26951-7_25], https://hal.inria.fr/hal-02358440

[30] P. GRUBBS, M.-S. LACHARITÉ, B. MINAUD, K. G. PATERSON.*Learning to Reconstruct: Statistical Learning Theory and Encrypted Database Attacks*, in "IEEE Symposium on Security and Privacy (S&P) 2019", San Francisco, United States, May 2019, https://hal.inria.fr/hal-01974962

[31] C. HÉBANT, D. H. PHAN, D. POINTCHEVAL.*Decentralized Evaluation of Quadratic Polynomials on Encrypted Data*, in "ISC 2019 - International Conference on Information Security", New York, United States, Z. LIN, C. PAPAMANTHOU, M. POLYCHRONAKIS (editors), ISC 2019 - International Conference on Information Security, Springer, September 2019, vol. LNCS, n^o 11723, p. 87-106 [*DOI :* 10.1007/978-3-030-30215-3_5], https://hal.archives-ouvertes.fr/hal-02345627

[32] L. KHATI, D. VERGNAUD.*Analysis and Improvement of an Authentication Scheme in Incremental Cryptography*, in "Selected Areas in Cryptography - SAC 2018", Calgary, Canada, C. CID, M. J. JACOBSON JR. (editors), Lecture Notes in Computer Science, Springer, January 2019, vol. 11349, p. 50-70 [*DOI :* 10.1007/978-3-030-10970-7_3], https://hal.inria.fr/hal-01893905

[33] L. KOWALCZYK, H. WEE.*Compact Adaptively Secure ABE for* NC^1 *from k-Lin*, in "EUROCRYPT 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques", Darmstadt, Germany, April 2019, p. 3-33 [*DOI :* 10.1007/978-3-030-17653-2_1], https://hal.inria.fr/hal-02358429

[34] T. RYFFEL, E. DUFOUR-SANS, R. GAY, F. BACH, D. POINTCHEVAL.*Partially Encrypted Machine Learning using Functional Encryption*, in "NeurIPS 2019 - Thirty-third Conference on Neural Information Processing Systems", Vancouver, Canada, Advances in Neural Information Processing Systems, December 2019, https://arxiv.org/abs/1905.10214 , https://hal.inria.fr/hal-02357181

Research Reports

[35] M. ABDALLA, M. BARBOSA.*Perfect Forward Security of SPAKE2*, IACR Cryptology ePrint Archive, October 2019, n^o Report 2019/1194, https://hal.inria.fr/hal-02317002

[36] M. ABDALLA, F. BENHAMOUDA, R. GAY.*From Single-Input to Multi-Client Inner-Product Functional Encryption*, IACR Cryptology ePrint Archive, May 2019, n^o Report 2019/487, https://hal.inria.fr/hal-02135963

[37] M. ABDALLA, F. BENHAMOUDA, M. KOHLWEISS, H. WALDNER.*Decentralizing Inner-Product Functional Encryption*, IACR Cryptology ePrint Archive, January 2019, n^o Report 2019/020, https://hal.inria.fr/hal-02317011

Team COML

Cognitive Machine Learning

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
Paris

THEME
Language, Speech and Audio

Table of contents

Team COML

Creation of the Team: 2017 May 04

Keywords:

Computer Science and Digital Science:

A2.5.1. - Software Architecture & Design
A2.5.4. - Software Maintenance & Evolution
A2.5.5. - Software testing
A3.4.2. - Unsupervised learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.8. - Deep learning
A5.7. - Audio modeling and processing
A5.7.1. - Sound
A5.7.3. - Speech
A5.7.4. - Analysis
A5.8. - Natural language processing
A6.3.3. - Data processing
A9.2. - Machine learning
A9.3. - Signal analysis
A9.4. - Natural language processing
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B1.2. - Neuroscience and cognitive science
B1.2.2. - Cognitive science
B2.2.6. - Neurodegenerative diseases
B2.5.2. - Cognitive disabilities
B9.6.1. - Psychology
B9.6.8. - Linguistics
B9.8. - Reproducibility
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientist

Justine Cassell [Inria, Advanced Research Position, from Oct 2019]

Faculty Member

Emmanuel Dupoux [Team leader, École des hautes études en sciences sociales, Professor, HDR]

External Collaborators

Nawal Abboub [École Normale Supérieure de Paris, from Apr 2019 until Jul 2019]
Ewan Dunbar [Univ Denis Diderot, from Apr 2019]

Technical Staff

Robin Algayres [École Normale Supérieure de Paris, Engineer, until Sep 2019]
Mathieu Bernard [Inria, Engineer]
Xuan-Nga Cao [École des hautes études en sciences sociales, Engineer]
Nicolas Hamilakis [École Normale Supérieure de Paris, Engineer]
Julien Karadayi [Inria, Engineer, from Apr 2019]
Manel Khentout [École Normale Supérieure de Paris, Engineer]
Marvin Lavechin [École Normale Supérieure de Paris, Engineer]
Malik Ould Arbi [CNRS, Engineer, from Feb 2019]
Hadrien Titeux [École Normale Supérieure de Paris, Engineer]
Kristina Madden [École Normale Supérieure de Paris, Designer]
Mohamed Zaiem [Univ Denis Diderot, Engineer, from Sep 2019]

PhD Students

Robin Algayres [École Normale Supérieure de Paris, PhD Student, from Oct 2019]
Rahma Chaabouni [École Normale Supérieure de Paris, PhD Student]
Juliette Millet [École Normale Supérieure de Paris, PhD Student]
Rachid Riad [École Normale Supérieure de Paris, PhD Student]
Neil Zeghidour [Facebook, PhD Student, until Feb 2019]
Ronan Riochet [PhD Student, co-advised E. Dupoux, I. Laptev, J. Sivic]

Administrative Assistants

Chantal Chazelas [Inria, Administrative Assistant, until Jun 2019]
Meriem Guemair [Inria, Administrative Assistant, until Nov 2019]
Catherine Urban [École des hautes études en sciences sociales, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

Brain-inspired machine learning algorithms combined with big data have recently reached spectacular results, equalling or beating humans on specific high level tasks (e.g. the game of go). However, there are still a lot of domains in which even humans infants outperform machines: unsupervised learning of rules and language, common sense reasoning, and more generally, cognitive flexibility (the ability to quickly transfer competence from one domain to another one).

The aim of the Cognitive Computing team is to *reverse engineer* such human abilities, i.e., to construct effective and scalable algorithms which perform as well (or better) than humans, when provided with similar data, study their mathematical and algorithmic properties and test their empirical validity as models of humans by comparing their output with behavioral and neuroscientific data. The expected results are more adaptable and autonomous machine learning algorithm for complex tasks, and quantitative models of cognitive processes which can used to predict human developmental and processing data. Most of the work is focused on speech and language and common sense reasoning.

3. Research Program

3.1. Background

In recent years, Artificial Intelligence (AI) has achieved important landmarks in matching or surpassing human level performance on a number of high level tasks (playing chess and go, driving cars, categorizing picture, etc., [31], [34], [39], [30], [36]). These strong advances were obtained by deploying on large amounts of data, massively parallel learning architectures with simple brain-inspired ‘neuronal’ elements. However, humans brains still outperform machines in several key areas (language, social interactions, common sense reasoning, motor skills), and are more flexible : Whereas machines require extensive expert knowledge and massive training for each particular application, humans learn autonomously over several time scales: over the developmental scale (months), humans infants acquire cognitive skills with noisy data and little or no expert feedback (weakly/unsupervised learning)[1]; over the short time scale (minutes, seconds), humans combine previously acquired skills to solve new tasks and apply rules systematically to draw inferences on the basis of extremely scarce data (learning to learn, domain adaptation, one- or zero-shot learning) [33].

The general aim of CoML, following the roadmap described in [1], is to bridge the gap in cognitive flexibility between humans and machines learning in language processing and common sense reasoning by reverse engineering how young children between 1 and 4 years of age learn from their environment. We conduct work along two axes: the first one, which we called *Developmental AI* is focused on building infant inspired machine learning algorithms. The second axis is devoted to using the developed algorithms to conduct *quantitative studies* of how infant learn across diverse environments.

3.2. Weakly/Unsupervised Learning

Much of standard machine learning is construed as regression or classification problems (mapping input data to expert-provided labels). Human infants rarely learn in this fashion, at least before going to school: they learn language, social cognition, and common sense autonomously (without expert labels) and when adults provide feedback, it is ambiguous and noisy and cannot be taken as a gold standard. Modeling or mimicking such achievement requires deploying unsupervised or weakly supervised algorithms which are less well known than their supervised counterparts.

We take inspiration from infant's landmarks during their first years of life: they are able to learn acoustic models, a lexicon, and susbtantive elements of language models and world models from raw sensory inputs. Building on previous work [3], [7], [11], we use DNN and Bayesian architectures to model the emergence of linguistic representations without supervision. Our focus is to establish how the labels in supervised settings can be replaced by weaker signals coming either from multi-modal input or from hierarchically organised linguistic levels.

At the level of phonetic representations, we study how cross-modal information (lips and self feedback from articulation) can supplement top-down lexical information in a weakly supervised setting. We use Siamese architectures or Deep CCA algorithms to combine the different views. We study how an attentional framework and uncertainty estimation can flexibly combine these informations in order to adapt to situations where one view is selectively degraded.

At the level of lexical representations, we study how audio/visual parallel information (ie. descriptions of images or activities) can help in segmenting and clustering word forms, and vice versa, help in deriving useful visual features. To achieve this, we will use architectures deployed in image captioning or sequence to sequence translation [37].

At the level of semantic and conceptual representations, we study how it is possible to learn elements of the laws of physics through the observation of videos (object permanence, solidity, spatio-temporal continuity, inertia, etc.), and how objects and relations between objects are mapped onto language.

3.3. Evaluating Machine Intelligence

Increasingly, complicated machine learning systems are being incorporated into real-life applications (e.g. self-driving cars, personal assistants), even though they cannot be formally verified, guaranteed statistically, nor even explained. In these cases, a well defined *empirical approach* to evaluation can offer interesting insights into the functioning and offer some control over these algorithms.

Several approaches exist to evaluate the 'cognitive' abilities of machines, from the subjective comparison of human and machine performance [38] to application-specific metrics (e.g., in speech, word error rate). A recent idea consist in evaluating an AI system in terms of it's *abilities* [32] , i.e., functional components within a more global cognitive architecture [35]. Psychophysical testing can offer batteries of tests using simple tasks that are easy to understand by humans or animals (e.g, judging whether two stimuli are same or different, or judging whether one stimulus is 'typical') which can be made selective to a specific component and to rare but difficult or adversarial cases. Evaluations of learning rate, domain adaptation and transfer learning are simple applications of these measures. Psychophysically inspired tests have been proposed for unsupervised speech and language learning [10], [6].

3.4. Documenting human learning

Infants learn their first language in a spontaneous fashion, across a lot of variation in amount of speech and the nature of the infant/adult interaction. In some linguistic communities, adults barely address infants until they can themselves speak. Despite these large variations in quantity and content, language learning proceeds at similar paces. Documenting such resilience is an essential step in understanding the nature of the learning algorithms used by human infants. Hence, we propose to collect and/or analyse large datasets of inputs to infants and correlate this with outcome measure (phonetic learning, vocabulary growth, syntactic learning, etc.).

4. Application Domains

4.1. Speech processing for underresourced languages

We plan to apply our algorithms for the unsupervised discovery of speech units to problems relevant to language documentation and the construction of speech processing pipelines for underresourced languages.

4.2. Tools for the analysis of naturalistic speech corpora

Daylong recordings of speech in the wild gives rise a to number of specific analysis difficulties. We plan to use our expertise in speech processing to develop tools for performing signal processing and helping annotation of such resources for the purpose of phonetic or linguistic analysis.

5. New Software and Platforms

5.1. intphys

IntPhys: A Benchmark for Visual Intuitive Physics Reasoning

KEYWORDS: Competition - Physical simulation - Artificial intelligence - Video Game

FUNCTIONAL DESCRIPTION: The intphys benchmark can be applied to any vision system, engineered, or trained, provided it can output a scalar when presented with a video clip, which should correspond to how physically plausible the video clip is. Our test set contains well matched videos of possible versus impossible events, and the metric consists in measuring how well the vision system can tell apart the possible from the impossible events..

- Contact: Mathieu Bernard
- URL: http://www.intphys.com

5.2. shennong

KEYWORDS: Speech processing - Python - Information extraction - Audio signal processing

FUNCTIONAL DESCRIPTION: Shennong is a Python library which implement the most used methods for speech features extraction. Features extraction is the first step of every speech processing pipeline.

Shennong provides the following functionalities: - implementation of the main methods from state of the art (including pre and post processing) - exhaustive documentation and tests - usage from a Python API or a command line tool - simple and coherent interface

- Contact: Mathieu Bernard
- URL: https://coml.lscp.ens.fr/docs/shennong

5.3. Seshat

Seshat Audio Annotation Platform

KEYWORDS: Audio - Speech - Web Application - Speech-text alignment

FUNCTIONAL DESCRIPTION: A web application to ease audio annotation campaigns, while also enabling the campaign manager to ensure that all annotations stick to a predefined format.

- Partner: ENS Paris
- Contact: Hadrien Titeux
- URL: https://github.com/bootphon/seshat

5.4. pyGammaAgreement

KEYWORDS: Reliability - Measures

FUNCTIONAL DESCRIPTION: Python library for measuring inter and intra annotator reliability for annotation sequences

- Contact: Emmanuel Dupoux

5.5. phonemizer

KEYWORD: Text

FUNCTIONAL DESCRIPTION:

- Conversion of a text into its phonemic representation
- Wrapper on speech synthesis programs espeak and festival
- Contact: Mathieu Bernard
- URL: https://github.com/bootphon/phonemizer

6. New Results

6.1. Unsupervised learning

Huamns learn to speak and to perceive the world in a largely self-supervised fashion. Yet, most of machine learning is still devoted to supervised algorithms that rely on abundant quantities of human labelled data. We have used humans as sources of inspiration for developing 3 novel machine learning benchmarks in order to push the field towards self-supervised learning.

- In the Zero Resource Speech Challenge 2019 [19], presented as a special session at Interspeech 2019, we propose to build a speech synthesizer without any text or phonetic labels: hence, TTS without T (text-to-speech without text). We provide raw audio for a target voice in an unknown language (the Voice dataset), but no alignment, text or labels. Participants must discover subword units in an unsupervised way (using the Unit Discovery dataset) and align them to the voice recordings in a way that works best for the purpose of synthesizing novel utterances from novel speakers, similar to the target speaker's voice. We describe the metrics used for evaluation, a baseline system consisting of unsupervised subword unit discovery plus a standard TTS system, and a topline TTS using gold phoneme transcriptions. We present an overview of the 19 submitted systems from 11 teams and discuss the main results.

- In [27], we introduce a new collection of spoken English audio suitable for training speech recognition systems under limited or no supervision. It is derived from open-source audio books from the LibriVox project. It contains over 60K hours of audio, which is, to our knowledge, the largest freely-available corpus of speech. The audio has been segmented using voice activity detection and is tagged with SNR, speaker ID and genre descriptions. Additionally, we provide baseline systems and evaluation metrics working under three settings: (1) the zero resource/unsupervised setting (ABX), (2) the semi-supervised setting (PER, CER) and (3) the distant supervision setting (WER). Settings (2) and (3) use limited textual resources (10 minutes to 10 hours) aligned with the speech. Setting (3) uses large amounts of unaligned text. They are evaluated on the standard LibriSpeech dev and test sets for comparison with the supervised state-of-the-art.
- In order to reach human performance on complex visual tasks, artificial systems need to incorporate a significant amount of understanding of the world in terms of macroscopic objects, movements, forces, etc. Inspired by work on intuitive physics in infants, we propose in [28] an evaluation framework which diagnoses how much a given system understands about physics by testing whether it can tell apart well matched videos of possible versus impossible events. The test requires systems to compute a physical plausibility score over an entire video. It is free of bias and can test a range of specific physical reasoning skills. We then describe the first release of a benchmark dataset aimed at learning intuitive physics in an unsupervised way, using videos constructed with a game engine. We describe two Deep Neural Network baseline systems trained with a future frame prediction objective and tested on the possible versus impossible discrimination task. The analysis of their results compared to human data gives novel insights in the potentials and limitations of next frame prediction architectures. This benchmark is currently being used in the DARPA project Machine Common Sense.

6.2. Language emergence in communicative agents

In this relatively new research topic, which is currently the focus of Rahma Chaabouni's PhD thesis, we study the inductive biases of neural systems by presenting them with few or no data.

- In [18], we study LSTMs' biases with respect to "natural" word-order constraints. To this end, we train them to communicate about trajectories in a grid world, using an artificial language that reflect or violate various natural language trends, such as the tendency to avoid redundancy or to minimize long-distance dependencies. We measure the speed of individual learning and the generational stability of language patterns in an iterative learning setting. Our results show a mixed picture. If LSTMs are affected by some "natural" word-order constraints, such as a preference for iconic orders and short-distance constructions, they have a preference toward redundant languages.
- In [25], we ask whether LSTMs have least-effort constraints and how this can affect their language. We let the neural systems develop their own language, to study a fundamental characteristic of natural language; Zipf's Law of Abbreviation (ZLA). In other words, we inverstigate if, even with the lack of the least-effort, LSTMs would produce a ZLA-like distribution like what we observe in natural language. Surprisingly, we find that networks develop an anti-efficient encoding scheme, in which the most frequent inputs are associated to the longest messages, and messages in general are skewed towards the maximum length threshold. This anti-efficient code appears easier to discriminate for the listener, and, unlike in human communication, the speaker does not impose a contrasting least-effort pressure towards brevity, as observed in [18]. Indeed, when the cost function includes a penalty for longer messages, the resulting message distribution starts respecting (ZLA). Our analysis stresses the importance of studying the basic features of emergent communication in a highly controlled setup, to ensure the latter will not strand too far from human language. Moreover, we present a concrete illustration of how different functional pressures can lead to successful communication codes that lack basic properties of human language, thus highlighting the role such pressures play in the latter.

- There is renewed interest in simulating language emergence among deep neural agents that communicate to jointly solve a task, spurred by the practical aim to develop language-enabled interactive AIs, and by theoretical questions about the evolution of human language. However, optimizing deep architectures connected by a discrete communication channel (such as that in which language emerges) is technically challenging. In [21], we introduce EGG, a toolkit that greatly simplifies the implementation of emergent-language communication experiments. EGG's modular design provides a set of building blocks that the user can combine to create new communication games, easily navigating the optimization and architecture space. We hope that the tool will lower the technical barrier, and encourage researchers from various backgrounds to do original work in this exciting area/

6.3. Evaluation of AI algorithms

Machine learning algorithms are typically evaluated in terms of end-to-end tasks, but it is very often difficult to get a grasp of how they achieve these tasks, what could be their break point, and more generally, how they would compare to the algorithms used by humans to do the same tasks. This is especially true of Deep Learning systems which are particularly opaque. The team develops evaluation methods based on psycholinguistic/linguistic criteria, and deploy them for systematic comparison of systems.

- Recurrent neural networks (RNNs) can learn continuous vector representations of symbolic structures such as sequences and sentences; these representations often exhibit linear regularities (analogies). Such regularities motivate our hypothesis that RNNs that show such regularities implicitly compile symbolic structures into tensor product representations (TPRs; Smolensky, 1990), which additively combine tensor products of vectors representing roles (e.g., sequence positions) and vectors representing fillers (e.g., particular words). To test this hypothesis, we introduce Tensor Product Decomposition Networks (TPDNs), which use TPRs to approximate existing vector representations. We demonstrate using synthetic data that TPDNs can successfully approximate linear and tree-based RNN autoencoder representations, suggesting that these representations exhibit interpretable compositional structure; we explore the settings that lead RNNs to induce such structure-sensitive representations. By contrast, further TPDN experiments show that the representations of four models trained to encode naturally-occurring sentences can be largely approximated with a bag of words, with only marginal improvements from more sophisticated structures. We conclude that TPDNs provide a powerful method for interpreting vector representations, and that standard RNNs can induce compositional sequence representations that are remarkably well approximated by TPRs; at the same time, existing training tasks for sentence representation learning may not be sufficient for inducing robust structural representations.
- LSTMs have proven very successful at language modeling. However, it remains unclear to what extent they are able to capture complex morphosyntactic structures. In [29], we examine whether LSTMs are sensitive to verb argument structures. We introduce a German grammaticality dataset in which ungrammatical sentences are constructed by manipulating case assignments (eg substituting nominative by accusative or dative). We find that LSTMs are better than chance in detecting incorrect argument structures and slightly worse than humans tested on the same dataset. Surprisingly, LSTMs are contaminated by heuristics not found in humans like a preference toward nominative noun phrases. In other respects they show human-similar results like biases for particular orders of case assignments.
- Pater (2019) proposes to use neural networks to model learning within existing grammatical frameworks. In [16] we argue that there is a fundamental gap to be bridged that does not receive enough attention : how can we use neural networks to examine whether it is possible to learn some linguistic representation (a tree, for example) when, after learning is finished, we cannot even tell if this is the type of representation that has been learned (all we see is a sequence of numbers)? Drawing a correspondence between an abstract linguistic representational system and an opaque parameter vector that can (or perhaps cannot) be seen as an instance of such a representation is an implementational mapping problem. Rather than relying on existing frameworks that propose

partial solutions to this problem, such as harmonic grammar, we suggest that fusional research of the kind proposed needs to directly address how to ‘find’ linguistic representations in neural network representations.

6.4. Learnability relevant descriptions of linguistic corpora

Evidently, infants are acquiring their language based on whatever linguistic input is available around them. The extent of variation that can be found across languages, cultures and socio-economic background provides strong constraints (lower bounds on data, higher bounds on noise, and variation and ambiguity) for language learning algorithms.

- Previous computational modeling suggests it is much easier to segment words from child-directed (CDS) than adult-directed speech (ADS). However, this conclusion is based on data collected in the laboratory, with CDS from play sessions and ADS between a parent and an experimenter, which may not be representative of ecologically-collected CDS and ADS. In [15], fully naturalistic ADS and CDS collected with a non-intrusive recording device as the child went about her day were analyzed with a diverse set of algorithms. The difference between registers was small compared to differences between algorithms, it reduced when corpora were matched, and it even reversed under some conditions. These results highlight the interest of studying learnability using naturalistic corpora and diverse algorithmic definitions.
- A number of unsupervised learning algorithms have been proposed in the last 20 years for modeling early word learning, some of which have been implemented computationally, but whose results remain difficult to compare across papers. In [14], we created a tool that is open source, enables reproducible results, and encourages cumulative science in this domain. WordSeg has a modular architecture: It combines a set of corpora description routines, multiple algorithms varying in complexity and cognitive assumptions (including several that were not publicly available, or insufficiently documented), and a rich evaluation package. In the paper, we illustrate the use of this package by analyzing a corpus of child-directed speech in various ways, which further allows us to make recommendations for experimental design of follow-up work. Supplementary materials allow readers to reproduce every result in this paper, and detailed online instructions further enable them to go beyond what we have done. Moreover, the system can be installed within container software that ensures a stable and reliable environment. Finally, by virtue of its modular architecture and transparency, WordSeg can work as an open-source platform, to which other researchers can add their own segmentation algorithms.

6.5. Test of the psychological validity of AI algorithms.

In this section, we focus on the utilisation of machine learning algorithms of speech and language processing to derive testable quantitative predictions in humans (adults or infants).

- In [24], we compare the performance of humans (English and French listeners) versus an unsupervised speech model in a perception experiment (ABX discrimination task). Although the ABX task has been used for acoustic model evaluation in previous research, the results have not, until now, been compared directly with human behaviour in an experiment. We show that a standard, well-performing model (DPGMM) has better accuracy at predicting human responses than the acoustic baseline. The model also shows a native language effect, better resembling native listeners of the language on which it was trained. However, the native language effect shown by the models is different than the one shown by the human listeners, and, notably , the models do not show the same overall patterns of vowel confusions.
- Word learning relies on the ability to master the sound con-trasts that are phonemic (i.e., signal meaning difference) in a given language. Though the timeline of phoneme development has been studied extensively over the past few decades, the mechanism of this development is poorly understood. In [20], we take inspiration from computational modeling work in language grounding

where phonetic and visual information is learned jointly. In this study, we varied the taxonomic distance of pairs of objects and tested how adult learners judged the phonemic status of the sound contrast associated with each of these pairs. We found that judgments were sensitive to gradients in the taxonomic structure, suggesting that learners use probabilistic information at the semantic level to optimize the accuracy of their judgements at the phonological level. The findings provide evidence for an interaction between phonological learning and meaning generalization in human learning.

6.6. Applications and tools for researchers

Some of CoMLs' activity is to produce speech and language technology tools that facilitate research into language development or clinical applications.

- Speech classifiers of paralinguistic traits traditionally learn from diverse hand-crafted low-level features, by selecting the relevant information for the task at hand. We explore an alternative to this selection, by learning jointly the classifier, and the feature extraction. Recent work on speech recognition has shown improved performance over speech features by learning from the waveform. In [24], we extend this approach to paralinguistic classification and propose a neural network that can learn a filterbank, a normalization factor and a compression power from the raw speech, jointly with the rest of the architecture. We apply this model to dysarthria detection from sentence-level audio recordings. Starting from a strong attention-based baseline on which mel-filterbanks out-perform standard low-level descriptors, we show that learning the filters or the normalization and compression improves over fixed features by 10% absolute accuracy. We also observe a gain over OpenSmile features by learning jointly the feature extraction, the normalization, and the compression factor with the architecture. This constitutes a first attempt at learning jointly all these operations from raw audio for a speech classification task.
- This paper [23] presents the problems and solutions addressed at the JSALT workshop when using a single microphone for speaker detection in adverse scenarios. The main focus was to tackle a wide range of conditions that go from meetings to wild speech. We describe the research threads we explored and a set of modules that was successful for these scenarios. The ultimate goal was to explore speaker detection; but our first finding was that an effective diarization improves detection, and not having a diarization stage impoverishes the performance. All the different configurations of our research agree on this fact and follow a main backbone that includes diarization as a previous stage. With this backbone, we analyzed the following problems: voice activity detection, how to deal with noisy signals, domain mismatch, how to improve the clustering; and the overall impact of previous stages in the final speaker detection. In this paper, we show partial results for speaker diarizarion to have a better understanding of the problem and we present the final results for speaker detection.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants with Industry

- **Facebook AI Research Grant** (2019, PI: E. Dupoux, 350K€) - Unrestricted Gift - The aim is to help the development of machine learning tools geared towards the psycholinguistic research community.
- **Google Research Award** (2019, PI E. Dunbar, 37K€) - Unrestricted Gift - Develop a first version of a universal synthesizer which can be tuned to specific dialects with sparse data.

8. Partnerships and Cooperations

8.1. Regional Initiatives

Collaboration with the Willow Team:

- co-advising with J. Sivic and I. Laptev of a PhD student: Ronan Riochet.
- construction of a naive physics benchmark (http://www.intphys.com)

Collaboration with the Almanach Team:

- co-advising with B. Sagot a PhD student: Robin Algayres.
- co-advising with B. Sagot a Master student: Charlotte Rochereau

8.2. National Initiatives

8.2.1. ANR

- **ANR-Transatlantic Platform Digging into Data - ACLEW** (2017–2020. 5 countries; Total budget: 1.4M€; coordinating PI : M. Soderstrom; Local PI: A. Cristia; Leader of tools development and co-PI : E. Dupoux) - Constructing tools for the Analysis of Children's Language Experiences Around the World.
- **CNRS Prematuration - BabyCloud**. (2018-2019; coordinating PIs : E. Dupoux and X.-N. Cao; 100€) - Enable the construction of a fully functionnal Baby Logger prototype; perform a market analysis and prepare the launch of a startup.
- **ANR GEOMPHON**. (2018-2021; coordinating PI : E. Dunbar; 299K€) - Study the effects of typologically common properties of linguistic sound systems on speech perception, human learning, and machine learning applied to speech.

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners

- Johns Hopkins University, Baltimore, USA: S. Kudanpur, H. Hermanksy
- RIKEN Institute, Tokyo, Japan: R. Mazuka

8.4. International Research Visitors

8.4.1. Visits of International Scientists

Justine Cassell (CMU, ARP, PRAIRIE Chair starting from Oct 2019)

8.4.2. Visits to International Teams

8.4.2.1. Research Stays Abroad

+ E. Dupoux, Research Scientist, JSALT Workshop, Montreal (July, 2019)

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- E. Dupoux & E. Dunbar: Co-organizers of the INTERSPEECH 2019 Zero Ressource Challenge Session.

9.1.1.2. Member of the Organizing Committees

- Executive committee of SIGMORPHON (Association for Computational Linguistics Special Interest Group, http://www.sigmorphon.org/).
- Executive committee of DARCLE http://www.darcle.org.

9.1.2. Scientific Events Selection

9.1.2.1. Reviewer

Invited editor for international conferences: Interspeech, NIPS, ACL, etc. (around 5-10 papers per conferences, 2 conferences per year)

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Member of the editorial board of: Mathématiques et Sciences Humaines, L'Année Psychologique, Frontiers in Psychology.

9.1.3.2. Reviewer - Reviewing Activities

Invited Reviewer for Frontiers in Psychology, Cognitive Science, Cognition, Transactions in Acoustics Signal Processing and Language, Speech Communication, etc. (around 4 papers per year)

9.1.4. Invited Talks

- Dec/4/2019, E. Dupoux, Invited Speaker, Symposium NeuroDevRob: Developmental AI, Cergy Pontoise
- Jul/15/2019, E. Dupoux, Invited Seminar, Microsoft Research: Inductive Biases and Language Emergence in Communicative Agents, Seattle
- Jun/26/2019, E. Dupoux, Invited Speaker, DARPA MCS Kickoff Meeting: Measuring Intuitive Physics Understanding in Artificial Systems, Washington DC
- May/26/2019, E. Dupoux, Invited Speaker, Facebook CogSci-AI Workshop: Learning speech like infants do, New York
- Mar/14/2019, E. Dupoux, GDR TAL, Artificial models of language acquisition, Paris
- Jan/13/2019, E. Dupoux, Invited Seminar, UTC Compiegne, Developmental AI

9.1.5. Scientific Expertise

E. Dupoux is invited expert for ERC, ANR, and other granting agencies, or tenure committees (around 2 per year).

9.1.6. Research Administration

E. Dupoux is on the Executive committee of the Foundation Cognition, the research programme IRIS-PSL "Sciences des Données et Données des Sciences", the industrial chair Almerys (2016-) and the collective organization DARCLE (http://www.darcle.org).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

E. Dupoux is co-director of the Cognitive Engineering track in the Cognitive Science Master (ENS, EHESS, Paris V).

- Master : E. Dupoux (with B. Sagot, ALMANACH, N. Zeghidour & R. Riad, COML), "Algorithms for speech and language processing", 30h, M2, (MVA), ENS Cachan, France
- Master : E. Dupoux, "Cognitive Engineering", 80h, M2, ITI-PSL, Paris France
- Doctorat : E. Dupoux, "Computational models of cognitive development", 32 h, Séminaire EHESS, Paris France
- Master: E. Dunbar, "Phonology" , 36 h, Master Sciences du Langage, Paris Diderot
- Master: E. Dunbar, "Statistics", 28h, Master Sciences du Langage, Paris Diderot
- Licence 3: E. Dunbar, "Phonology", 36h, Licence Sciences du Langage, Paris Diderot
- Licence 3: E. Dunbar, "Experimental methods", 36h, Licence Sciences du Langage, Paris Diderot

9.2.2. Supervision

- defended PhD: Neil Zeghidour, Learning speech features from raw signals, Feb 2015-13 Mar 2019, co-advised E. Dupoux, N. Usunier (Facebook-CIFRE)
- PhD in progress : Rahma Chaabouni, Language learning in artificial agents, Sept 2017, co-advised E. Dupoux, M. Baroni (Facebook-CIFRE)
- PhD in progress : Ronan Riochet, Learning models of intuitive physics, Sept 2017, co-advised E. Dupoux, I. Laptev, J. Sivic
- PhD in progress : Rachid Riad, "Speech technology for biomarkers in neurodegenerative diseases" , Sept 2018, co-advised E. Dupoux, A.-C. Bachoud-Levi
- PhD in progress : Robin Algayres "Audio word embeddings and word segmentation" , from Oct 2019, co-advised E. Dupoux, B. Sagot
- PhD in progress: Juliette Millet, "Modeling L2 Speech perception", from Sept 2018, advised E. Dunbar Bachoud-Lévi

9.2.3. Juries

E. Dupoux participated in the HDR jury of Jacques Cartier Jul 3, 2019.

9.3. Popularization

E. Dupoux presented the startup project BabyCloud at the CNRS innovatives SHS Salon, May 14-16, Lille.

10. Bibliography

Major publications by the team in recent years

[1] E. DUPOUX.*Cognitive Science in the era of Artificial Intelligence: A roadmap for reverse-engineering the infant language-learner*, in "Cognition", 2018

[2] A. FOURTASSI, E. DUPOUX.*A Rudimentary Lexicon and Semantics Help Bootstrap Phoneme Acquisition*, in "Proceedings of the 18th Conference on Computational Natural Language Learning (CoNLL)", Baltimore, Maryland USA, Association for Computational Linguistics, June 2014, p. 191-200 [*DOI :* 10.3115/v1/W14-1620]

[3] A. FOURTASSI, T. SCHATZ, B. VARADARAJAN, E. DUPOUX.*Exploring the Relative Role of Bottom-up and Top-down Information in Phoneme Learning*, in "Proceedings of the 52nd Annual meeting of the ACL", Baltimore, Maryland, Association for Computational Linguistics, 2014, vol. 2, p. 1-6 [*DOI :* 10.3115/v1/P14-2001]

[4] Y. HOSHEN, R. J. WEISS, K. W. WILSON.*Speech acoustic modeling from raw multichannel waveforms*, in "Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on", IEEE, 2015, p. 4624–4628

[5] T. LINZEN, E. DUPOUX, Y. GOLDBERG.*Assessing the ability of LSTMs to learn syntax-sensitive dependencies*, in "Transactions of the Association for Computational Linguistics", 2016, vol. 4, p. 521-535

[6] T. LINZEN, E. DUPOUX, B. SPECTOR.*Quantificational features in distributional word representations*, in "Proceedings of the Fifth Joint Conference on Lexical and Computational Semantics", 2016, p. 1-11 [*DOI :* 10.18653/v1/S16-2001]

[7] A. Martin, S. Peperkamp, E. Dupoux. *Learning Phonemes with a Proto-lexicon*, in "Cognitive Science", 2013, vol. 37, p. 103-124 [*DOI :* 10.1111/j.1551-6709.2012.01267.x]

[8] S. Mehri, K. Kumar, I. Gulrajani, R. Kumar, S. Jain, J. Sotelo, A. Courville, Y. Bengio. *SampleRNN: An unconditional end-to-end neural audio generation model*, in "arXiv preprint arXiv:1612.07837", 2016

[9] T. N. Sainath, R. J. Weiss, A. Senior, K. W. Wilson, O. Vinyals. *Learning the speech front-end with raw waveform CLDNNs*, in "Sixteenth Annual Conference of the International Speech Communication Association", 2015

[10] T. Schatz, V. Peddinti, F. Bach, A. Jansen, H. Hynek, E. Dupoux. *Evaluating speech features with the Minimal-Pair ABX task: Analysis of the classical MFC/PLP pipeline*, in "INTERSPEECH-2013", Lyon, France, International Speech Communication Association, 2013, p. 1781-1785

[11] R. Thiollière, E. Dunbar, G. Synnaeve, M. Versteegh, E. Dupoux. *A Hybrid Dynamic Time Warping-Deep Neural Network Architecture for Unsupervised Acoustic Modeling*, in "INTERSPEECH-2015", 2015, p. 3179-3183

[12] 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 abs/1609.03499", 2016

Publications of the year

Doctoral Dissertations and Habilitation Theses

[13] N. Zeghidour. *Learning representations of speech from the raw waveform*, PSL Research University, March 2019, https://tel.archives-ouvertes.fr/tel-02278616

Articles in International Peer-Reviewed Journal

[14] M. Bernard, R. Thiollière, A. Saksida, G. Loukatou, E. Larsen, M. C. Johnson, L. Fibla, E. Dupoux, R. Daland, X.-N. Cao, A. Cristia. *WordSeg: Standardizing unsupervised word form segmentation from text*, in "Behavior Research Methods", April 2019 [*DOI :* 10.3758/s13428-019-01223-3], https://hal.archives-ouvertes.fr/hal-02274072

[15] A. Cristia, E. Dupoux, N. Bernstein Ratner, M. Soderstrom. *Segmentability Differences Between Child-Directed and Adult-Directed Speech: A Systematic Test With an Ecologically Valid Corpus*, in "Open Mind", 2019, vol. 3, p. 13-22 [*DOI :* 10.1162/opmi_a_00022], https://hal.archives-ouvertes.fr/hal-02274050

[16] E. Dunbar. *Generative grammar, neural networks, and the implementational mapping problem: Response to Pater*, in "Language", 2019, vol. 95, n° 1, p. e87-e98 [*DOI :* 10.1353/lan.2019.0013], https://hal.archives-ouvertes.fr/hal-02274522

[17] M. Maldonado, E. Dunbar, E. Chemla. *Mouse tracking as a window into decision making*, in "Behavior Research Methods", June 2019, vol. 51, n° 3, p. 1085-1101 [*DOI :* 10.3758/s13428-018-01194-x], https://hal.archives-ouvertes.fr/hal-02274523

International Conferences with Proceedings

[18] R. CHAABOUNI, E. KHARITONOV, A. LAZARIC, E. DUPOUX, M. BARONI.*Word-order biases in deep-agent emergent communication*, in "ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics", Florence, Italy, July 2019, https://arxiv.org/abs/1905.12330 , https://hal.archives-ouvertes.fr/hal-02274157

[19] E. DUNBAR, R. ALGAYRES, J. KARADAYI, M. BERNARD, J. BENJUMEA, X.-N. CAO, L. MISKIC, C. DUGRAIN, L. ONDEL, A. W. BLACK, L. BESACIER, S. SAKTI, E. DUPOUX.*The Zero Resource Speech Challenge 2019: TTS without T*, in "Interspeech 2019 - 20th Annual Conference of the International Speech Communication Association", Graz, Austria, September 2019, https://hal.archives-ouvertes.fr/hal-02274112

[20] A. FOURTASSI, E. DUPOUX.*Phoneme learning is influenced by the taxonomic organization of the semantic referents*, in "Cognitive Science Society", Montreal, Canada, July 2019, https://hal.archives-ouvertes.fr/hal-02274093

[21] E. KHARITONOV, R. CHAABOUNI, D. BOUCHACOURT, M. BARONI.*EGG: a toolkit for research on Emergence of lanGuage in Games*, in "Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations", Hong Kong, China, November 2019, https://arxiv.org/abs/1907.00852 [*DOI :* 10.18653/V1/D19-3010], https://hal.archives-ouvertes.fr/hal-02274229

[22] R. T. MCCOY, T. LINZEN, E. DUNBAR, P. SMOLENSKY.*RNNs Implicitly Implement Tensor Product Representations*, in "ICLR 2019 - International Conference on Learning Representations", New Orleans, United States, May 2019, https://arxiv.org/abs/1812.08718 - Accepted to ICLR 2019, https://hal.archives-ouvertes.fr/hal-02274498

[23] J. MILLET, N. JUROV, E. DUNBAR.*Comparing unsupervised speech learning directly to human performance in speech perception*, in "CogSci 2019 - 41st Annual Meeting of Cognitive Science Society", Montréal, Canada, July 2019, https://hal.archives-ouvertes.fr/hal-02274499

[24] J. MILLET, N. ZEGHIDOUR.*Learning to detect dysarthria from raw speech*, in "ICASSP-2019 - IEEE International Conference on Acoustics, Speech and Signal Processing", Brighton, United Kingdom, May 2019, https://arxiv.org/abs/1811.11101 , https://hal.archives-ouvertes.fr/hal-02274504

Other Publications

[25] R. CHAABOUNI, E. KHARITONOV, E. DUPOUX, M. BARONI.*Anti-efficient encoding in emergent communication*, August 2019, https://arxiv.org/abs/1905.12561 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02274205

[26] P. GARCÍA, J. VILLALBA, H. BREDIN, J. DU, D. CASTAN, A. CRISTIA, L. BULLOCK, L. GUO, K. OKABE, P. S. NIDADAVOLU, S. KATARIA, S. CHEN, L. GALMANT, M. LAVECHIN, L. SUN, M.-P. GILL, B. BEN-YAIR, S. ABDOLI, X. WANG, W. BOUAZIZ, H. TITEUX, E. DUPOUX, K. A. LEE, N. DEHAK.*Speaker detection in the wild: Lessons learned from JSALT 2019*, December 2019, https://arxiv.org/abs/1912.00938 - Submitted to ICASSP 2020, https://hal.archives-ouvertes.fr/hal-02417632

[27] J. KAHN, M. RIVIÈRE, W. ZHENG, E. KHARITONOV, Q. XU, P.-E. MAZARÉ, J. KARADAYI, V. LIPTCHINSKY, R. COLLOBERT, C. FUEGEN, T. LIKHOMANENKO, G. SYNNAEVE, A. JOULIN, A. MOHAMED, E.

DUPOUX.*Libri-Light: A Benchmark for ASR with Limited or No Supervision*, December 2019, https://arxiv.org/abs/1912.07875 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02417621

[28] R. RIOCHET, M. Y. CASTRO, M. BERNARD, A. LERER, R. FERGUS, V. IZARD, E. DUPOUX.*IntPhys: A Benchmark for Visual Intuitive Physics Reasoning*, August 2019, https://arxiv.org/abs/1803.07616 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02274273

[29] C. ROCHEREAU, B. SAGOT, E. DUPOUX.*Modeling German Verb Argument Structures: LSTMs vs. Humans*, December 2019, https://arxiv.org/abs/1912.00239 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02417640

References in notes

[30] D. A. FERRUCCI.*Introduction to "this is watson"*, in "IBM Journal of Research and Development", 2012, vol. 56, n^{o} 3.4, p. 1–1

[31] K. HE, X. ZHANG, S. REN, J. SUN.*Delving deep into rectifiers: Surpassing human-level performance on imagenet classification*, in "Proceedings of the IEEE International Conference on Computer Vision", 2015, p. 1026–1034

[32] J. HERNÁNDEZ-ORALLO, F. MARTÍNEZ-PLUMED, U. SCHMID, M. SIEBERS, D. L. DOWE.*Computer models solving intelligence test problems: Progress and implications*, in "Artificial Intelligence", 2016, vol. 230, p. 74–107

[33] B. M. LAKE, T. D. ULLMAN, J. B. TENENBAUM, S. J. GERSHMAN.*Building machines that learn and think like people*, in "arXiv preprint arXiv:1604.00289", 2016

[34] C. LU, X. TANG.*Surpassing human-level face verification performance on LFW with GaussianFace*, in "arXiv preprint arXiv:1404.3840", 2014

[35] S. T. MUELLER.*A partial implementation of the BICA cognitive decathlon using the Psychology Experiment Building Language (PEBL)*, in "International Journal of Machine Consciousness", 2010, vol. 2, n^{o} 02, p. 273–288

[36] D. SILVER, A. HUANG, C. J. MADDISON, A. GUEZ, L. SIFRE, G. VAN DEN DRIESSCHE, J. SCHRITTWIESER, I. ANTONOGLOU, V. PANNEERSHELVAM, M. LANCTOT, S. DIELEMAN, D. GREWE, J. NHAM, N. KALCHBRENNER, I. SUTSKEVER, T. LILLICRAP, M. LEACH, K. KAVUKCUOGLU, T. GRAEPEL, D. HASSABIS.*Mastering the game of Go with deep neural networks and tree search*, in "Nature", 2016, vol. 529, n^{o} 7587, p. 484–489

[37] I. SUTSKEVER, O. VINYALS, Q. V. LE.*Sequence to sequence learning with neural networks*, in "Advances in neural information processing systems", 2014, p. 3104–3112

[38] A. M. TURING.*Computing machinery and intelligence*, in "Mind", 1950, vol. 59, n^{o} 236, p. 433–460

[39] W. XIONG, J. DROPPO, X. HUANG, F. SEIDE, M. SELTZER, A. STOLCKE, D. YU, G. ZWEIG.*Achieving human parity in conversational speech recognition*, in "arXiv preprint arXiv:1610.05256", 2016

Project-Team COMMEDIA

Computational mathematics for bio-medical applications

IN COLLABORATION WITH: Laboratoire Jacques-Louis Lions (LJLL)

IN PARTNERSHIP WITH:
CNRS

Sorbonne Université

RESEARCH CENTER
Paris

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team COMMEDIA

Creation of the Project-Team: 2019 June 01

Keywords:

Computer Science and Digital Science:

A6.1.1. - Continuous Modeling (PDE, ODE)
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

Other Research Topics and Application Domains:

B2.2.1. - Cardiovascular and respiratory diseases
B2.4.1. - Pharmaco kinetics and dynamics

1. Team, Visitors, External Collaborators

Research Scientists

Miguel Ángel Fernández Varela [Team leader, Inria, Senior Researcher, from Jan 2019, HDR]
Céline Grandmont [Inria, Senior Researcher, from Jan 2019, HDR]
Damiano Lombardi [Inria, Researcher, from Jan 2019]
Marina Vidrascu [Inria, Emeritus, from Jan 2019, HDR]

Faculty Members

Muriel Boulakia [Sorbonne Université, Associate Professor, from Jan 2019]
Olga Mula [Université Paris-Dauphine, Assistant Professor, from Sep 2019, en de´le´gation chez Inria]

Technical Staff

Daniele Carlo Corti [Inria, Engineer, from Jan 2019]

PhD Students

Mocia Agbalessi [Casis, PhD Student, from Apr 2019, granted by CIFRE]
Ludovic Boilevin-Kayl [Inria, PhD Student, from Jan 2019 until Jul 2019]
Felipe Galarce Marin [Inria, PhD Student, from Jan 2019, CORDI-S]
Fannie Gerosa [Inria, PhD Student, from Jan 2019, CORDI-S]
Fabien Raphel [Notocord Systems, PhD Student, from Apr 2019, granted by CIFRE]

Post-Doctoral Fellow

Jean-Jerome Casanova [Inria, Post-Doctoral Fellow, from Jan 2019 until Aug 2019]

Administrative Assistant

Julien Guieu [Inria, Administrative Assistant, from Jun 2019]

2. Overall Objectives

2.1. Overall Objectives

COMMEDIA is a joint project-team of the Inria Research Center of Paris and the Jacques-Louis Lions Laboratory (LJLL) of Sorbonne Université and CNRS (UMR7598). The research activity of COMMEDIA focuses on the numerical simulation of bio-fluid flows in the human body, more specifically, blood flows in the cardiovascular system and air flows in the respiratory system. These simulations are intended to complement available clinical data with the following purpose: help clinicians or bio-engineers to enhance the understanding of physiological phenomena, to improve diagnosis and therapy planning or to optimize medical devices. The main main objectives of COMMEDIA are:

- the development of appropriate mathematical models and efficient numerical methods for the simulations and for the interaction of simulations with measured data;
- the mathematical analysis of these models and numerical techniques;
- the development and validation of scientific computing software which implements these numerical techniques.

A distinctive feature of the mathematical models considered in COMMEDIA is that they often couple different types of partial differential equations (PDEs). This heterogeneous character in the models is a mathematical manifestation of the multi-physics nature of the considered problems.

3. Research Program

3.1. Multi-physics modeling and simulation

The research activity in terms of modeling and simulation (i.e., the so-called forward problem) is driven by two application domains related to the cardiovascular and the respiratory systems.

3.1.1. Cardiovascular hemodynamics

We distinguish between *cardiac hemodynamics* (blood flow inside the four chambers of the heart) and *vascular hemodynamics* (blood flow in the vessels of the body).

Cardiac hemodynamics. The numerical simulation of cardiac hemodynamics presents many difficulties. We can mention, for instance, the large deformation of the cardiac chambers and the complex fluid-structure interaction (FSI) phenomena between blood, the valves and the myocardium. Blood flow can be described by the incompressible Navier-Stokes equations which have to be coupled with a bio-physical model of the myocardium electro-mechanics and a mechanical model of the valves. The coupling between the fluid and the solid media is enforced by kinematic and dynamic coupling conditions, which guarantee the continuity of velocity and stresses across the interface. In spite of the significant advances achieved since the beginning of this century (see, e.g., [61], [69], [60], [63], [53]), the simulation of all the fluid-structure interaction phenomena involved in the heart hemodynamics remains a complex and challenging problem.

Heart valves are definitely a bottleneck of the problem, particularly due to their fast dynamics and the contact phenomena at high pressure-drops. Computational cost is recognized as one of the key difficulties, related to the efficiency of the FSI coupling method and the robustness of the contact algorithm. Furthermore, the numerical discretization of these coupled systems requires to deal with unfitted fluid and solid meshes, which are known to complicate the accuracy and/or the robustness of the numerical approximations (see Section 3.3.2 below).

The ultimate goal of the proposed research activity is the simulation of the complete fluid-structure-contact interaction phenomena involved within the heart. Most of this work will be carried out in close collaboration with the M3DISIM project-team, which has a wide expertise on the modeling, simulation and estimation of myocardium electro-mechanics. We will also consider simplified approaches for cardiac hemodynamics (see, e.g., [34], [48], [51]). The objective is to develop mathematically sound models of reduced valve dynamics with the purpose of enhancing the description of the pressure dynamics right after the opening/closing of the valve (traditional models yield spurious pressure oscillations).

Vascular hemodynamics. The modeling and simulation of vascular hemodynamics in large vessels has been one of the core research topics of some members of COMMEDIA, notably as regards the fluid-structure interaction phenomena. Here we propose to investigate the modeling of pathological scenarios, such as the hemorrhage phenomena in smaller vessels. Modeling of hemorrhage is motivated by the medical constatation that, after a primary vessel wall rupture, secondary vessel wall ruptures are observed. Biologists postulate that the mechanical explanation of this phenomena might be in the change of applied stress due to blood bleeding. We propose to model and simulate the underlying coupled system, blood vessel flow through the external tissue, to estimate the effect of the subsequent stress variation.

3.1.2. Respiratory flows

The motivation of the proposed research activities is to develop a hierarchy of easily parametrizable models allowing to describe and efficiently simulate the physical, mechanical and biological phenomena related to human respiration, namely, ventilation, particle deposition, gas diffusion and coupling with the circulatory system.

Ventilation. The current modeling approaches (either 3D–0D coupled models where the 3D Navier-Stokes equations are solved in truncated geometries of the bronchial tree with appropriate lumped boundary conditions, or 0D–3D coupled models where the lung parenchyma is described by a 3D elastic media irrigated by a simplified bronchial tree) provide satisfactory results in the case of mechanical ventilation or normal breathing. Realistic volume-flow phase portraits can also be simulated in the case of forced expiration (see [36], [45], [66]), but the magnitude of the corresponding pressure is not physiological. The current models must be enriched since they do not yet correctly describe all the physiological phenomena at play. We hence propose to extend the 0D–3D (bronchial tree–parenchyma) model developed in the team, by considering a non-linear, viscoelastic and possibly poro-elastic description of the parenchyma with appropriate boundary conditions that describe ribs and adjacent organs and taking into account an appropriate resistive model.

So far, the motion of the trachea and proximal bronchi has been neglected in the ventilation models (see, e.g., [67]). These features can be critical for the modeling of pathologic phenomena such as sleep apnea and occlusion of the airways. This would be a long-term goal where fluid-structure interaction and the possible contact phenomena will be taken into account, as in the simulation of cardiac hemodynamics (see Section 3.1.1).

Aerosol and gas diffusion. The dynamics of aerosols in the lung have been widely studied from the mathematical modeling standpoint. They can be described by models at different scales: the microscopic one for which each particle is described individually, the mesoscopic (or kinetic) one for which a density of probability is considered, or the macroscopic one where reaction-diffusion equations describing the behavior of the constituant concentration are considered. The objective of COMMEDIA will mainly be to develop the kinetic approach that allows a precise description of the deposition area at controlled computational costs. Part of this study could be done in collaboration with colleagues from the Research Center for Respiratory Diseases at Inserm Tours (UMR1100).

The macroscopic description is also appropriate for the diffusion of gases (oxygen and carbon dioxide) in the bronchial tree (see [62]). Regarding the influence of the carrier gas, if the patient inhales a different mixture of air such as a Helium-Oxygen mixture, the diffusion mechanisms could be modified. In this context, the goal is to evaluate if the cross-diffusion (and thus the carrier gas) modifies the quantities of oxygen diffused. Part of this work will be carried out in collaboration with members of the LJLL and of the MAP5.

As a long term goal, we propose to investigate the coupling of these models to models of diffusion in the blood or to perfusion models of the parenchyma, and thus, have access thanks to numerical simulations to new indices of ventilation efficiency (such as dissolved oxygen levels), depending on the pathology considered or the resting or exercise condition of the patient.

3.2. Simulation with data interaction

The second research axis of COMMEDIA is devoted to the interaction of numerical simulations with measured data. Several research directions related to two specific applications are described below: blood flows and cardiac electrophysiology, for which the mathematical models have been validated against experimental data. This list is not exhaustive and additional problems (related to cardiac and respiratory flows) shall be considered depending on the degree of maturity of the developed models.

3.2.1. Fluid flow reconstruction from medical imaging

A first problem which is currently under study at COMMEDIA is the reconstruction of the flow state from Doppler ultrasound measurements. This is a cheap and largely available imaging modality where the measure can be interpreted as the average on a voxel of the velocity along the direction of the ultrasound beam. The goal is to perform a full-state estimation in a time compatible with a realistic application.

A second problem which is relevant is the flow and wall dynamics reconstruction using 4D-flow MRI. This imaging modality is richer than Doppler ultrasound and provides directly a measure of the 3D velocity field in the voxels. This enables the use of direct estimation methods at a reduced computational cost with respect to the traditional variational data assimilation approaches. Yet, the sensitivity of the results to subsampling and noise is still not well understood.

We also propose to address the issues related to uncertainty quantification. Indeed, measurements are corrupted by noise and the parameters as well as the available data of the system are either hidden or not known exactly (see [59]). This uncertainty makes the estimation difficult and has a large impact on the precision of the reconstruction, to be quantified in order to provide a reliable tool.

3.2.2. Inverse problem in electro-cardiography

The objective of the inverse problem in electro-cardiography is to recover information about the cardiac electrical activity from electrical measurements on the body surface (for instance from electrocardiograms). We propose to investigate approaches based on recent methods for the Cauchy problem reported in [42]. Basically, the idea consists in regularizing the discrete inverse problem using stabilized finite element methods, without the need of integrating a priori knowledge of the solution, only regularity on the exact solution is required.

3.2.3. Safety pharmacology

One of the the most important problems in pharmacology is cardio-toxicity (see [58]). The objective is to predict whether or not a molecule alters in a significant way the normal functioning of the cardiac cells. This problem can be formulated as inferring the impact of a drug on the ionic currents of each cell based on the measured electrical signal (e.g., electrograms from Micro-Electrodes Arrays). The proposed approach in collaboration with two industrial partners (NOTOCORD and Ncardia) consists in combining available realistic data with virtual ones obtained by numerical simulations. These two datasets can be used to construct efficient classifiers and regressors using machine learning tools (see [41]) and hence providing a rapid way to estimate the impact of a molecule on the electrical activity. The methodological aspects of this work are addressed in Section 3.3.3.

3.3. Methodological core

The work described in this section is aimed at investigating fundamental mathematical and numerical problems which arise in the first two research axes.

3.3.1. Mathematical analysis of PDEs

The mathematical analysis of the multi-scale and multi-physics models are a fundamental tool of the simulation chain. Indeed, well-posedness results provide precious insights on the properties of solutions of the systems which can, for instance, guide the design of the numerical methods or help to discriminate between different modeling options.

Fluid-structure interaction. Most of the existing results concern the existence of solutions locally in time or away from contacts. One fundamental problem, related to the modeling and simulation of valve dynamics (see Sections 3.1.1 and 3.3.2), is the question of whether or not the model allows for contact (see [57], [55]). The proposed research activity is aimed at investigating the case of both immersed rigid or elastic structures and explore if the considered model allows for contact and if existence can be proved beyond contact. The question of the choice of the model is crucial and considering different types of fluid (newtonian or non newtonian), structure (smooth or rough, elastic, viscoelastic, poro-elastic), or various interface conditions has an influence on whether the model allows contact or not.

Fluid–structure mixture. The main motivation to study fluid-solid mixtures (i.e., porous media consisting of a skeleton and connecting pores filled with fluid) comes from the modeling of the lung parenchyma and cerebral hemorrhages (see Sections 3.1.1–3.1.2). The Biot model is the most widely used in the literature for the modeling of poro-elastic effects in the arterial wall. Here, we propose to investigate the recent model proposed by the M3DISIM project-team in [47], which allows for nonlinear constitutive behaviors and viscous effects, both in the fluid and the solid. Among the questions which will be addressed, some of them in collaboration with M3DISIM, we mention the justification of the model (or its linearized version) by means of homogenization techniques and its well-posedness.

Fluid–particle interaction. Mathematical analysis studies on the Navier-Stokes-Vlasov system for fluid-particle interaction in aerosols can be found in [38], [39]. We propose to extend these studies to more realistic models which take into account, for instance, changes in the volume of the particles due to humidity.

3.3.2. Numerical methods for multi-physics problems

In this section we describe the main research directions that we propose to explore as regards the numerical approximation of multi-physics problems.

Fluid-structure interaction. The spatial discretization of fluid-structure interaction (FSI) problems generally depends on the amount of solid displacement within the fluid. Problems featuring moderate interface displacements can be successfully simulated using (moving) fitted meshes with an arbitrary Lagrangian-Eulerian (ALE) description of the fluid. This facilitates, in particular, the accurate discretization of the interface conditions. Nevertheless, for problems involving large structural deflections, with solids that might come into contact or that might break up, the ALE formalism becomes cumbersome. A preferred approach in this case is to combine an Eulerian formalism in the fluid with an unfitted mesh discretization, in which the fluid-structure interface deforms independently of a background fluid mesh. In general, traditional unfitted mesh approaches (such as the immersed boundary and the fictitious domain methods [65], [37], [54], [35]) are known to be inaccurate in space. These difficulties have been recently circumvented by a Nitsche-based cut-FEM methodolgy (see [32], [43]). The superior accuracy properties of cut-FEM approaches comes at a price: these methods demand a much more involved computer implementation and require a specific evaluation of the interface intersections.

As regards the time discretization, significant advances have been achieved over the last decade in the development and the analysis of time-splitting schemes that avoid strong coupling (fully implicit treatment of the interface coupling), without compromising stability and accuracy. In the vast majority these studies, the spatial discretization is based on body fitted fluid meshes and the problem of accuracy remains practically open for the coupling with thick-walled structures (see, e.g., [52]). Within the unfitted mesh framework, splitting schemes which avoid strong coupling are much more rare in the literature.

Computational efficiency is a major bottleneck in the numerical simulation of fluid-structure interaction problems with unfitted meshes. The proposed research activity is aimed at addressing these issues. Another fundamental problem that we propose to face is the case of topology changes in the fluid, due to contact or fracture of immersed solids. This challenging problem (fluid-structure-contact-fracture interaction) has major role in many applications (e.g., heart valves repair or replacement, break-up of drug-loaded micro-capsules) but most of the available studies are still merely illustrative. Indeed, besides the numerical issues discussed above, the stability and the accuracy properties of the numerical approximations in such a singular setting are not known.

Fluid–particle interaction and gas diffusion.

Aerosols can be described through mesoscopic equations of kinetic type, which provide a trade-off between model complexity and accuracy. The strongly coupled fluid-particle system involves the incompressible Navier-Stokes equations and the Vlasov equation. The proposed research activity is aimed at investigating the theoretical stability of time-splitting schemes for this system. We also propose to extend these studies to more complex models that take into account the radius growth of the particles due to humidity, and for which stable, accurate and mass conservative schemes have to be developed.

As regards gas diffusion, the mathematical models are generally highly non-linear (see, e.g., [62], [64], [40]). Numerical difficulties arise from these strong non linearities and we propose to develop numerical schemes able to deal with the stiff geometrical terms and that guarantee mass conservation. Moreover, numerical diffusion must be limited in order to correctly capture the time scales and the cross-diffusion effects.

3.3.3. Statistical learning and mathematical modeling interactions

Machine learning and in general statistical learning methods (currently intensively developed and used, see [33]) build a relationship between the system observations and the predictions of the QoI based on the *a posteriori* knowledge of a large amount of data. When dealing with biomedical applications, the available observations are signals (think for instance to images or electro-cardiograms, pressure and Doppler measurements). These data are high dimensional and the number of available individuals to set up precise classification/regression tools could be prohibitively large. To overcome this major problem and still try to exploit the advantages of statistical learning approaches, we try to add, to the a posteriori knowledge of the available data an *a priori* knowledge, based on the mathematical modeling of the system. A large number of numerical simulations is performed in order to explore a set of meaningful scenarios, potentially missing in the dataset. This *in silico* database of virtual experiments is added to the real dataset: the number of individuals is increased and, moreover, this larger dataset can be used to compute semi-empirical functions to reduce the dimension of the observed signals.

Several investigations have to be carried out to systematically set up this framework. First, often there is not a single mathematical model describing a physiological phenomenon, but hierarchies of model of different complexity. Every model is characterized by a model error. How can this be accounted for? Moreover, several statistical estimators can be set up and eventually combined together in order to improve the estimations (see [70]). Other issues have an actual impact and has to be investigated: what is the optimal number of *in silico* experiments to be added? What are the most relevant scenarios to be simulated in relation to the statistical learning approach considered in order to obtain reliable results? In order to answer to these questions, discussions and collaborations with statistics and machine learning groups have to be developed.

3.3.4. Tensor approximation and HPC

Tensor methods have a recent significant development because of their pertinence in providing a compact representation of large, high-dimensional data. Their applications range from applied mathematics and numerical analysis to machine learning and computational physics. Several tensor decompositions and methods are currently available (see [56]). Contrary to matrices, for tensors of order higher or equal to three, there does not exist, in general, a best low rank approximation, the problem being ill posed (see [68]). Two main points will be addressed: (i) The tensor construction and the multi-linear algebra operations involved when solving high-dimensional problems are still sequential in most of the cases. The objective is to design efficient parallel methods for tensor construction and computations; (ii) When solving high-dimensional problems, the tensor is not assigned; instead, it is specified through a set of equations and tensor data. Our goal is to devise numerical methods able to (dynamically) adapt the rank and the discretization (possibly even the tensor format) to respect the chosen error criterion. This could, in turn, improve the efficiency and reduce the computational burden.

These sought improvements could make the definition of parsimonious discretizations for kinetic theory and uncertainty quantification problems (see Section 3.2.1) more efficient and suitable for a HPC paradigm. This work will be carried out in collaboration with Olga Mula (Université Paris-Dauphine) and the ALPINES and MATHERIALS project-teams.

4. Application Domains

4.1. Cardiovascular hemodynamics

The heart is a double pump whose purpose is to deliver blood to the tissue and organs of the body. This function is made possible through the opening and closing of the heart valves. Cardiac diseases generally manifest by affecting the pumping function of the heart. Numerical simulations of cardiac hemodynamics, in normal and pathological conditions, are recognized as a tool of paramount importance for improving the understanding, diagnosis and treatment of cardiac pathologies, and also for the development of implantable devices (see, e.g., [63], [46]). As an example, we can mention the case of cardiac mitral valve regurgitation, one of the most common heart valve diseases. For this pathology, clinical data are known to be insufficient for determining the optimal timing for surgery, the best surgical strategy and the long-term outcome of a surgical repair. Contrary to imaging techniques, numerical simulations provide local information, such as pressure and stresses, which are of fundamental importance for the prediction of the mechanical behavior of native valves and of implantable devices.

4.2. Respiratory flows

Respiration involves the transport of air through the airways from the mouth to the alveoli of the lungs. These units where diffusion of oxygen and carbon dioxide take place, are surrounded by a viscoelastic medium (the parenchyma) consisting of blood vessels and collagen fibers. Air flows due to the displacement of the diaphragm, which drives the pulmonary parenchyma. Accidental inhalations of foreign bodies or pathologies such as asthma, emphysema and fibrosis might prevent the lung of fulfilling its function. Therapies mostly use aerosols (set of small particles, solid or liquid), which must reach the specific areas of the lung targeted for treatment. Understanding the airflow mechanisms within the respiratory network is a fundamental ingredient for predicting the particles motion and their deposition (see, e.g., [44]). Moreover, understanding of the gas diffusion in the lung is also of major importance since the main fonction of this organ is to deliver oxygen to the blood.

4.3. Safety pharmacology

The problem of safety pharmacology can be summarized as follows: given a molecule which is a candidate to become a drug, is its use dangerous due to side effects? Among all the different problems to be addressed, one of the most relevant questions in pharmacology is cardio-toxicity (see [58]). More precisely, the objective is to determine whether or not a molecule alters in a significant way the normal functioning of the cardiac cells. To answer these questions, the CiPA initiative promotes the introduction of novel techniques and their standardisation (see [50]). One of the proposed tests of the CiPA panel is to measure the the electrical activity using Micro-Electrodes Array: these are microchips that record the electrical activity of an ensemble of cells. The task is to infer the impact of a drug on the ionic currents of each cell based on the electrical signal measured (electrograms) and, in perspective, to be able to assess whether a molecule can induce arrhythmia (see [49]).

5. Highlights of the Year

5.1. Highlights of the Year

- Major results have been obtained on the simulation of fluid-structure interaction [1] and of cardiac hemodynamics [8].

6. New Software and Platforms

6.1. FELiScE

Finite Elements for Life SCiences and Engineering problems

KEYWORDS: Finite element modelling - Cardiac Electrophysiology - Cardiovascular and respiratory systems

FUNCTIONAL DESCRIPTION: FELiScE is a finite element code which the M3DISIM and REO project-teams have decided to jointly develop in order to build up on their respective experiences concerning finite element simulations. One specific objective of this code is to provide in a unified software environment all the state-of-the-art tools needed to perform simulations of the complex respiratory and cardiovascular models considered in the two teams – namely involving fluid and solid mechanics, electrophysiology, and the various associated coupling phenomena. FELISCE is written in C++, and may be later released as an opensource library. FELiScE was registered in July 2014 at the Agence pour la Protection des Programmes under the Inter Deposit Digital Number IDDN.FR.001.350015.000.S.P.2014.000.10000.

- Participants: Matteo Aletti, Daniele Carlo Corti, Dominique Chapelle, Miguel Ángel Fernández, Benoit Fabreges, Axel Fourmont, Jean-Frédéric Gerbeau, Fannie Gerosa, Sébastien Gilles, Mikel Landajuela Larma, Damiano Lombardi, Philippe Moireau, Irène Vignon-Clementel and Marina Vidrascu
- Contact: Miguel Ángel Fernández
- URL: http://felisce.gforge.inria.fr

6.2. FELiScE-NS

KEYWORDS: Incompressible flows - Thin-walled solids

FUNCTIONAL DESCRIPTION: FELiScE-NS is a set finite elements solvers for incompressible fluids (fractional-step schemes) and non-linear thin-walled structures (3D shells, and 2D curved beams) developed in the framework of the FELiScE library. FELiSCe-NS was registered in 2018 at the Agence pour la Protection des Programmes Inter Deposit Digital Number IDDN.FR.001.270015.000.S.A.2018.000.31200.

- Participants: Benoit Fabreges, Miguel Ángel Fernández, Axel Fourmont, Jean-Frédéric Gerbeau and Marina Vidrascu
- Contact: Miguel Ángel Fernández

6.3. DCIMaL

KEYWORD: Cardiac Electrophysiology

FUNCTIONAL DESCRIPTION: DCIMaL is a Python and C++ software for safety pharmacology studies and particularly field potentials signals measured with micro-electrode array (MEA). The software includes a solver for field potential simulations and a dictionary of entries corresponding to features which can be extracted from real or simulated potential signals. It also includes an algorithm for drug classification (channel blockade or torsadogenic risk) and a tool for estimating ion channel activity (based on the CMAES library). DCIMaL was registered in 2018 at the Agence pour la Protection des Programmes Inter Deposit Digital Number IDDN.FR.001.270003.000.S.P.2018.000.31230

- Participants: Fabien Raphel, Jean-Frédéric Gerbeau and Damiano Lombardi
- Contact: Damiano Lombardi

7. New Results

7.1. Cardiovascular hemodynamics

Participant: Miguel Ángel Fernández Varela.

Mitral regurgitation is one of the most prevalent valvular heart disease. Proper evaluation of its severity is necessary to choose appropriate treatment. The PISA method, based on Color Doppler echocardiography, is widely used in the clinical setting to estimate various relevant quantities related to the severity of the disease. In [19], the use of a pipeline to quickly generate image-based numerical simulation of intracardiac hemodynamics is investigated. The pipeline capabilities are evaluated on a database of twelve volunteers. Full pre-processing is achieved completely automatically in 55 minutes, on average, with small registration errors compared to the image spatial resolution. This pipeline is then used to study the intracardiac hemodynamics in the presence of diseased mitral valve. A strong variability among the simulated cases, mainly due to the valve geometry and regurgitation specifics, is found. The results from those numerical simulations is used to assess the potential limitations of the PISA method with respect to different MR types. While the PISA method provides reasonable estimates in the case of a simple circular regurgitation, it is shown that unsatisfying estimates are obtained in the case of non-circular leakage. Moreover, it is shown that the choice of high aliasing velocities can lead to difficulties in quantifying MR.

7.2. Respiratory flows

Participant: Céline Grandmont.

In [21], we propose a coupled fluid-kinetic model taking into account the radius growth of aerosol particles due to humidity in the respiratory system. We aim to numerically investigate the impact of hygroscopic effects on the particle behaviour. The air flow is described by the incompressible Navier-Stokes equations, and the aerosol by a Vlasov-type equation involving the airhumidity and temperature, both quantities satisfying a convection-diffusion equation with a sourceterm. Conservations properties are checked and an explicit time-marching scheme is proposed. Two-dimensional numerical simulations in a branched structure show the influence of the particle size variations on the aerosol dynamics.

7.3. Fluid flow reconstruction from medical imaging

Participants: Muriel Boulakia, Miguel Ángel Fernández Varela, Felipe Galarce Marin, Damiano Lombardi, Olga Mula, Colette Voisembert.

In [22], we are interested in designing and analyzing a finite element data assimilation method for laminar steady flow described by the linearized incompressible Navier-Stokes equation. We propose a weakly consistent stabilized finite element method which reconstructs the whole fluid flow from velocity measurements in a subset of the computational domain. Using the stability of the continuous problem in the form of a three balls inequality, we derive quantitative local error estimates for the velocity. Numerical simulations illustrate these convergences properties and we finally apply our method to the flow reconstruction in a blood vessel.

In [29] a state estimation problem is investigated, that consists in reconstructing the blood flow from ultrasound Doppler images. The method proposed is based on a reduced-order technique. Semi-realistic 3D configurations are tested.

7.4. Safety pharmacology

Participants: Damiano Lombardi, Fabien Raphel.

In [31] a greedy method is used to classify molecules action on cardiac myocites. The method is applied to a realistic dataset: the experiments were performed at Ncardia (Netherlands). The experimental dataset is complemented by a synthetic dataset, obtained by simulating experimental meaningful scenarios. The results obtained are encouraging.

7.5. Mathematical analysis of PDEs

Participants: Muriel Boulakia, Jean-Jerome Casanova, Céline Grandmont.

In [23], we consider a reaction-diffusion equation where the reaction term is given by a cubic function and we are interested in the numerical reconstruction of the time-independent part of the source term from measurements of the solution. For this identification problem, we present an iterative algorithm based on Carleman estimates which consists of minimizing at each iteration strongly convex cost functionals. Despite the nonlinear nature of the problem, we prove that our algorithm globally converges and the convergence speed evaluated in weighted norm is linear. In the last part of the paper, we illustrate the effectiveness of our algorithm with several numerical reconstructions in dimension one or two.

In [25] a coupled system of pdes modelling the interaction between a two–dimensional incompressible viscous fluid and a one–dimensional elastic beam located on the upper part of the fluid domain boundary is considered. A good functional framework to define weak solutions in case of contact between the elastic beam and the bottom of the fluid cavity is designed.It is then prove that such solutions exist globally in time regardless a possible contact by approximating the beam equation by a damped beam and letting this additional viscosity vanishes.

7.6. Numerical methods for multi-physics problems

Participants: Miguel Ángel Fernández Varela, Fannie Gerosa.

In [24] we introduce a mixed dimensional Stokes-Darcy coupling where a d-dimensional Stokes' flow is coupled to a Darcy model on the $d-1$ dimensional boundary of the domain. The porous layer introduces tangential creeping flow along the boundary and allows for the modelling of boundary flow due to surface roughness. This leads to a new model of flow in fracture networks with reservoirs in an impenetrable bulk matrix. Exploiting this modelling capability, we then formulate a fluid-structure interaction method with contact, where the porous layer allows for mechanically consistent contact and release. Physical seepage in the contact zone due to rough surfaces is modelled by the porous layer. Some numerical examples are reported, both on the Stokes-Darcy coupling alone and on the fluid-structure interaction with contact in the porous boundary layer.

Unfitted mesh finite element approximations of immersed incompressible fluid-structure interaction problems which efficiently avoid strong coupling without compromising stability and accuracy are rare in the literature. Moreover, most of the existing approaches introduce additional unknowns or are limited by penalty terms which yield ill conditioning issues. In [28], we introduce a new unfitted mesh semi-implicit coupling scheme which avoids these issues. To this purpose, we provide a consistent generalization of the projection based semi-implicit coupling paradigm of [Int. J. Num. Meth. Engrg.,69(4):794-821, 2007] to the unfitted mesh Nitsche-XFEM framework.

7.7. Statistical learning and mathematical modeling interactions

Participants: Damiano Lombardi, Fabien Raphel.

In [30] a greedy dimension reduction method is proposed to deal with classification problems. The method proposed can be seen as a goal oriented dimension reduction method. Elements of a Stiefel manifold (whose dimension is not fixed a priori) are computed in such a way that the classification score is maximised. Several examples are proposed to illustrate the method features and to highlight its differences with classical reduction methods used in classification.

7.8. Tensor approximation and HPC

Participant: Damiano Lombardi.

In [26] a hierarchical adaptive piece-wise tensor decomposition is proposed to approximate high-dimensional functions. Neither the subtensor partitioning nor the rank of the approximation in each of the partitions are fixed a priori. Instead, they are computed to fulfill a prescribed accuracy. Two main contributions are proposed. A greedy error distribution scheme, that allows to adaptively construct the approximation in each of the partitions and a hierarchical tree algorithm that optimise the subtensor partitioning to minimise the storage. Several example on challenging functions are proposed.

7.9. Miscellaneous

Participants: Damiano Lombardi, Olga Mula.

In [20] an approximated formulation of the multi-marginal optimal transport problem (Kantorovich formulation) is proposed. In the formulation, called MCOT, the constraints on the marginal densities are replaced by moments of the densities. This formulation allows to deal simply with a wide spectrum of high-dimensional multi-marginal problems, with non-standard (martingale) constraints.

In [27] a reduced-order modeling framework is proposed, in which a set of model instances is part of a metric space. The introduction of the exponential and logarithmic maps (Riemannian geometry) makes it possible to reduce in an effective way solutions that are classically challenging for standard model reduction methods. Some examples on 1D hyperbolic equations and Wasserstein distance are proposed.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Notocord Systems

Participants: Damiano Lombardi, Fabien Raphel.

This work is devoted to the investigation on new approaches and efficient algorithms in the context of safety pharmacology and the analysis of biological signals.

8.1.2. Casis

Participants: Mocia Agbalessi, Miguel Ángel Fernández Varela, Damiano Lombardi.

This work is devoted to the combination of 4D-MRI data and fluid-structure interaction models of blood flow to asses indicators of aneurysm rupture.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

9.1.1.1. ANR Project “IFSMACS”

Participants: Muriel Boulakia, Céline Grandmont [local coordinator].

Period: 2015-2019.
The objective of this project, coordinated by Takéo Takahashi (Inria Nancy Grand-Est), is the mathematical analysis of systems involving structures immersed in a fluid. This includes the asymptotic analysis, the study of the controllability and stabilization of fluid-structure interaction systems, the understanding of the motion of self-propelled structures and the analysis and development of numerical methods to simulate fluid-structure systems.

9.1.1.2. ANR Project “ADAPT”

Participants: Maria Fuente-Ruiz, Damiano Lombardi [coordinator], Olga Mula.

Period: 2018-2022.

Adaptive Dynamical Approximations by Parallel Tensor methods. The main goal of the ANR is to investigate the numerical approximation of the solution of high-dimensional problems. In particular, the applications that motivate this study are the Uncertainty Quantification and the Kinetic theory. The main objective is to construct in an adaptive way parsimonious discretisations starting from arbitrarily chosen separated discretisations.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

- Céline Grandmont
  - Co-organizer of Inria-LJLL meeting in scientific computing
- Damiano Lombardi
  - Co-organisation (with Sanjay Pant, Swansea University) of two mini-symposia: Machine learning methods in biomedical engineering and Uncertainty Quantification and Bayesian inference in biomedical applications at CMBE 2019, June 2019, Sendai, Japan
  - Co-organizer of Inria-LJLL meeting in scientific computing
- Olga Mula
  - Co-organizer of Summer School on Sparity for Physics, Signals and Learning, at Inria Paris, June 2019
  - Co-organiser of the GdT Stat-Num at Paris Dauphine

10.1.1.1. Member of the Organizing Committees

- Céline Grandmont
  - Member of the local organizing comittee of WCCM 2020

10.1.2. Scientific Events: Selection

- Muriel Boulakia
  - Member of the scientific committee of LIA COPDESC and Lions-Magenes Days, LJJL, November 2019, Paris
- Miguel Ángel Fernández Varela
  - Member of the scientific committee of XDMS 2019, July 2019, Lugano

10.1.2.1. Member of the Editorial Boards

- Céline Grandmont
  - Member of the editorial board of Mathematical Modelling of Natural Phenomena
  - Member of the editorial board of Journal of Mathematical Fluid Mechanics (since November 2019)
- Olga Mula
  - Member of the editorial board of Calcolo (since October 2019)

10.1.2.2. Reviewer - Reviewing Activities

- Muriel Boulakia
  - Expert for « Appel à projets générique », ANR 2019

10.1.3. Research Administration

- Miguel Ángel Fernández Varela
  - Head of Science, Inria Paris (from September 2019)

  - Deputy Head of Science, Inria Paris (until August 2019)
  - Member of the Inria Evaluation Committee
- Céline Grandmont
  - Member of the Inria Evaluation Committee
  - Member of the Inria Parity Committee
  - Member of the scientific board of the EDMH, Paris Saclay

10.1.4. Conferences

- Muriel Boulakia
  - Seminar, Modeling and scientific computation, November 2019, LAGA, Univ. Paris 13
  - Invited speaker, Control and stabilization issues for PDE, September 2019, Toulouse
  - Contributed talk in minisymposium, Applied Inverse Problems Conference, July 2019, Grenoble
  - Seminar, Numerical Analysis and PDE, Université Lille, March 2019, Lille
  - Invited speaker, Workshop on fluid-structure interaction, March 2019, Politecnico di Milano, Italy
- Miguel Ángel Fernández Varela
  - Seminar, Laboratoire Mathématiques Jean Leray, Univ Nates, Nantes
  - Contributed talk in minisymposium, ICIAM, July 2019, Valencia, Spain
  - Invited speaker, iHEART conference on Modelling the Cardiac Function, July 2019, Varese, Italy
- Felipe Galarce Marin
  - Seminar at Inria Chile, July 2019, Santiago, Chile
  - Semina at Universidad Técnica Federico Santa Maria, July 2019, Valparaiso, Chile
  - Seminar at Pontificia Universidad Católica de Valparaíso, July 2019, Valparíso, Chile
  - Contributed talk in minisymposium, CSMA, May 2019, Giens
  - Contributed talk in minisymposium, Enumath, September/October 2019, Egmond aan Zee, The Netherlands
- Fannie Gerosa
  - Contributed talk in minisymposium, 9th biennial congress of SMAI, May 2019, Guidel Plages, France
  - Contributed talk in minisymposium, 15th US National Congress on Computational Mechanics, July/August 2019, Austin, Texas
- Céline Grandmont
  - Plenary speaker, International Meeting on Applied Mathematics & Evolution, La Rochelle, April 2019
  - Keynote speaker, PDE 2019: Partial Differential Equations in Fluids and Solids, Berlin, Sep. 2019
  - Invited Speaker, International Conference "Multiscale Modeling in Fluid Mechanics and Fluid-Structure Interaction" , Oct. 2019, Vilnius
  - Seminar, ULB, Bruxelles, Nov. 2019
  - Seminar, CMAP, Ecole Polytechnique, Dec. 2019
- Damiano Lombardi
  - Contributed talk in minisymposium, ICIAM, July 2019, Valencia, Spain

  - Contributed talk at the Lions-Magenes days, November 2019, Paris.
  - Contributed talk at Mortech, November 2019, Paris.
- Olga Mula
  - Contributed talk at Enumath, September/October 2019, Egmond aan Zee, The Netherlands
  - Contributed talk at INdaM workshop on kinetic equations, November 2019, Rome, Italy
  - Contributed talk at Mortech, November 2019, Paris.
- Fabien Raphel
  - Contributed poster at The Safety Pharmacology Society (SPS) annual meeting, September 2019, Barcelona, Spain
  - Contributed talk at The Safety Pharmacology Society (SPS) webinar, November 2019, Paris
  - Invited talk at Math jobs forum University of Nantes, January 2019, Nantes

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence:

- Muriel Boulakia
  - Projects on differential equations, 34h, L3, Polytech Sorbonne, Sorbonne Université
  - Nonlinear systems and optimization, 35h, L3, Polytech Sorbonne, Sorbonne Université
  - Numerical approximation of functions, 36h, L3, Sorbonne Université
- Fannie Gerosa
  - Function series, integrals and application to differential equation, 38h, L3, Sorbonne Université
- Damiano Lombardi
  - Cours/TD analysis and scientific computing, 32 h, ENPC, L3.
  - Model reduction in large dimensions, 9h, ENPC, L3.
  - TD, Function approximation, Sorbonne University , 18h, L3

Master:

- Muriel Boulakia
  - Preparatory course for teaching admission examination “Agrégation”, 40h, M2, Sorbonne Université
- Miguel Ángel Fernández Varela
  - Modeling and numerical methods for hemodynamics, 30h, M2, Sorbonne Université
- Damiano Lombardi
  - Mini-cours Modeling of cardiac electrophysiology, 3h, M2, Ecole des Mines Paristech

Others:

- Céline Grandmont
  - Invited Lecturer, Hausdorff School on “Modeling and Analysis of Evolutionary Problems in Materials Science” , 6h, Bonn, Sept. 2019

– Invited Lecturer, Institut d'Etudes Scientifiques de Cargèse, Summer school : *Why modelling the pulmonary ventilation? Music, exercise and medicine*, 1h30, November 2019

10.2.2. Supervision

PhD: Ludovic Boilevin-Kayl, Modeling of cardiac implantable devices, Sorbonne Université, defended on July 10, 2019. Supervisors: J.-F. Gerbeau & M.A. Fernández Varela

PhD: Alexandre This, Fusion data/simulation for the assessment of mitral regurgitation, Sorbonne Université, defended on May 28, 2019. Supervisors: O. Bonnefous, M.A. Fernández Varela, J.-F. Gerbeau & H. Morales

PhD in progress: Mocia Agbalessi, Modeling and patient specific fluid-structure interaction simulations of aortic pathological configurations. Since April, 2019, Supervisors: M.A. Fernández Varela & D. Lombardi

PhD in progress: Felipe Galarce, Enhancing hemodynamics measurements with mathematical modeling, since December 2017. Supervisors: J.-F. Gerbeau, D. Lombardi & O. Mula

PhD in progress: Fannie Gerosa, Immersed boundary methods for fluid-structure interaction with topological changes, since January 2018. Supervisor: M.A. Fernández Varela

PhD in progress: Fabien Raphel, Mathematical modeling and learning of biomedical signals for safety pharmacology. Since April 2019. Supervisors: J.-F. Gerbeau & D. Lombardi

10.2.3. Juries

- Miguel Ángel Fernández Varela
  - Hiring committees: Inria Paris (CRCN, Vice-président), Université de Franche-Comté (MdC)
  - PHD committee: P. Quemar, Université Paris-Nord (reviewer), D. Nolte, Groningen University, The Netherlands (reviewer), R.J. Rebolledo Cormack, Universidad de Concepción, Chile (reviewer)
- Céline Grandmont
  - Member of the "agrégation" jury in mathematics
  - Hiring committees: Inria CRCN, Inria DR2
  - PHD committee: N. Lauzeral, Ecole centrale de Nantes (reviewer), F. Vergnet, Paris Sud Univ (member), A. Dekkers, Centrale Suppelec (reviewer)
  - HDR committee: A. Decoene, Paris-Sud Univ (reviewer)

10.3. Popularization

10.3.1. Interventions

- Céline Grandmont
  - Meeting with junior high school students, Collège Iqbal Masih, Saint-Denis. Interview: https://www.youtube.com/watch?v=eJVaVB7z8tU&t=12s

11. Bibliography

Major publications by the team in recent years

[1] L. Boilevin-Kayl, M. A. Fernández, J.-F. Gerbeau.*A loosely coupled scheme for fictitious domain approximations of fluid-structure interaction problems with immersed thin-walled structures*, in "SIAM Journal on Scientific Computing", February 2019, vol. 41, n^o 2, p. 351-374 [*DOI :* 10.1137/18M1192779], https://hal.inria.fr/hal-01811290

[2] M. BOULAKIA.*Quantification of the unique continuation property for the nonstationary Stokes problem*, in "Mathematical Control and Related Fields", March 2016, https://hal.inria.fr/hal-01094490

[3] M. BOULAKIA, S. GUERRERO, T. TAKAHASHI.*Well-posedness for the coupling between a viscous incompressible fluid and an elastic structure*, in "Nonlinearity", 2019, vol. 32, p. 3548-3592 [*DOI :* 10.1088/1361-6544/AB128C], https://hal.inria.fr/hal-01939464

[4] C. GRANDMONT, M. HILLAIRET.*Existence of global strong solutions to a beam-fluid interaction system*, in "Archive for Rational Mechanics and Analysis", 2016 [*DOI :* 10.1007/S00205-015-0954-Y], https://hal.inria.fr/hal-01138736

[5] M. LANDAJUELA, M. VIDRASCU, D. CHAPELLE, M. A. FERNÁNDEZ.*Coupling schemes for the FSI forward prediction challenge: comparative study and validation*, in "International Journal for Numerical Methods in Biomedical Engineering", 2017, vol. 33, n^{o} 4, e02813 [*DOI :* 10.1002/CNM.2813], https://hal.inria.fr/hal-01239931

[6] D. LOMBARDI, S. PANT.*A non-parametric k-nearest neighbor entropy estimator*, in "Physical Reviev E", January 2016 [*DOI :* 10.1103/PHYSREVE.93.013310], https://hal.inria.fr/hal-01272527

[7] N. POZIN, S. MONTESANTOS, I. KATZ, M. PICHELIN, I. VIGNON-CLEMENTEL, C. GRANDMONT.*Predicted airway obstruction distribution based on dynamical lung ventilation data: a coupled modeling-machine learning methodology*, in "International Journal for Numerical Methods in Biomedical Engineering", May 2018, vol. 34, n^{o} 9 [*DOI :* 10.1002/CNM.3108], https://hal.archives-ouvertes.fr/hal-01568065

[8] A. THIS, L. BOILEVIN-KAYL, M. A. FERNÁNDEZ, J.-F. GERBEAU.*Augmented Resistive Immersed Surfaces valve model for the simulation of cardiac hemodynamics with isovolumetric phases*, in "International Journal for Numerical Methods in Biomedical Engineering", May 2019, forthcoming [*DOI :* 10.1002/CNM.3223], https://hal.inria.fr/hal-01944798

[9] E. TIXIER, F. RAPHEL, D. LOMBARDI, J.-F. GERBEAU.*Composite biomarkers derived from Micro-Electrode Array measurements and computer simulations improve the classification of drug-induced channel block*, in "Frontiers in Physiology", 2018, vol. 8, n^{o} 1096, p. 1-30 [*DOI :* 10.3389/FPHYS.2017.01096], https://hal.archives-ouvertes.fr/hal-01570819

Publications of the year

Doctoral Dissertations and Habilitation Theses

[10] L. BOILEVIN-KAYL.*Modeling and numerical simulation of implantable cardiovascular devices*, Sorbonne Université, July 2019, https://hal.inria.fr/tel-02217259

[11] A. THIS.*Image/Model Fusion for the Quantification of Mitral Regurgitation Severity*, Sorbonne Université, May 2019, https://hal.inria.fr/tel-02176167

Articles in International Peer-Reviewed Journal

[12] L. BOILEVIN-KAYL, M. A. FERNÁNDEZ, J.-F. GERBEAU.*A loosely coupled scheme for fictitious domain approximations of fluid-structure interaction problems with immersed thin-walled structures*, in "SIAM Journal on Scientific Computing", February 2019, vol. 41, n^{o} 2, p. 351-374 [*DOI :* 10.1137/18M1192779], https://hal.inria.fr/hal-01811290

[13] L. BOILEVIN-KAYL, M. A. FERNÁNDEZ, J.-F. GERBEAU.*Numerical methods for immersed FSI with thin-walled structures*, in "Computers and Fluids", January 2019, vol. 179, p. 744-763 [*DOI :* 10.1016/J.COMPFLUID.2018.05.024], https://hal.inria.fr/hal-01704575

[14] M. BOULAKIA, S. GUERRERO, T. TAKAHASHI.*Well-posedness for the coupling between a viscous incompressible fluid and an elastic structure*, in "Nonlinearity", 2019, vol. 32, p. 3548-3592 [*DOI :* 10.1088/1361-6544/AB128C], https://hal.inria.fr/hal-01939464

[15] E. BURMAN, M. A. FERNÁNDEZ, S. FREI.*A Nitsche-based formulation for fluid-structure interactions with contact*, in "Modelisation Mathématique et Analyse Numérique", October 2019, forthcoming, https://hal.inria.fr/hal-01784841

[16] M. A. FERNÁNDEZ, M. LANDAJUELA.*Splitting schemes and unfitted mesh methods for the coupling of an incompressible fluid with a thin-walled structure*, in "IMA Journal of Numerical Analysis", January 2019 [*DOI :* 10.1093/IMANUM/DRY098], https://hal.inria.fr/hal-01309462

[17] C. GRANDMONT, M. HILLAIRET, J. LEQUEURRE.*Existence of local strong solutions to fluid-beam and fluid-rod interaction systems*, in "Annales de l'Institut Henri Poincaré (C) Non Linear Analysis", July 2019, vol. 36, n^{o} 4, p. 1105-1149 [*DOI :* 10.1016/J.ANIHPC.2018.10.006], https://hal.inria.fr/hal-01567661

[18] A. THIS, L. BOILEVIN-KAYL, M. A. FERNÁNDEZ, J.-F. GERBEAU.*Augmented Resistive Immersed Surfaces valve model for the simulation of cardiac hemodynamics with isovolumetric phases*, in "International Journal for Numerical Methods in Biomedical Engineering", May 2019, forthcoming [*DOI :* 10.1002/CNM.3223], https://hal.inria.fr/hal-01944798

[19] A. THIS, H. G. MORALES, O. BONNEFOUS, M. A. FERNÁNDEZ, J.-F. GERBEAU.*A pipeline for image based intracardiac CFD modeling and application to the evaluation of the PISA method*, in "Computer Methods in Applied Mechanics and Engineering", 2019, vol. 358, n^{o} 1 [*DOI :* 10.1016/J.CMA.2019.112627], https://hal.archives-ouvertes.fr/hal-02142416

Other Publications

[20] A. ALFONSI, R. COYAUD, V. EHRLACHER, D. LOMBARDI.*Approximation of Optimal Transport problems with marginal moments constraints*, May 2019, https://arxiv.org/abs/1905.05663 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02128374

[21] L. BOUDIN, C. GRANDMONT, B. GREC, S. MARTIN, A. MECHERBET, F. NOËL.*Fluid-kinetic modelling for respiratory aerosols with variable size and temperature*, April 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02092574

[22] M. BOULAKIA, E. BURMAN, M. A. FERNÁNDEZ, C. VOISEMBERT.*Data assimilation finite element method for the linearized Navier-Stokes equations in the low Reynolds regime*, October 2019, working paper or preprint, https://hal.inria.fr/hal-02318504

[23] M. BOULAKIA, M. DE BUHAN, E. SCHWINDT.*Numerical reconstruction based on Carleman estimates of a source term in a reaction-diffusion equation.* *, July 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02185889

[24] E. BURMAN, M. A. FERNÁNDEZ, S. FREI, F. GEROSA.*3D-2D Stokes-Darcy coupling for the modelling of seepage with an application to fluid-structure interaction with contact*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02417042

[25] J.-J. CASANOVA, C. GRANDMONT, M. HILLAIRET.*On an existence theory for a fluid-beam problem encompassing possible contacts*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02396915

[26] V. EHRLACHER, L. GRIGORI, D. LOMBARDI, H. SONG.*Adaptive hierarchical subtensor partitioning for tensor compression*, September 2019, working paper or preprint, https://hal.inria.fr/hal-02284456

[27] V. EHRLACHER, D. LOMBARDI, O. MULA, F.-X. VIALARD.*Nonlinear model reduction on metric spaces. Application to one-dimensional conservative PDEs in Wasserstein spaces*, September 2019, https://arxiv.org/abs/1909.06626 - working paper or preprint, https://hal.inria.fr/hal-02290431

[28] M. A. FERNÁNDEZ, F. GEROSA.*An unfitted mesh semi-implicit coupling scheme for fluid-structure interaction with immersed solids*, September 2019, working paper or preprint, https://hal.inria.fr/hal-02288723

[29] F. GALARCE, J.-F. GERBEAU, D. LOMBARDI, O. MULA.*State estimation with nonlinear reduced models. Application to the reconstruction of blood flows with Doppler ultrasound images*, December 2019, https://arxiv.org/abs/1904.13367 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02403686

[30] D. LOMBARDI, F. RAPHEL.*A greedy dimension reduction method for classification problems*, September 2019, working paper or preprint, https://hal.inria.fr/hal-02280502

[31] F. RAPHEL, T. DE KORTE, D. LOMBARDI, S. BRAAM, J.-F. GERBEAU.*A greedy classifier optimisation strategy to assess ion channel blocking activity and pro-arrhythmia in hiPSC-cardiomyocytes*, September 2019, working paper or preprint, https://hal.inria.fr/hal-02276945

References in notes

[32] F. ALAUZET, B. FABRÈGES, M. A. FERNÁNDEZ, M. LANDAJUELA.*Nitsche-XFEM for the coupling of an incompressible fluid with immersed thin-walled structures*, in "Comput. Methods Appl. Mech. Engrg.", 2016, vol. 301, p. 300–335

[33] E. ALPAYDIN.*Introduction to machine learning*, MIT press, 2009

[34] M. ASTORINO, J. HAMERS, S. C. SHADDEN, J.-F. GERBEAU.*A robust and efficient valve model based on resistive immersed surfaces*, in "Int. J. Numer. Meth. Biomed. Engng.", 2012, vol. 28, n^{o} 9, p. 937–959

[35] F. BAAIJENS.*A fictitious domain/mortar element method for fluid-structure interaction*, in "Int. Jour. Num. Meth. Fluids", 2001, vol. 35, p. 743-761

[36] L. BAFFICO, C. GRANDMONT, B. MAURY.*Multiscale modeling of the respiratory tract*, in "Math. Models Methods Appl. Sci.", 2010, vol. 20, n^{o} 1, p. 59–93

[37] D. BOFFI, N. CAVALLINI, L. GASTALDI.*Finite element approach to immersed boundary method with different fluid and solid densities*, in "Math. Models Methods Appl. Sci.", 2011, vol. 21, n^{o} 12, p. 2523–2550

[38] L. BOUDIN, L. DESVILLETTES, C. GRANDMONT, A. MOUSSA.*Global existence of solutions for the coupled Vlasov and Navier-Stokes equations*, in "Differential Integral Equations", 2009, vol. 22, n^{o} 11-12, p. 1247–1271

[39] L. BOUDIN, C. GRANDMONT, A. MOUSSA.*Global existence of solutions to the incompressible Navier-Stokes-Vlasov equations in a time-dependent domain*, in "J. Differential Equations", 2017, vol. 262, n^{o} 3, p. 1317–1340, https://doi.org/10.1016/j.jde.2016.10.012

[40] L. BOUDIN, D. GÖTZ, B. GREC.*Diffusion models of multicomponent mixtures in the lung*, in "CEMRACS 2009: Mathematical modelling in medicine", ESAIM Proc., EDP Sci., Les Ulis, 2010, vol. 30, p. 90–103

[41] L. BREIMAN.*Classification and regression trees*, Routledge, 2017

[42] E. BURMAN.*Stabilized finite element methods for nonsymmetric, noncoercive, and ill-posed problems. Part I: Elliptic equations*, in "SIAM J. Sci. Comput.", 2013, vol. 35, n^{o} 6, p. 2752-2780

[43] E. BURMAN, M. A. FERNÁNDEZ.*An unfitted Nitsche method for incompressible fluid-structure interaction using overlapping meshes*, in "Comput. Methods Appl. Mech. Engrg.", 2014, vol. 279, p. 497–514

[44] R. K. CALAY, J. KURUJAREON, A. E. HOLDO.*Numerical simulation of respiratory flow patterns within human lung*, in "Respir. Physiol. Neurobiol.", 2002, vol. 130, n^{o} 2, p. 201–221

[45] P. CAZEAUX, C. GRANDMONT.*Homogenization of a multiscale viscoelastic model with nonlocal damping, application to the human lungs*, in "Math. Models Methods Appl. Sci.", 2015, vol. 25, n^{o} 6, p. 1125–1177

[46] K. B. CHANDRAN.*Role of Computational Simulations in Heart Valve Dynamics and Design of Valvular Prostheses*, in "Cardiovasc. Eng. Technol.", 2010, vol. 1, n^{o} 1, p. 18–38

[47] D. CHAPELLE, P. MOIREAU.*General coupling of porous flows and hyperelastic formulations—From thermodynamics principles to energy balance and compatible time schemes*, in "Eur. J. Mech. B Fluids.", 2014, vol. 46, p. 82–96

[48] C. CHNAFA, S. MENDEZ, F. NICOUD.*Image-Based Simulations Show Important Flow Fluctuations in a Normal Left Ventricle: What Could be the Implications?*, in "Ann. Biomed. Eng.", 2016, vol. 44, n^{o} 11, p. 3346–3358

[49] T. COLATSKY, B. FERMINI, G. GINTANT, J. B. PIERSON, P. SAGER, Y. SEKINO, D. G. STRAUSS, N. STOCKBRIDGE.*The comprehensive in vitro proarrhythmia assay (CiPA) initiative—update on progress*, in "J. Pharmacol. Toxicol. Methods", 2016, vol. 81, p. 15–20

[50] W. J. CRUMB, J. VICENTE, L. JOHANNESEN, D. G. STRAUSS.*An evaluation of 30 clinical drugs against the comprehensive in vitro proarrhythmia assay (CiPA) proposed ion channel panel*, in "J. Pharmacol. Toxicol. Methods", 2016, vol. 81, p. 251–262

[51] M. FEDELE, E. FAGGIANO, L. DEDÈ, A. QUARTERONI.*A patient-specific aortic valve model based on moving resistive immersed implicit surfaces*, in "Biomech. Model. Mechanobiol.", 2017, vol. 16, n^{o} 5, p. 1779–1803

[52] M. A. FERNÁNDEZ, J. MULLAERT.*Convergence and error analysis for a class of splitting schemes in incompressible fluid-structure interaction*, in "IMA J. Numer. Anal.", 2016, vol. 36, n^{o} 4, p. 1748–1782

[53] H. GAO, L. FENG, N. QI, C. BERRY, B. GRIFFITH, X. LUO.*A coupled mitral valve-left ventricle model with fluid-structure interaction*, in "Med. Eng. Phys.", 09 2017, vol. 47, p. 128–136

[54] R. GLOWINSKI, T. PAN, T. HESLA, D. JOSEPH.*A distributed Lagrange mutiplier/fictitious domain method for particulate flows*, in "Int. J. of Multiphase Flow", 1999, vol. 25, p. 755-794

[55] C. GRANDMONT, M. HILLAIRET.*Existence of global strong solutions to a beam-fluid interaction system*, in "Arch. Ration. Mech. Anal.", 2016, vol. 220, n^{o} 3, p. 1283–1333

[56] L. GRASEDYCK, D. KRESSNER, C. TOBLER.*A literature survey of low-rank tensor approximation techniques*, in "GAMM-Mitt.", 2013, vol. 36, n^{o} 1, p. 53–78

[57] M. HILLAIRET.*Lack of collision between solid bodies in a 2D incompressible viscous flow*, in "Comm. Partial Differential Equations", 2007, vol. 32, n^{o} 7-9, p. 1345–1371

[58] H. M. HIMMEL.*Drug-induced functional cardiotoxicity screening in stem cell-derived human and mouse cardiomyocytes: effects of reference compounds*, in "J. Pharmacol. Toxicol. Methods", 2013, vol. 68, n^{o} 1, p. 97–111

[59] J. KAIPIO, E. SOMERSALO.*Statistical and computational inverse problems*, Applied Mathematical Sciences, Springer-Verlag, New York, 2005, vol. 160

[60] D. KAMENSKY, M.-C. HSU, Y. YU, J. A. EVANS, M. S. SACKS, T. J. R. HUGHES.*Immersogeometric cardiovascular fluid-structure interaction analysis with divergence-conforming B-splines*, in "Comput. Methods Appl. Mech. Engrg.", 2017, vol. 314, p. 408–472

[61] M. C. LAI, C. S. PESKIN.*An immersed boundary method with formal second-order accuracy and reduced numerical viscosity*, in "J. Comp. Phys.", 2000, vol. 160, n^{o} 2, p. 705–719

[62] S. MARTIN, B. MAURY.*Modeling of the oxygen transfer in the respiratory process*, in "ESAIM Math. Model. Numer. Anal.", 2013, vol. 47, n^{o} 4, p. 935–960

[63] R. MITTAL, J. H. SEO, V. VEDULA, Y. J. CHOI, H. LIU, H. H. HUANG, S. JAIN, L. YOUNES, T. ABRAHAM, R. T. GEORGE.*Computational modeling of cardiac hemodynamics: current status and future outlook*, in "J. Comput. Phys.", 2016, vol. 305, p. 1065–1082

[64] J. M. OAKES, S. C. SHADDEN, C. GRANDMONT, I. E. VIGNON-CLEMENTEL.*Aerosol transport throughout inspiration and expiration in the pulmonary airways*, in "Int. J. Numer. Methods Biomed. Eng.", 2017, vol. 33, n^{o} 9

[65] C. S. PESKIN.*The immersed boundary method*, in "Acta Numer.", 2002, vol. 11, p. 479–517

[66] C. J. ROTH, M. ISMAIL, L. YOSHIHARA, W. A. WALL.*A comprehensive computational human lung model incorporating inter-acinar dependencies: Application to spontaneous breathing and mechanical ventilation*, in "Int. J. Numer. Method. Biomed. Eng.", 2016, vol. 33, n^{o} 1, e02787

[67] L. YOSHIHARA, C. J. ROTH, W. A. WALL.*Fluid-structure interaction including volumetric coupling with homogenised subdomains for modeling respiratory mechanics*, in "Int. J. Numer. Method. Biomed. Eng.", 2016, vol. 33, n^{o} 4, e2812

[68] V. DE SILVA, L.-H. LIM.*Tensor rank and the ill-posedness of the best low-rank approximation problem*, in "SIAM J. Matrix Anal. Appl.", 2008, vol. 30, n^{o} 3, p. 1084–1127

[69] R. VAN LOON, P. D. ANDERSON, J. DE HART, F. P. T. BAAIJENS.*A combined fictitious domain/adaptive meshing method for fluid-structure interaction in heart valves*, in "Int. J. Numer. Methods Fluids", 2004, vol. 46, n^{o} 5, p. 533–544

[70] M. J. VAN DER LAAN, S. ROSE.*Targeted learning*, Springer Series in Statistics, Springer, New York, 2011

Project-Team DELYS

DistributEd aLgorithms and sYStems

IN COLLABORATION WITH: Laboratoire d'informatique de Paris 6 (LIP6)

IN PARTNERSHIP WITH:
Sorbonne Université

RESEARCH CENTER
Paris

THEME
Distributed Systems and middleware

Table of contents

Project-Team DELYS

Creation of the Team: 2018 January 01, updated into Project-Team: 2019 January 01

Keywords:

Computer Science and Digital Science:

A1.1.1. - Multicore, Manycore
A1.1.9. - Fault tolerant systems
A1.2.5. - Internet of things
A1.3.2. - Mobile distributed systems
A1.3.3. - Blockchain
A1.3.4. - Peer to peer
A1.3.5. - Cloud
A1.3.6. - Fog, Edge
A1.5.2. - Communicating systems
A2.6. - Infrastructure software
A2.6.1. - Operating systems
A2.6.2. - Middleware
A2.6.3. - Virtual machines
A2.6.4. - Ressource management
A3.1.8. - Big data (production, storage, transfer)
A7.1.1. - Distributed algorithms

Other Research Topics and Application Domains:

B6.4. - Internet of things

1. Team, Visitors, External Collaborators

Research Scientists

Mesaac Makpangou [Inria, Researcher, HDR]
Marc Shapiro [Inria, Senior Researcher, HDR]
Pascale Vicat - Blanc [Inria, Senior Researcher, from Apr 2019 until Oct 2019, HDR]

Faculty Members

Pierre Sens [Team leader, Sorbonne Univ., Professor, HDR]
Luciana Bezerra Arantes [Sorbonne Univ., Associate Professor]
Philippe Darche [Univ. René Descartes, Associate Professor]
Swan Dubois [Sorbonne Univ., Associate Professor]
Jonathan Lejeune [Sorbonne Univ., Associate Professor]
Franck Petit [Sorbonne Univ., Professor, HDR]
Julien Sopena [Sorbonne Univ., Associate Professor]

External Collaborator

Sébastien Monnet [Univ Savoie Mont-Blanc]

Technical Staff

Ilyas Toumlilt [Inria, Engineer, from Sep 2019]

PhD Students

Sébastien Bouchard [Inria, until Sep 2019]

Marjorie Bournat [Sorbonne Univ., until Aug 2019]
Gabriel Le Bouder [Sorbonne Univ., from Oct. 2019]
Arnaud Favier [Inria]
Denis Jeanneau [Sorbonne Univ., until Jun 2019]
Francis Laniel [Sorbonne Univ.]
Sreeja Nair [Sorbonne Univ.]
Laurent Prosperi [Inria, from Oct 2019]
Benoît Martin [Sorbonne-Univ., from Oct 2019]
Jonathan Sid Otmane [Orange Labs]
Ilyas Toumlilt [Sorbonne Univ., until Apr 2019]
Dimitrios Vasilas [Sorbonne Univ.]

Post-Doctoral Fellows

Sara Hamouda [Inria, from Feb 2019]
Sébastien Bouchard [Sorbonne Univ., Oct. – Dec. 2019]

2. Overall Objectives

2.1. Overall Objectives

The research of the Delys team addresses the theory and practice of distributed systems, including multicore computers, clusters, networks, peer-to-peer systems, cloud and fog computing systems, and other communicating entities such as swarms of robots. It addresses the challenges of correctly communicating, sharing information, and computing in such large-scale, highly dynamic computer systems. This includes addressing the core problems of communication, consensus and fault detection, scalability, replication and consistency of shared data, information sharing in collaborative groups, dynamic content distribution, and multi- and many-core concurrent algorithms.

Delys is a joint research team between LIP6 (Sorbonne University/CNRS) and Inria Paris.

3. Research Program

3.1. Research rationale

DELYS addresses both theoretical and practical issues of *Computer Systems*, leveraging our dual expertise in theoretical and experimental research. Our approach is a "virtuous cycle," triggered by issues with real systems, of algorithm design which we prove correct and evaluate theoretically, and then implement and test experimentally feeding back to theory. The major challenges addressed by DELYS are the sharing of information and guaranteeing correct execution of highly-dynamic computer systems. Our research covers a large spectrum of distributed computer systems: multicore computers, mobile networks, cloud computing systems, and dynamic communicating entities. This holistic approach enables handling related problems at different levels. Among such problems we can highlight consensus, fault detection, scalability, search of information, resource allocation, replication and consistency of shared data, dynamic content distribution, and concurrent and parallel algorithms.

Two main evolutions in the Computer Systems area strongly influence our research project:

(1) Modern computer systems are **increasingly distributed**, **dynamic** and composed of multiple devices **geographically spread over heterogeneous platforms**, spanning multiple management domains. Years of research in the field are now coming to fruition, and are being used by millions of users of web systems, peer-to-peer systems, gaming and social applications, cloud computing, and now fog computing. These new uses bring new challenges, such as *adaptation to dynamically-changing conditions*, where knowledge of the system state can only be partial and incomplete.

(2) **Heterogeneous architectures and virtualisation are everywhere**. The parallelism offered by distributed clusters and *multicore* architectures is opening highly parallel computing to new application areas. To be successful, however, many issues need to be addressed. Challenges include obtaining a consistent view of shared resources, such as memory, and optimally distributing computations among heterogeneous architectures. These issues arise at a more fine-grained level than before, leading to the need for different solutions down to OS level itself.

The scientific challenges of the distributed computing systems are subject to many important features which include scalability, fault tolerance, dynamics, emergent behaviour, heterogeneity, and virtualisation at many levels. Algorithms designed for traditional distributed systems, such as resource allocation, data storage and placement, and concurrent access to shared data, need to be redefined or revisited in order to work properly under the constraints of these new environments. Sometimes, classical "*static*" problems, (*e.g.*, Election Leader, Spanning Tree Construction, ...) even need to be redefined to consider the unstable nature of the distributed system. In particular, DELYS will focus on a number of key challenges:

Consistency in geo-scale systems. Distributed systems need to scale to large geographies and large numbers of attached devices, while executing in an untamed, unstable environment. This poses difficult scientific challenges, which are all the more pressing as the cloud moves more and more towards the edge, IoT and mobile computing. A key issue is how to share data effectively and consistently across the whole spectrum. Delys has made several key contributions, including CRDTs, the Transactional Causal Consistency Plus model, the AntidoteDB geo-distributed database, and its edge extension EdgeAnt.

Rethinking distributed algorithms. From a theoretical point of view the key question is how to adapt the fundamental building blocks to new architectures. More specifically, how to rethink the classical algorithms to take into account the dynamics of advanced modern systems. Since a recent past, there have been several papers that propose models for dynamic systems: there is practically a different model for each setting and currently there is no unification of models. Furthermore, models often suffer of lack of realism. One of the key challenge is to identify which assumptions make sense in new distributed systems. DELYS's objectives are then (1) to identify under which realistic assumptions a given fundamental problem such as mutual exclusion, consensus or leader election can be solved and (2) to design efficient algorithms under these assumptions.

Resource management in heterogeneous systems. The key question is how to manage resources on large and heterogeneous configurations. Managing resources in such systems requires fully decentralized solutions, and to rethink the way various platforms can collaborate and interoperate with each other. In this context, data management is a key component. The fundamental issue we address in ow to efficiently and reliably share information in highly distributed environments.

Adaptation of runtimes. One of the main challenge of the OS community is how to adapt runtime supports to new architectures. With the increasingly widespread use of multicore architectures and virtualised environments, internal runtime protocols need to be revisited. Especially, memory management is crucial in OS and virtualisation technologies have highly impact on it. On one hand, the isolation property of virtualisation has severe side effects on the efficiency of memory allocation since it needs to be constantly balanced between hosted OSs. On the other hand, by hiding the physical machine to OSs, virtualisation prevents them to efficiently place their data in memory on different cores. Our research will thus focus on providing solutions to efficiently share memory between OSs without jeopardizing isolation properties.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Awards

In 2019, DELYS obtained a best paper award at ICDCN 2019 (distributed computing track) .

Alejandro Tomsic, former PhD student of the Delys group, was awarded the Prize for Best French PhD in Systems and Networking 2019 for his thesis titled “Exploring the design space of highly-available distributed transactions” . The prize is awarded yearly by ASF (the French chapter of ACM Sigops) and by RSD (the French research network in Networked and Distributed Systems). The award ceremony and presentation of Alejandro’s work took place at the COMPAS 2019 conference in Anglet, France.

The US patent, titled “Distributing computing system implementing a non-speculative hardware transactional memory and a method for using same for distributed computing,” was awarded to inventors Julien Peeters, Nicolas Ventroux, Tanguy Sassolas and Marc Shapiro in April 2019, with number US 10 416 925 B2 [4].
BEST PAPERS AWARDS :
[28]
E. MAUFFRET, D. JEANNEAU, L. ARANTES, P. SENS.*The Weakest Failure Detector to Solve the Mutual Exclusion Problem in an Unknown Dynamic Environment*, in "20th International Conference on Distributed Computing and Networking (ICDCN 2019)", Bangalore, India, January 2019, Extended version: https://hal.archives-ouvertes.fr/hal-01661127v3, https://hal.archives-ouvertes.fr/hal-01929224

[7]
A. Z. TOMSIC.*Exploring the design space of highly-available distributed transactions*, Université Pierre et Marie Curie, Paris, France, April 2018

5. New Results

5.1. Distributed Algorithms for Dynamic Networks and Fault Tolerance

Participants: Luciana Bezerra Arantes [correspondent], Sébastien Bouchard, Marjorie Bournat, João Paulo de Araujo, Swan Dubois, Laurent Feuilloley, Denis Jeanneau, Jonathan Lejeune, Franck Petit, Pierre Sens, Julien Sopena.

Nowadays, distributed systems are more and more heterogeneous and versatile. Computing units can join, leave or move inside a global infrastructure. These features require the implementation of *dynamic* systems, that is to say they can cope autonomously with changes in their structure in terms of physical facilities and software. It therefore becomes necessary to define, develop, and validate distributed algorithms able to managed such dynamic and large scale systems, for instance mobile *ad hoc* networks, (mobile) sensor networks, P2P systems, Cloud environments, robot networks, to quote only a few.

The fact that computing units may leave, join, or move may result of an intentional behavior or not. In the latter case, the system may be subject to disruptions due to component faults that can be permanent, transient, exogenous, evil-minded, etc. It is therefore crucial to come up with solutions tolerating some types of faults.

In 2019, we obtained the following results.

5.1.1. Failure detectors

Mutual exclusion is one of the fundamental problems in distributed computing but existing mutual exclusion algorithms are unadapted to the dynamics and lack of membership knowledge of current distributed systems (e.g., mobile ad-hoc networks, peer-to-peer systems, etc.). Additionally, in order to circumvent the impossibility of solving mutual exclusion in asynchronous message passing systems where processes can crash, some solutions include the use of $(\mathcal{T}+\Sigma^l)$, which is the weakest failure detector to solve mutual exclusion in known static distributed systems. In [28], we define a new failure detector $\mathcal{T}\Sigma^{lr}$ which is equivalent to $(\mathcal{T}+\Sigma^l)$ in known static systems, and prove that $\mathcal{T}\Sigma^{lr}$ is the weakest failure detector to solve mutual exclusion in unknown dynamic systems with partial memory losses. We consider that crashed processes may recover.

Assuming a message-passing environment with a majority of correct processes, the necessary and sufficient information about failures for implementing a general state machine replication scheme ensuring consistency is captured by the Ω failure detector. We show in [19] that in such a message-passing environment, Ω is also the weakest failure detector to implement an eventually consistent replicated service, where replicas are expected to agree on the evolution of the service state only after some (a priori unknown) time.

5.1.2. Scheduler Tolerant to Temporal Failures in Clouds

Cloud platforms offer different types of virtual machines which ensure different guarantees in terms of availability and volatility, provisioning the same resource through multiple pricing models. For instance, in Amazon EC2 cloud, the user pays per hour for on-demand instances while spot instances are unused resources available for a lower price. Despite the monetary advantages, a spot instance can be terminated or hibernated by EC2 at any moment. Using both hibernation prone spot instances (for cost sake) and on-demand instances, we propose in [31] a static scheduling for applications which are composed of independent tasks (bag-of-task) with deadline constraints. However, if a spot instance hibernates and it does not resume within a time which guarantees the application's deadline, a temporal failure takes place. Our scheduling, thus, aims at minimizing monetary costs of bag-of-tasks applications in EC2 cloud, respecting its deadline and avoiding temporal failures. Performance results with task execution traces, configuration of Amazon EC2 virtual machines, and EC2 market history confirms the effectiveness of our scheduling and that it tolerates temporal failures. In [30], we extend our approach for dynamic scheduling.

5.1.3. Gathering of Mobile Agents

Gathering a group of mobile agents is a fundamental task in the field of distributed and mobile systems. It consists of bringing agents that initially start from different positions to meet all together in finite time. In the case when there are only two agents, the gathering problem is often referred to as the rendezvous problem.

In [14] we show that rendezvous under the strong scenario is possible for agents with asynchrony restricted in the following way: agents have the same measure of time but the adversary can impose, for each agent and each edge, the speed of traversing this edge by this agent. The speeds may be different for different edges and different agents but all traversals of a given edge by a given agent have to be at the same imposed speed. We construct a deterministic rendezvous algorithm for such agents, working in time polynomial in the size of the graph, in the length of the smaller label, and in the largest edge traversal time.

5.1.4. Perpetual self-stabilizing exploration of dynamic environments

In [15], we deal with the classical problem of exploring a ring by a cohort of synchronous robots. We focus on the perpetual version of this problem in which it is required that each node of the ring is visited by a robot infinitely often. We assume that the robots evolve in ring-shape TVGs, *i.e.*, the static graph made of the same set of nodes and that includes all edges that are present at least once over time forms a ring of arbitrary size. We also assume that each node is infinitely often reachable from any other node. In this context, we aim at providing a self-stabilizing algorithm to the robots (*i.e.*, the algorithm must guarantee an eventual correct behavior regardless of the initial state and positions of the robots). We show that this problem is deterministically solvable in this harsh environment by providing a self-stabilizing algorithm for three robots.

5.1.5. Torus exploration by oblivious robots

In [17], we deal with a team of autonomous robots that are endowed with motion actuators and visibility sensors. Those robots are weak and evolve in a discrete environment. By weak, we mean that they are anonymous, uniform, unable to explicitly communicate, and oblivious. We first show that it is impossible to solve the terminating exploration of a simple torus of arbitrary size with less than 4 or 5 such robots, respectively depending on whether the algorithm is probabilistic or deterministic. Next, we propose in the SSYNC model a probabilistic solution for the terminating exploration of torus-shaped networks of size $\ell \times L$, where $7 \leq \ell \leq L$, by a team of 4 such weak robots. So, this algorithm is optimal *w.r.t.* the number of robots.

5.1.6. Explicit communication among stigmergic robots

In [18], we investigate avenues for the exchange of information (explicit communication) among deaf and mute mobile robots scattered in the plane. We introduce the use of movement-signals (analogously to flight signals and bees waggle) as a mean to transfer messages, enabling the use of distributed algorithms among robots. We propose one-to-one deterministic movement protocols that implement explicit communication among semi-synchronous robots. Our protocols enable the use of distributing algorithms based on message exchanges among swarms of stigmergic robots. They also allow robots to be equipped with the means of communication to tolerate faults in their communication devices.

5.1.7. Gradual stabilization

In [13], we introduce the notion of *gradual stabilization under* (τ, ρ)*-dynamics* (gradual stabilization, for short). A gradually stabilizing algorithm is a self-stabilizing algorithm with the following additional feature: after up to τ*dynamic steps* of a given type ρ occur starting from a legitimate configuration, it first quickly recovers to a configuration from which a specification offering a minimum quality of service is satisfied.

It then gradually converges to specifications offering stronger and stronger safety guarantees until reaching a configuration (1) from which its initial (strong) specification is satisfied again, and (2) where it is ready to achieve gradual convergence again in case of up to τ new dynamic steps of type ρ. A gradually stabilizing algorithm being also self-stabilizing, it still recovers within finite time (yet more slowly) after any other finite number of transient faults, including for example more than τ arbitrary dynamic steps or other failure patterns such as memory corruptions. We illustrate this new property by considering three variants of a synchronization problem respectively called *strong*, *weak*, and *partial* unison. We propose a self-stabilizing unison algorithm which achieves gradual stabilization in the sense that after one dynamic step of a certain type *BULCC* (such a step may include several topological changes) occurs starting from a configuration which is legitimate for the strong unison, it maintains clocks almost synchronized during the convergence to strong unison: it satisfies partial unison immediately after the dynamic step, then converges in at most one round to weak unison, and finally re-stabilizes to strong unison.

5.2. Distributed systems and Large-scale data distribution

Participants: Guillaume Fraysse, Saalik Hatia, Mesaac Makpangou, Sreeja Nair, Jonathan Sid-Otmane, Pierre Sens, Marc Shapiro, Ilyas Toumlilt, Dimitrios Vasilas.

5.2.1. Proving the safety of highly-available distributed objects

To provide high availability in distributed systems, object replicas allow concurrent updates. Although replicas eventually converge, they may diverge temporarily, for instance when the network fails. This makes it difficult for the developer to reason about the object's properties , and in particular, to prove invariants over its state. For the sub-class of state-based distributed systems, we propose a proof methodology for establishing that a given object maintains a given invariant, taking into account any concurrency control. Our approach allows reasoning about individual operations separately. We demonstrate that our rules are sound, and we illustrate their use with some representative examples. We automate the rule using Boogie, an SMT-based tool.

This work is accepted for publication at the 29th European Symposium on Programming (ESOP), April 2020, Dublin, Ireland [34]. Preliminary results were presented at the Workshop on Principles and Practice of Consistency for Distributed Data (PaPoC), March 2019, Dresden, Germany [29].

5.3. Resource management in system software

Participants: Jonathan Lejeune, Marc Shapiro, Julien Sopena, Francis Laniel.

5.3.1. MemOpLight: Leveraging applicative feedback to improve container memory consolidation

The container mechanism supports consolidating several servers on the same machine, thus amortizing cost. To ensure performance isolation between containers, Linux relies on memory limits. However these limits are static, but application needs are dynamic; this results in poor performance. To solve this issue, MemOpLight reallocates memory to containers based on dynamic applicative feedback. MemOpLight rebalances physical memory allocation, in favor of under-performing ones, with the aim of improving overall performance. Our research explores the issues, addresses the design of MemOpLight, and validates it experimentally. Our approach increases total satisfaction by 13% compared to the default.

It is standard practice in Infrastructure as a Service to *consolidate* several logical servers on the same physical machine, thus amortizing cost. However, the execution of one logical server should not disturb the others: the logical servers should remain *isolated* from one another.

To ensure both consolidation and isolation, a recent approach is "containers," a group of processes with sharing and isolation properties. To ensure *memory performance isolation*, *i.e.*, guaranteeing to each container enough memory for it to perform well, the administrator limits the total amount of physical memory that a container may use at the expense of others. In previous work, we showed that these limits impede memory consolidation [26]. Furthermore, the metrics available to the kernel to evaluate its policies (*e.g.*, frequency of page faults, I/O requests, use of CPU cycles, *etc.*), are not directly relevant to performance as experienced from the application perspective, which is better characterized by, for instance, response time or throughput measured at application level.

To solve these problems, we propose a new approach, called the Memory Optimization Light (MemOpLight). It is based on application-level feedback from containers. Our mechanism aims to rebalance memory allocation in favor of unsatisfied containers, while not penalizing the satisfied ones. By doing so, we guarantee application satisfaction, while consolidating memory; this also improves overall resource consumption.

Our main contributions are the following:

- An experimental demonstration of the limitations of the existing Linux mechanisms.
- The design of a simple feedback mechanism from application to the kernel.
- An algorithm for adapting container memory allocation.
- And implementation in Linux and experimental confirmation.

This work is currently under submission at a major conference. Some preliminary results are published at NCA 2019 [26].

6. Bilateral Contracts and Grants with Industry

6.1. Bilateral Contracts with Industry

DELYS has a CIFRE contract with Scality SA:

- Dimitrios Vasilas is advised by Marc Shapiro and Brad King. He works on secondary indexing in large-scale storage systems under weak consistency.

DELYS has three contracts with Orange within the I/O Lab joint laboratory:

- Guillaume Fraysse is advised by Jonathan Lejeune, Julien Sopena, and Pierre Sens. He works on distributed resources allocation in virtual network environments.
- Jonathan Sid-Otmane is advised by Marc Shapiro. He studies the applications of distributed databases to the needs of the telco industry in the context of 5G.
- José Alves Esteves Jurandir is advised by Pierre Sens. He works on network slice placement stategies.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. ANR

7.1.1.1. AdeCoDS (2019–2023)

Title: Programming, verifying, and synthesizing Adequately-Consistent Distributed Systems (AdeCoDS).

Members: Université de Paris (project leader), Sorbonne-Université LIP6, ARM, Orange.

Funding: The total funding of AdeCoDS from ANR is 523 471 euros, of which 162 500 euros for Delys.

Objectives The goal of the project is to provide a framework for programming distributed systems that are both correct and efficient (available and performant). The idea is to offer to developers a programming framework where it is possible, for a given application, (1) to build implementations that are correct under specific assumptions on the consistency level guaranteed by the infrastructure (e.g., databases and libraries of data structures), and (2) to discover in a systematic way the different trade-offs between the consistency level guaranteed by the infrastructure and the type and the amount of synchronization they need to use in their implementation in order ensure its correctness. For that, the project will develop a methodology based on combining (1) automated verification and synthesis methods, (2) language-based methods for correct programming, and (3) techniques for efficient system design.

7.1.1.2. ESTATE - (2016–2021)

Members: LIP6 (DELYS, project leader), LaBRI (Univ. de Bordeaux); Verimag (Univ. de Grenoble).

Funding: ESTATE is funded by ANR (PRC) for a total of about 544 000 euros, of which 233 376 euros for DELYS.

Objectives: The core of ESTATE consists in laying the foundations of a new algorithmic framework for enabling Autonomic Computing in distributed and highly dynamic systems and networks. We plan to design a model that includes the minimal algorithmic basis allowing the emergence of dynamic distributed systems with self-* capabilities, *e.g.*, self-organization, self-healing, self-configuration, self-management, self-optimization, self-adaptiveness, or self-repair. In order to do this, we consider three main research streams:

(i) building the theoretical foundations of autonomic computing in dynamic systems, (ii) enhancing the safety in some cases by establishing the minimum requirements in terms of amount or type of dynamics to allow some strong safety guarantees, (iii) providing additional formal guarantees by proposing a general framework based on the Coq proof assistant to (semi-)automatically construct certified proofs.

The coordinator of ESTATE is Franck Petit.

7.1.1.3. RainbowFS - (2016–2020)

Members: LIP6 (DELYS, project leader), Scality SA, CNRS-LIG, Télécom Sud-Paris, Université Savoie-Mont-Blanc.

Funding: is funded by ANR (PRC) for a total of 919 534 euros, of which 359 554 euros for DELYS.

Objectives: RainbowFS proposes a “just-right” approach to storage and consistency, for developing distributed, cloud-scale applications. Existing approaches shoehorn the application design to some predefined consistency model, but no single model is appropriate for all uses. Instead, we propose tools to co-design the application and its consistency protocol. Our approach reconciles the conflicting requirements of availability and performance vs. safety: common-case operations are designed to be asynchronous; synchronisation is used only when strictly necessary to satisfy the application’s integrity invariants. Furthermore, we deconstruct classical consistency models into orthogonal primitives that the developer can compose efficiently, and provide a number of tools for quick, efficient and correct cloud-scale deployment and execution. Using this methodology, we will develop an entreprise-grade, highly-scalable file system, exploring the rainbow of possible semantics, and we demonstrate it in a massive experiment.

The coordinator of RainbowFS is Marc Shapiro.

7.1.2. LABEX

7.1.2.1. SMART - (2012–2019)

Members: ISIR (Sorbonne Univ./CNRS), LIP6 (Sorbonne Univ./CNRS), LIB (Sorbonne Univ./INSERM), LJLL (Sorbonne Univ./CNRS), LTCI (Institut Mines-Télécom/CNRS), CHArt-LUTIN (Univ. Paris 8/EPHE), L2E (Sorbonne Univ.), STMS (IRCAM/CNRS).

Funding: Sorbonne Universités, ANR.

Description: The SMART Labex project aims globally to enhancing the quality of life in our digital societies by building the foundational bases for facilitating the inclusion of intelligent artifacts in our daily life for service and assistance. The project addresses underlying scientific questions raised by the development of Human-centered digital systems and artifacts in a comprehensive way. The research program is organized along five axes and DELYS is responsible of the axe “Autonomic Distributed Environments for Mobility.”

The project involves a PhD grant of 100 000 euros over 3 years.

7.2. European Initiatives

7.2.1. FP7 & H2020 Projects

7.2.1.1. LightKone

Title: Lightweight Computation for Networks at the Edge

Programm: H2020-ICT-2016-2017

Duration: January 2017 - December 2019

Coordinator: Université Catholique de Louvain

Partners:

Université Catholique de Louvain (Belgium)

Technische Universitaet Kaiserslautern (Germany)

INESC TEC - Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciencia (Portugal)

Faculdade de Ciencias E Tecnologiada Universidade Nova de Lisboa (Portugal)

Universitat Politecnica De Catalunya (Spain)

Scality (France)

Gluk Advice B.V. (Netherlands)

Inria contact: Marc Shapiro

The goal of LightKone is to develop a scientifically sound and industrially validated model for doing general-purpose computation on edge networks. An edge network consists of a large set of heterogeneous, loosely coupled computing nodes situated at the logical extreme of a network. Common examples are networks of Internet of Things, mobile devices, personal computers, and points of presence including Mobile Edge Computing. Internet applications are increasingly running on edge networks, to reduce latency, increase scalability, resilience, and security, and permit local decision making. However, today’s state of the art, the gossip and peer-to-peer models, give no solution for defining general-purpose computations on edge networks, i.e., computation with shared mutable state. LightKone will solve this problem by combining two recent advances in distributed computing, namely synchronisation-free programming and hybrid gossip algorithms, both of which are successfully used separately in industry. Together, they are a natural combination for edge computing. We will cover edge networks both with and without data center nodes, and applications focused on collaboration, computation, and both. Project results will be new programming models and algorithms that advance scientific understanding, implemented in new industrial applications and a startup company, and evaluated in large-scale realistic settings.

7.3. International Initiatives

7.3.1. Participation in Other International Programs

7.3.1.1. Spanish research ministry project

Title: BFT-DYNASTIE - Byzantine Fault Tolerance: Dynamic Adaptive Services for Partitionable Systems

French Partners: Labri, Irisa, LIP6

International Partners (Institution - Laboratory - Researcher):

> University of the Basque Country UPV - Spain, EPFL - LSD - Switzerland, Friedrich-Alexander-Universitat Erlangen-Nurenberg - Deutschland, University of Sydney - Australia

Duration: 2017–2019

The project BFT-DYNASTIE is aimed at extending the model based on the alternation of periods of stable and unstable behavior to all aspects of fault-tolerant distributed systems, including synchrony models, process and communication channel failure models, system membership, node mobility, and network partitioning. The two main and new challenges of this project are: the consideration of the most general and complex to address failure model, known as Byzantine, arbitrary or malicious, which requires qualified majorities and the use of techniques form the security area; and the operation of the system in partitioned mode, which requires adequate reconciliation mechanisms when two partitions merge.

7.3.1.2. Spanish research ministry project

Title: BFT-DYNASTIE - Byzantine Fault Tolerance: Dynamic Adaptive Services for Partitionable Systems

French Partners: Labri, Irisa, LIP6

International Partners (Institution - Laboratory - Researcher):

> University of the Basque Country UPV - Spain, EPFL - LSD - Switzerland, Friedrich-Alexander-Universitat Erlangen-Nurenberg - Deutschland, University of Sydney - Australia

Duration: 2017–2019

The project BFT-DYNASTIE is aimed at extending the model based on the alternation of periods of stable and unstable behavior to all aspects of fault-tolerant distributed systems, including synchrony models, process and communication channel failure models, system membership, node mobility, and network partitioning. The two main and new challenges of this project are: the consideration of the most general and complex to address failure model, known as Byzantine, arbitrary or malicious, which requires qualified majorities and the use of techniques form the security area; and the operation of the system in partitioned mode, which requires adequate reconciliation mechanisms when two partitions merge.

7.3.1.3. STIC Amsud

Title: ADMITS - Architecting Distributed Monitoring and Analytics for IoT in Disaster Scenarios

International Partners (Institution - Laboratory - Researcher):

> Universidad Diego Portales and Universidad Tecnica Federico Santa Maria (Chile)
>
> Universidade Federal de Uberlandia, Universidade Federal do Rio Grande do Norte and Instituto Federal Sul-Rio-Grandense (Brazil)
>
> Universidad de la Republica (Uuruguay)

Duration: 2019 - 2020

Start year: 2019

Develop algorithms, protocols and architectures to enable a decentralized distributed computing environment to provide support for failure monitoring and data analytics in Internet-of-Things (IoT) disaster scenarios.

7.4. International Research Visitors

7.4.1. Visits of International Scientists

- AMOZARRAIN Ugaitz, PhD Student, University of San Sebastian (Spain), Feb. 2019 - Mar. 2019
- CORREA Leonardo, PhD Student, Federal University of Rio Grande do Sul (Brazil), Jan 2019 - Oct. 2019
- GOUVEIA LIMA Luan Teylo, PhD Student, UFF (Brazil), Sep. 2019-Mar. 2020
- PELC Andrzej, Professor, Université du Québec en Outaouais (Canada), Sep. 2019 - Oct. 2019
- DIEUDONNE Yoann, Associate Professor, Amiens Univ., Sep. 2019-Oct. 2019
- LONG Darrell, Professor, Univ. California Santa Cruz (USA), Feb. 2019 - Mar. 2019
- PARIS Jehan-François, Professor, University of Houston (USA), Feb. 2019 - Mar. 2019

7.4.2. Visits to International Teams

Marc Shapiro spent three weeks visiting Technical University Kaiserslautern during the Spring.

Luciana Arantes and Pierre Sens have been invited for 10 days at New-York University Shanghai

Luciana Arantes visited the network team at Pontifical Catholic University of Rio de Janeiro - PUC (Brazil)

Luciana Arantes and Pierre Sens visited the computer science department at Universidade Federal Fluminense - UFF (Brazl)

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events: Organisation

8.1.1.1. Chair of Conference Steering Committees

- Marc Shapiro is a member of the steering committee of the yearly Workshop on Principles and Practice of Consistency (PaPoC), co-located with EuroSys.

8.1.2. Scientific Events: Selection

8.1.2.1. Member of the Conference Program Committees

Marc Shapiro, member of PC of the European Conference on Computer Systems (EuroSys) 2020, in Heraklion (Greece).

Marc Shapiro, member of PC of the European Conference on Computer Systems (EuroSys) 2019, in Dresden (Germany).

Marc Shapiro, member of Progamme Committed of the ACM Symposium on Principles of Distributed Computing (PODC) 2019, in Salerno, Italy.

Marc Shapiro, member of the PC of the Workshop on Planetary-Scale Distributed Systems (W-PSDS) 2019, in Lyon.

Marc Shapiro, member of the Technical Program Committee of the Workshop on Advanced tools, programming languages, and PLatforms for Implementing and Evaluating algorithms for Distributed systems (ApPLIED) 2019, in Budapest, Hungary.

Pierre Sens, 29th International Symposium on Software Reliability Engineering (ISSRE 2019), 19th IEEE International Symposium on Network Computing and Applications (NCA 2019), 38th IEEE International Symposium on Reliable Distributed Systems (SRDS 2019), Gestion de Données – Principes, Technologies et Applications (BDA 2019).

Swan Dubois, 21st International Symposium on Stabilization, Safety, and Security of Distributed Systems (SSS 2019), 21èmes Rencontres Francophones pour les Aspects Algorithmiques des Télécommunications (AlgoTel'18).

Jonathan Lejeune, 22nd Innovation in Clouds, Internet and Networks (ICIN 2019)

8.1.2.2. Reviewer

Swan Dubois, 26th International Colloquium on Structural Information and Communication Complexity (SIROCCO'19).

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

Pierre Sens, International Journal of High Performance Computing and Networking (IJHPCN)

8.1.3.2. Reviewer - Reviewing Activities

Swan Dubois, Distributed Computing, Theoretical Computer Science.

Luciana Arantes, Journal of Parallel and Distributed Computing.

Pierre Sens, Journal of Parallel and Distributed Computing

8.1.4. Invited Talks

Marc Shapiro gave the keynote presentation titled "Living on the edge, safely; or: Life without consensus" at the 7th International Conference on Networked Systems, in Marrakech, Morroco, June 2019.

Marc Shapiro gave an invited talk on "The programming continuum, from core to edge," at the Workshop on Verification of Distributed Systems (VDS), June 2019, Marrakech, Morroco.

Marc Shapiro gave two invited talks at the Dagstuhl Seminar on "Programming Languages for Distributed Systems and Distributed Data Management," October 2019.

Marc Shapiro was invited to speak about "Living Without Consensus," at the seminar "Taking Stock of Distributed Computing," at Collège de France, April 2019. The seminar was organised in conjunction with the *Chaire informatique et sciences numériques* of Rachid Guerraoui.

Pierre Sens, *Probabilistic Byzantine Tolerance Scheduling in Hybrid Cloud Environments*. Research Seminar, University of Fluminense, Brazil, October 2019

Pierre Sens, *Fault tolerance in large and dynamic distributed systems*. Research Seminar, University of Fluminense, Brazil, October 2019

Luciana Arabtes, *A Communication-Efficient Causal Broadcast Protocol*. Research Seminar, University of Fluminense, PUC-Rio, Brazil, October 2019

8.1.5. Leadership within the Scientific Community

- Marc Shapiro is Vice-Chair for Research of Société informatique de France, the French learned society in Informatics.

8.1.6. Research Administration

Pierre Sens, since 2016: Member of Section 6 of the national committee for scientific research CoNRS

Franck Petit, Pierre Sens, since 2012: Member of the Executive Committee of Labex SMART, Co-Chairs of Track 4, Autonomic Distributed Environments for Mobility.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Julien Sopena is Member of “Directoire des formations et de l’insertion professionnelle” of Sorbonne Université, France

Master: Julien Sopena is responsible of Computer Science Master’s degree in Distributed systems and applications (in French, SAR), Sorbonne Universités, France

Master: Luciana Arantes, Swan Dubois, Jonathan Lejeune, Franck Petit, Pierre Sens, Julien Sopena, Advanced distributed algorithms, M2, Sorbonne Université, France

Master: Jonathan Lejeune, Designing Large-Scale Distributed Applications, M2, Sorbonne Université, France

Master: Maxime Lorrillere, Julien Sopena, Linux Kernel Programming, M1, Sorbonne Université, France

Master: Luciana Arantes, Swan Dubois, Jonathan Lejeune, Pierre Sens, Julien Sopena, Operating systems kernel, M1, Sorbonne Université, France

Master: Luciana Arantes, Swan Dubois, Franck Petit, Distributed Algorithms, M1, Sorbonne Université, France

Master: Franck Petit, Autonomic Networks, M2, Sorbonne Université, France

Master: Franck Petit, Distributed Algorithms for Networks, M1, Sorbonne Université, France

Master: Jonathan Lejeune, Julien Sopena, Client-server distributed systems, M1, Sorbonne Université, France.

Master: Julien Sopena, Marc Shapiro, Ilyas Toumlilt, Francis Laniel. Kernels and virtual machines (*Noyaux et machines virtuelles*, NMV), M2, Sorbonne Université, France.

Licence: Pierre Sens, Luciana Arantes, Julien Sopena, Principles of operating systems, L3, Sorbonne Université, France

Licence: Swan Dubois, Initiation to operating systems, L3, Sorbonne Université, France

Licence: Swan Dubois, Multi-threaded Programming, L3, Sorbonne Université, France

Licence: Jonathan Lejeune, Oriented-Object Programming, L3, Sorbonne Université, France

Licence: Franck Petit, Advanced C Programming, L2, Sorbonne Université, France

Licence: Swan Dubois, Jonathan Lejeune, Julien Sopena, Introduction to operating systems, L2, Sorbonne Université, France

Licence: Mesaac Makpangou, C Programming Language, 27 h, L2, Sorbonne Université, France

Ingénieur 4ème année : Marc Shapiro, Introduction aux systèmes d’exploitation, 26 h, M1, Polytech Sorbonne Université, France.

Licence : Philippe Darche (coordinator), Architecture of Internet of Things (IoT), 2 × 32h, L3, Institut Universitaire Technologique (IUT) Paris Descartes, France.

Engineering School: Philippe Darche (coordinator), Solid-State Memories, 4th year, ESIEE, France.

DUT: Philippe Darche (coordinator), Introduction to Computer Systems - Data representation, 60h, Institut Universitaire Technologique (IUT) Paris Descartes, France.

DUT: Philippe Darche (coordinator), Computer Architecture, 32h, Institut Universitaire Technologique (IUT) Paris Descartes, France.

DUT: Philippe Darche (coordinator), Computer Systems Programming, 80h, Institut Universitaire Technologique (IUT) Paris Descartes, France.

8.2.2. Supervision

PhD: João Paulo de Araujo, “L’exécution efficace d’algorithmes distribués dans les réseaux véhiculaires”, funded by CNPq (Brésil), Sorbonne Univ, Apr. 2019. Advised by Pierre Sens and Luciana Arantes.

PhD: Sébastien Bouchard, “Gathering with faulty robots”, Sorbonne Univ., Sep. 2019. Advised by Swan Dubois, Franck Petit, Yoann Dieudonné (University of Picardy Jules Verne)

PhD: Marjorie Bournat, “Speculation and Graceful Degradability for Robots in Highly Dynamic Environments”, Sorbonne Univ., Jun. 2019. Advised by Swan Dubois, Franck Petit, Yoann Dieudonné (University of Picardy Jules Verne)

PhD: Damien Carver, “HACHE : HorizontAl Cache cHorEgraphy - Toward automatic resizing of shared I/O caches.”, Sorbonne Univ., May 2019. Advised by Sébastien Monnet, Julien Sopena, Dimitri Refauvelet (Magency).

CIFRE PhD in progress: José Alves Esteves, “Adaptation dynamique en environnements répartis contraints”, Sorbonne Univ., since Sep. 2019. Advised by Pierre Sens and Amina Boubendir Orange Labs.

PhD in progress: Arnaud Favier, “Algorithmes de coordination répartis dans des réseaux dynamiques”, Sorbonne Univ., since Sep. 2018. Advised by Pierre Sens and Luciana Arantes.

PhD in progress: Saalik Hatia, “Efficient management of memory and storage for CRDTs”, Sorbonne Univ., since Oct. 2018. Advised by Marc Shapiro.

CIFRE PhD in progress: Guillaume Fraysse, Orange Lab - Inria, “Ubiquitous Resouces for Service Availability”, since Jul. 2017. Advised by Pierre Sens, Imen Grida Ben Yahia (Orange-Lab) , Jonathan Lejeune, Julien Sopena.

PhD in progress: Francis Laniel, “Vers une utilisation efficace de la mémoire non volatile pour économiser l’énergie.”Sorbonne Univ., since Sept. 2017. Advised by Marc Shapiro, Julien Sopena, Jonathan Lejeune.

PhD in progress: Gabriel Le Bouder, “Autonomic synchronization”, Sorbonne Univ., since Sep. 2019. Advised by Franck Petit.

PhD in progress: Benoît Martin, “Protocol de cohérence hybride: de la cohérence causal à la cohérence forte”, Sorbonne Univ., since Sep. 2019. Advised by Mesaac Makpangou and Marc Shapiro.

PhD in progress: Sreeja Nair, “Just-Right Consistency for massive gee-replicated storage”, Sorbonne Univ., since Apr. 2018. Advised by Marc Shapiro.

PhD in progress: Laurent Prosperi, “Abstractions, langage et runtime pour les systèmes distribués”, Sorbonne Univ., since Sep. 2019. Advised by Marc Shapiro.

CIFRE PhD in progress: Jonathan Sid-Otmane. “Étude des critères de distribution et de l’usage d’une base de données distribuée pour un OS Telco”, since Dec. 2017. Advised by Marc Shapiro, with Sofiane Imadali and Frédéric Martelli, Orange Labs.

PhD in progress: Ilyas Toumlilt, “Bridging the CAP gap, all the way to the edge”, Sorbonne Univ., since Sep. 2016. Advised by Marc Shapiro.

CIFRE PhD in progress: Dimitrios Vasilas, “Indexing in large-scale storage systems”, Sorbonne Univ., since Sep. 2016. Advised by Marc Shapiro, with Brad King, Scality.

PhD in progress: Daniel Wladdimiro, “Adaptation dynamique en environnements répartis contraints”, Sorbonne Univ., since Sep. 2019. Advised by Pierre Sens and Luciana Arantes.

PhD in progress: Daniel Wilhelm, “Algorithmes de diffusion causale dans les systèmes répartis dynamique”, Sorbonne Univ., since Oct. 2019, Pierre Sens and Luciana Arantes.

8.2.3. Juries

Franck Petit was the reviewer of:

- David Ilcinkas, HDR, LaBRI, Bordeaux

Pierre Sens was the reviewer of:

- Heithem Abbes, HDR, LIPN, Villetaneuse
- Xavier Etchevers, HDR, LIG, Univ. Grenoble
- Nicolas Aussel, PhD, SAMOVAR, Telecom Sud Paris
- Paul Chaignon, PhD, LORIA, Nancy
- Jad Darrous, PhD, LIP, ENS-Lyon
- Umar Ozeer, PhD, LIG, Univ. Grenoble
- Bharati Sinha, PhD, National Institute of Technology Kurukshetra, India
- Amir Teshome Wonjiga, PhD, IRISA, Rennes

Pierre Sens was Chair of

- Julien Loudet, PhD, DAVID, UVSQ
- Asma Berriri, PhD, SAMOVAR, Telecom Sud Paris

Marc Shapiro was a reviewer for the CSD (mid-thesis commitee) for Khaled Zaouk, PhD student of Yanlei Diao at École Polytechnique.

8.3. Popularization

8.3.1. Articles and contents

- Swan Dubois and Franck Petit coauthored the book entitled "*Introduction to Distributed Self-Stabilizing Algorithms*" [35], that aims at being a comprehensive and pedagogical introduction to the concept of *self-stabilization*.
- Marc Shapiro co-authored two entries for the Encyclopedia of Big Data Technologies (published by Springer, 2019):
  - "Database Consistency Models" [38].
  - "Conflict-Free Replicated Data Types CRDTs" [37].
- Marc Shapiro contributed the post "Living at the edge, safely" to the LightKone blog [6].

9. Bibliography

Major publications by the team in recent years

[1] V. Balegas, N. Preguiça, R. Rodrigues, S. Duarte, C. Ferreira, M. Najafzadeh, M. Shapiro.*Putting Consistency back into Eventual Consistency*, in "Euro. Conf. on Comp. Sys. (EuroSys)", Bordeaux, France, April 2015, p. 6:1–6:16, https://doi.org/10.1145/2741948.2741972

[2] L. Gidra, G. Thomas, J. Sopena, M. Shapiro, N. Nguyen.*NumaGiC: a garbage collector for big data on big NUMA machines*, in "Int. Conf. on Archi. Support for Prog. Lang. and Systems (ASPLOS)", Istanbul, Turkey, Assoc. for Computing Machinery, March 2015, p. 661–673, http://dx.doi.org/10.1145/2694344.2694361

[3] A. Gotsman, H. Yang, C. Ferreira, M. Najafzadeh, M. Shapiro.*'Cause I'm Strong Enough: Reasoning about Consistency Choices in Distributed Systems*, in "Symp. on Principles of Prog. Lang. (POPL)", St. Petersburg, FL, USA, 2016, p. 371–384, http://dx.doi.org/10.1145/2837614.2837625

[4] J. PEETERS, N. VENTROUX, T. SASSOLAS, M. SHAPIRO.*Distributing computing system implementing a non-speculative hardware transactional memory and a method for using same for distributed computing*, United States Patent and Trademark Office (USPTO), September 2019, n° US 10 416 925 B2

[5] M. SHAPIRO, N. PREGUIÇA, C. BAQUERO, M. ZAWIRSKI.*Conflict-free Replicated Data Types*, in "Int. Symp. on Stabilization, Safety, and Security of Dist. Sys. (SSS)", Grenoble, France, X. DÉFAGO, F. PETIT, V. VILLAIN (editors), Lecture Notes in Comp. Sc., Springer-Verlag, October 2011, vol. 6976, p. 386–400, http://lip6.fr/Marc.Shapiro/papers/CRDTs_SSS-2011.pdf

[6] M. SHAPIRO.*Living at the edge, safely*, LightKone European Project, May 2019

[7] *Best Paper*
A. Z. TOMSIC.*Exploring the design space of highly-available distributed transactions*, Université Pierre et Marie Curie, Paris, France, April 2018.

[8] M. ZAWIRSKI, N. PREGUIÇA, S. DUARTE, A. BIENIUSA, V. BALEGAS, M. SHAPIRO.*Write Fast, Read in the Past: Causal Consistency for Client-side Applications*, in "Int. Conf. on Middleware (MIDDLEWARE)", Vancouver, BC, Canada, ACM/IFIP/Usenix, December 2015, p. 75–87

Publications of the year

Doctoral Dissertations and Habilitation Theses

[9] S. BOUCHARD.*On the Deterministic Gathering of Mobile Agents*, Sorbonne Université, September 2019, https://tel.archives-ouvertes.fr/tel-02320156

[10] M. BOURNAT.*Graceful Degradation and Speculation for Robots in Highly Dynamic Environments*, Sorbonne Université, June 2019, https://hal.inria.fr/tel-02177304

[11] D. CARVER.*Advanced Consolidation for Dynamic Containers*, EDITE de Paris, May 2019, https://tel.archives-ouvertes.fr/tel-02393773

[12] J. P. DE ARAUJO.*A communication-efficient causal broadcast publish/subscribe system*, Sorbonne Université ; LIP6 - Laboratoire d'Informatique de Paris 6, April 2019, https://tel.archives-ouvertes.fr/tel-02105743

Articles in International Peer-Reviewed Journal

[13] K. ALTISEN, S. DEVISMES, A. DURAND, F. PETIT.*Gradual stabilization*, in "Journal of Parallel and Distributed Computing", January 2019, vol. 123, p. 26-45 [*DOI :* 10.1016/J.JPDC.2018.09.002], https://hal.archives-ouvertes.fr/hal-02420362

[14] S. BOUCHARD, M. BOURNAT, Y. DIEUDONNÉ, S. DUBOIS, F. PETIT.*Asynchronous approach in the plane: a deterministic polynomial algorithm*, in "Distributed Computing", August 2019, vol. 32, n° 4, p. 317-337 [*DOI :* 10.1007/S00446-018-0338-2], https://hal.inria.fr/hal-02413273

[15] M. BOURNAT, A. K. DATTA, S. DUBOIS.*Self-stabilizing robots in highly dynamic environments*, in "Theoretical Computer Science", June 2019, vol. 772, p. 88-110 [*DOI :* 10.1016/J.TCS.2018.11.026], https://hal.inria.fr/hal-02413327

[16] A. CASTEIGTS, S. DUBOIS, F. PETIT, J. ROBSON.*Robustness: A new form of heredity motivated by dynamic networks*, in "Theoretical Computer Science", August 2019 [*DOI :* 10.1016/J.TCS.2019.08.008], https://hal.archives-ouvertes.fr/hal-02420608

[17] S. DEVISMES, A. LAMANI, F. PETIT, S. TIXEUIL.*Optimal torus exploration by oblivious robots*, in "Computing", September 2019, vol. 101, n^{o} 9, p. 1241-1264 [*DOI :* 10.1007/S00607-018-0595-8], https://hal.archives-ouvertes.fr/hal-02420598

[18] Y. DIEUDONNÉ, S. DOLEV, F. PETIT, M. SEGAL.*Explicit Communication Among Stigmergic Robots*, in "International Journal of Foundations of Computer Science", March 2019, vol. 30, n^{o} 02, p. 315-332 [*DOI :* 10.1142/S0129054119500072], https://hal.archives-ouvertes.fr/hal-02420394

[19] S. DUBOIS, R. GUERRAOUI, P. KUZNETSOV, F. PETIT, P. SENS.*The weakest failure detector for eventual consistency*, in "Distributed Computing", December 2019, vol. 32, n^{o} 6, p. 479-492 [*DOI :* 10.1007/s00446-016-0292-9], https://hal.inria.fr/hal-02413314

[20] M. S. MOSLI BOUKSIAA, F. TRAHAY, A. LESCOUET, G. VORON, R. DULONG, A. GUERMOUCHE, E. BRUNET, G. THOMAS.*Using differential execution analysis to identify thread interference*, in "IEEE Transactions on Parallel and Distributed Systems", December 2019, vol. 30, n^{o} 12, p. 2866-2878 [*DOI :* 10.1109/TPDS.2019.2927481], https://hal.archives-ouvertes.fr/hal-02179717

International Conferences with Proceedings

[21] K. ALTISEN, S. DEVISMES, A. DURAND.*Silence dans la forêt !*, in "ALGOTEL 2019 - 21èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications", Saint Laurent de la Cabrerisse, France, June 2019, https://hal.archives-ouvertes.fr/hal-02118549

[22] S. BOUCHARD, Y. DIEUDONNÉ, A. PELC, F. PETIT.*Trouver un trésor plus rapidement avec des conseils angulaires*, in "ALGOTEL 2019 - 21èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications", Saint Laurent de la Cabrerisse, France, June 2019, https://hal.inria.fr/hal-02118362

[23] D. CARVER, R. GOUICEM, J.-P. LOZI, J. SOPENA, B. LEPERS, W. ZWAENEPOEL, N. PALIX, J. L. LAWALL, G. MULLER.*Fork/Wait and Multicore Frequency Scaling: a Generational Clash*, in "10th Workshop on Programming Languages and Operating Systems", Huntsville, Canada, ACM Press, October 2019, p. 53-59 [*DOI :* 10.1145/3365137.3365400], https://hal.inria.fr/hal-02349987

[24] C. COURTAUD, J. SOPENA, G. MULLER, D. GRACIA.*Improving Prediction Accuracy of Memory Interferences for Multicore Platforms*, in "RTSS 2019 - 40th IEEE Real-Time Systems Symposium", Hong-Kong, China, IEEE, December 2019, https://hal.inria.fr/hal-02401625

[25] S. HAMOUDA, J. MILTHORPE.*Resilient Optimistic Termination Detection for the Async-Finish Model*, in "34th International Conference on High Performance Computing", Frankfurt, Germany, Lecture Notes in Computer Science, Springer, May 2019, vol. 11501, p. 291-311 [*DOI :* 10.1007/978-3-030-20656-7_15], https://hal.inria.fr/hal-02169496

[26] F. LANIEL, D. CARVER, J. SOPENA, F. WAJSBURT, J. LEJEUNE, M. SHAPIRO.*Highlighting the Container Memory Consolidation Problems in Linux*, in "NCA 2019 - 18th IEEE International Symposium on Network Computing and Applications", Cambridge, United States, IEEE, September 2019, p. 1-4 [*DOI :* 10.1109/NCA.2019.8935034], https://hal.archives-ouvertes.fr/hal-02424007

[27] L. Le Frioux, S. Baarir, J. Sopena, F. Kordon.*Modular and Efficient Divide-and-Conquer SAT Solver on Top of the Painless Framework*, in "TACAS 2019 - 25th International Conference on Tools and Algorithms for the Construction and Analysis of Systems", Prague, Czech Republic, T. Vojnar, L. Zhang (editors), Lecture Notes in Computer Science, April 2019, vol. 11427, p. 135-151 [*DOI :* 10.1007/978-3-030-17462-0_8], https://hal.archives-ouvertes.fr/hal-02093520

[28] *Best Paper*
E. Mauffret, D. Jeanneau, L. Arantes, P. Sens.*The Weakest Failure Detector to Solve the Mutual Exclusion Problem in an Unknown Dynamic Environment*, in "20th International Conference on Distributed Computing and Networking (ICDCN 2019)", Bangalore, India, January 2019, Extended version: https://hal.archives-ouvertes.fr/hal-01661127v3, https://hal.archives-ouvertes.fr/hal-01929224.

[29] S. S. Nair, G. Petri, M. Shapiro.*Invariant Safety for Distributed Applications*, in "Workshop on Principles and Practice of Consistency for Distributed Data (PaPoC)", Dresden, Germany, Workshop on Principles and Practice of Consistency for Distributed Data (PaPoC), March 2019, https://arxiv.org/abs/1903.02759 , https://hal.inria.fr/hal-02052012

[30] L. Teylo, L. Arantes, P. Sens, L. M. de A. Drummond.*A Hibernation Aware Dynamic Scheduler for Cloud Environments*, in "ICCP 2019 - 48th International Conference on Parallel Processing - Workshop", Kyoto, Japan, ACM Press, August 2019, p. 1-10 [*DOI :* 10.1145/3339186.3339205], https://hal.inria.fr/hal-02391614

[31] L. Teylo, L. M. de A. Drummond, L. Arantes, P. Sens.*A Bag-of-Tasks Scheduler Tolerant to Temporal Failures in Clouds*, in "SBAC-PAD 2019 - International Symposium on Computer Architecture and High Performance Computing", Campo Grande, Brazil, October 2019, https://hal.inria.fr/hal-02284965

Conferences without Proceedings

[32] A. Durand, M. Raynal, G. Taubenfeld.*Pannes de processus liées à la contention*, in "ALGOTEL 2019 - 21èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications", Saint Laurent de la Cabrerisse, France, June 2019, p. 1-4, https://hal.archives-ouvertes.fr/hal-02118917

[33] L. Feuilloley, P. Fraigniaud, J. Hirvonen, A. Paz, M. Perry.*Vérification de preuves distribuées : compromis temps-espace*, in "ALGOTEL 2019 - 21èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications", Saint Laurent de la Cabrerisse, France, June 2019, https://hal.archives-ouvertes.fr/hal-02118043

[34] S. S. Nair, G. Petri, M. Shapiro.*Proving the safety of highly-available distributed objects*, in "ESOP 2020 - 29th European Symposium on Programming", Dublin, Ireland, April 2020, https://hal.archives-ouvertes.fr/hal-02424317

Scientific Books (or Scientific Book chapters)

[35] K. Altisen, S. Devismes, S. Dubois, F. Petit.*Introduction to Distributed Self-Stabilizing Algorithms*, Synthesis Lectures on Distributed Computing Theory, Morgan & Claypool, April 2019, vol. 8, n° 1, p. 1-165 [*DOI :* 10.2200/S00908ED1V01Y201903DCT015], https://hal.archives-ouvertes.fr/hal-02178659

Research Reports

[36] K. ALTISEN, S. DEVISMES, A. DURAND, C. JOHNEN, F. PETIT.*Self-stabilizing Systems in Spite of High Dynamics*, LaBRI, CNRS UMR 5800, November 2019, https://hal.archives-ouvertes.fr/hal-02376832

Scientific Popularization

[37] N. PREGUIÇA, C. BAQUERO, M. SHAPIRO.*Conflict-free Replicated Data Types (CRDTs)*, in "Encyclopedia of Big Data Technologies", S. SAKR, A. ZOMAYA (editors), Encyclopedia of Big Data Technologies, Springer International Publishing, 2019, https://arxiv.org/abs/1805.06358 [*DOI :* 10.1007/978-3-319-63962-8_185-1], https://hal.inria.fr/hal-01793586

[38] M. SHAPIRO, P. SUTRA.*Database Consistency Models*, in "Encyclopedia of Big Data Technologies", S. SAKR, A. ZOMAYA (editors), Springer, 2019, https://arxiv.org/abs/1804.00914 [*DOI :* 10.1007/978-3-319-63962-8_203-1], https://hal.inria.fr/hal-01756780

Project-Team DYOGENE

Dynamics of Geometric Networks

IN COLLABORATION WITH: Département d'Informatique de l'Ecole Normale Supérieure

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Paris

RESEARCH CENTER
Paris

THEME
Networks and Telecommunications

Table of contents

Project-Team DYOGENE

Creation of the Project-Team: 2013 July 01

Keywords:

Computer Science and Digital Science:

A1.2.4. - QoS, performance evaluation
A6.1.4. - Multiscale modeling
A6.2.3. - Probabilistic methods
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.3. - Geometry, Topology
A8.6. - Information theory
A8.7. - Graph theory
A8.8. - Network science
A8.9. - Performance evaluation
A9.2. - Machine learning
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B4.3. - Renewable energy production
B6.2.2. - Radio technology
B6.3.4. - Social Networks

1. Team, Visitors, External Collaborators

Research Scientists

Bartlomiej Blaszczyszyn [Inria, Senior Researcher, HDR]
Marc Lelarge [Team leader, Inria, Senior Researcher, HDR]
François Baccelli [Inria, Senior Researcher, HDR]
Ana Busic [Inria, Researcher]
Christine Fricker [Inria, Researcher, HDR]
Georgina Hall [Inria, Starting Research Position, until Aug 2019]
Laurent Massoulié [Inria, Researcher, HDR]
Sean Meyn [Inria, Advanced Research Position (Inria International Chair), from May 2019 until Jul 2019]

External Collaborators

Pierre Bremaud [Emeritus]
Pierre Nicodème [Universite Sorbonne Paris-Nord]
Pierre-Louis Xech [Microsoft Research, until Mar 2019]
Holger Keeler [University of Melbourne, from Oct 2019]

Technical Staff

Walid Ghanem [Inria, Engineer, from Apr 2019 until Sep 2019]
Pierre Popineau [Inria-ERC, Engineer, from Feb 2019]
Bharath Roy Choudhury [Inria-ERC, Engineer, from Mar 2019 until Aug 2019]

PhD Students

Arnaud Cadas [PSL, PhD Student]

Michel Davydov [ENS Paris-Saclay, PhD Student, from Sep 2019]
Alexis Galland [Ministère des Armées, PhD Student]
Luca Ganassali [Inria, PhD Student, from Sep 2019]
Md Umar Hashmi [Inria, PhD Student, until Dec 2019]
Hadrien Hendrikx [Inria, PhD Student]
Sayeh Khaniha [Inria-ENS-ERC, PhD Student, from Feb 2019]
Quentin Le Gall [Orange, PhD Student]
Leo Miolane [École polytechnique, PhD Student, until Aug 2019]
Bharath Roy Choudhury [Inria-ENS-ERC PhD Student, from Sep 2019]
Sébastien Samain [Inria, PhD Student]
Ilia Shilov [Inria, PhD Student, from Nov 2019]
Ludovic Stephan [École Normale Supérieure de Paris, PhD Student]
Antoine Brochard [Huawei, PhD Student, from Sep 2018]

Post-Doctoral Fellows

Simon Coste [Inria-ERC, Post-Doctoral Fellow, from Oct 2019]
Sanket Sanjay Kalamkar [Inria-ERC, Post-Doctoral Fellow, from Mar 2019]

Visiting Scientists

Michel Davydov [ENS Paris-Saclay, from Jun 2019 until Aug 2019]
Josu Doncel [University of the Basque Country, from Jul 2019 until Sep 2019]
Natasa Dragovic [UT-Austin, from Sep 2019 until Oct 2019]
Ahmad Alammouri [UT-Austin, from April 2019 until June 2019]

Administrative Assistants

Helene Bessin Rousseau [Inria, Administrative Assistant]
Helene Milome [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

The general scientific focus of DYOGENE is on the development of network mathematics. The following theories lie within our research interest: dynamical systems, queuing theory, optimization and control, information theory, stochastic processes, random graphs, stochastic geometry.

Our theoretical developments are motivated by and applied in the context of communication networks (Internet, wireless, mobile, cellular, peer-to-peer), social and economic networks, power grids.

We collaborate with many industrial partners. Our current industrial relations involve EDF, Huawei, Microsoft, Nokia, Orange, Safran.

More specifically, the scientific focus of DYOGENE defined in 2013 was on geometric network dynamics arising in communications. By geometric networks we understand networks with a nontrivial, discrete or continuous, geometric definition of the existence of links between the nodes. In stochastic geometric networks, this definition leads to random graphs or stochastic geometric models.

A first type of geometric network dynamics is the one where the nodes or the links change over time according to an exogeneous dynamics (e.g. node motion and geometric definition of the links). We will refer to this as dynamics of geometric networks below. A second type is that where links and/or nodes are fixed but harbor local dynamical systems (in our case, stemming from e.g. information theory, queuing theory, social and economic sciences). This will be called dynamics on geometric networks. A third type is that where the dynamics of the network geometry and the local dynamics interplay. Our motivations for studying these systems stem from many fields of communications where they play a central role, and in particular: message passing algorithms; epidemic algorithms; wireless networks and information theory; device to device networking; distributed content delivery; social and economic networks, power grids.

3. Research Program

3.1. Initial research axes

The following research axes have been defined in 2013 when the project-team was created.

- Algorithms for network performance analysis, led by A. Bouillard and A. Busic.
- Stochastic geometry and information theory for wireless network, led by F. Baccelli and B. Blaszczyszyn.
- The cavity method for network algorithms, led by M. Lelarge.

Our scientific interests keep evolving. Research areas which received the most of our attention in 2019 are summarized in the following sections.

3.2. Distributed network control and smart-grids

Theory and algorithms for distributed control of networks with applications to the stabilization of power grids subject to high volatility of renewable energy production are being developed by A. Busic in collaboration with Sean Meyn [Prof. at University of Florida and Inria International Chair].

3.3. Mathematics of wireless cellular networks

A comprehensive approach involving information theory, queueing and stochastic geometry to model and analyze the performance of large cellular networks, validated and implemented by Orange is being led by B. Blaszczyszyn in collaboration with F. Baccelli and M. K. Karray [Orange Labs]. A new collaboration between the Standardization and Research Lab at Nokia Bell Labs and ERC NEMO led by F. Baccelli has been started in 2019.

3.4. High-dimensional statistical inference and distributed learning

We computed information theoretic bounds for unsupervised and semi-supervised learning and proved complexity bounds for distributed optimization of convex functions using a network of computing units.

3.5. Stochastic Geometry

In collaboration with Mir-Omid Haji-Mirsadeghi [Sharif University, Tehran, Iran] and Ali Khezeli [School of Mathematical Sciences, Tehran, Iran] F. Baccelli develops a theory of unimodular random metric spaces.

The distortion properties of unconstrained one-bit compression were analyzed by F. Baccelli in collaboration with E. O'Reilly [Caltech] using high dimensional hyperplane tessellations.

In collaboration with D. Yogeshwaran [Indian Statistical Institute, Bangalore] and J. E. Yukich [Lehigh University] B. Błaszczyszyn develops a limit theory (Laws of Large Numbers and Central Limit Theorems) for functionals of spatially correlated point processes.

4. Application Domains

4.1. Physical communication networks

Internet, wireless, mobile, cellular networks, transportation networks, distributed systems (cloud, call centers). In collaboration with Nokia Bell Labs and Orange Labs.

4.2. Abstract networks

Social interactions, human communities, economic networks.

4.3. Power grids

Energy networks. In collaboration with EDF.

5. Highlights of the Year

5.1. Highlights of the Year

ERC NEMO

ERC NEMO, led by F. Baccelli, started in January 2019; see 9.3.1.1.

Inria International Chair

Sean Meyn obtained Inria International Chair for the period 2019–2023 and joined Dyogene.

Markov Lecture

L. Massoulié gave the Markov Lecture at the Annual Meeting of the INFORMS society, in October in Seattle https://connect.informs.org/aps/participate46/markov.

6. New Software and Platforms

6.1. Platforms

6.1.1. CapRadio

Cellular network dimensioning toolbox *CapRadio* is being developed by Orange in a long-term collaboration between TREC/DYOGENE represented by B. Błaszczyszyn, and Orange Labs, represented by M. K. Karray. This year we are working on taking into account the "massive MIMO" in 5G cellular networks; see 8.1.1.

7. New Results

7.1. Distributed network control and smart-grids

1. Distributed Control of Thermostatically Controlled Loads: Kullback-Leibler Optimal Control in Continuous Time [20] The paper develops distributed control techniques to obtain grid services from flexible loads. The Individual Perspective Design (IPD) for local (load level) control is extended to piecewise deterministic and diffusion models for thermostatically controlled load models. The IPD design is formulated as an infinite horizon average reward optimal control problem, in which the reward function contains a term that uses relative entropy rate to model deviation from nominal dynamics. In the piecewise deterministic model, the optimal solution is obtained via the solution to an eigenfunction problem, similar to what is obtained in prior work. For a jump diffusion model this simple structure is absent. The structure for the optimal solution is obtained, which suggests an ODE technique for computation that is likely far more efficient than policy-or value-iteration.

2. Optimal Control of Dynamic Bipartite Matching Models [23] A dynamic bipartite matching model is given by a bipartite matching graph which determines the possible matchings between the various types of supply and demand items. Both supply and demand items arrive to the system according to a stochastic process. Matched pairs leave the system and the others wait in the queues, which induces a holding cost. We model this problem as a Markov Decision Process and study the discounted cost and the average cost case. We first consider a model with two types of supply and two types of demand items with an N-shaped matching graph. For linear cost function, we prove that an optimal matching policy gives priority to the end edges of the matching graph and is of threshold type for the diagonal edge. In addition, for the average cost problem, we compute the optimal threshold value. According to our numerical experiments, threshold-type policies perform also very well for more general bipartite graphs.

3. Kullback-Leibler-Quadratic Optimal Control of Flexible Power Demand [24] A new stochastic control methodology is introduced for distributed control, motivated by the goal of creating virtual energy storage from flexible electric loads, i.e. Demand Dispatch. In recent work, the authors have introduced Kullback-Leibler-Quadratic (KLQ) optimal control as a stochastic control methodology for Markovian models. This paper develops KLQ theory and demonstrates its applicability to demand dispatch. In one formulation of the design, the grid balancing authority simply broadcasts the desired tracking signal, and the heterogeneous population of loads ramps power consumption up and down to accurately track the signal. Analysis of the Lagrangian dual of the KLQ optimization problem leads to a menu of solution options, and expressions of the gradient and Hessian suitable for Monte-Carlo-based optimization. Numerical results illustrate these theoretical results.

4. Bike sharing systems: a new incentive rebalancing method based on spatial outliers detection [8] Since its launch, Velib' (the Bike Sharing System-BSS-in Paris) has emerged in the Parisian landscape and has been a model for similar systems in many cities. A major problem with BSS is the stations' heterogeneity caused by the attractivity of some stations located in particular areas. In this paper, we focus on spatial outliers defined as stations having a behavior significantly different from their neighboring stations. First, we propose an improved version of Moran scatterplot to exploit the similarity between neighbors, and we test it on a real dataset issued from Velib' system to identify outliers. Then, we design a new method that globally improves the resources' availability in bike stations by adapting the users' trips to the resources' availability. Results show that with a partial collaboration of the users or a limitation to the rush hours, the proposed method enhances significantly the resources' availability in Velib' system.

5. Stochastic Battery Operations using Deep Neural Networks [25] In this paper, we introduce a scenario-based optimal control framework to account for the forecast uncertainty in battery arbitrage problems. Due to the uncertainty of prices and variations of forecast errors, it is challenging for battery operators to design profitable strategies in electricity markets. Without any explicit assumption or model for electricity price forecasts' uncertainties, we generate future price scenarios via a data-driven, learning-based approach. By aiding the predictive control with such scenarios representing possible realizations of future markets, our proposed real-time controller seeks the optimal charge/discharge levels to maximize profits. Simulation results on a case-study of California-based batteries and prices show that our proposed method can bring higher profits for different battery parameters.

6. Aggregate capacity for TCLs providing virtual energy storage with cycling constraints [26] The coordination of thermostatically controlled loads (TCLs) is challenging due to the need to meet individual loads quality of service (QoS), such as indoor temperature constraints. Since these loads are usually on/off type, cycling rate is one of their QoS metrics; frequent cycling between on and off states is detrimental to them. While significant prior work has been done on the coordination of air conditioning TCLs, the question of cycling QoS has not been investigated in a principled manner. In this work we propose a method to characterize aggregate capacity of a collection of air conditioning TCLs that respects the loads cycling rate constraints (maximum number of cycles in a given time period). The development is done within the framework of randomized local control in which a load makes on/off decisions probabilistically. This characterization allows us to propose a reference planning problem to generate feasible reference trajectories for the ensemble that respect cycling constraints. The reference planning problem manifests itself in the form a Nonlinear Programming problem (NLP), that can be efficiently solved. Our proposed method is compared to previous methods in the literature that do not enforce aggregate cycling. Enforcing individual cycling constraint without taking that into account in reference generation leads to poor reference tracking.

7. Optimal Storage Arbitrage under Net Metering using Linear Programming [29] We formulate the optimal energy arbitrage problem for a piecewise linear cost function for energy storage devices using linear programming (LP). The LP formulation is based on the equivalent minimization of the epigraph. This formulation considers ramping and capacity constraints, charging and discharging efficiency losses of the storage, inelastic consumer load and local renewable generation in presence of net-metering which facilitates selling of energy to the grid and incentivizes consumers to install renewable generation and energy storage. We consider the case where the consumer loads, electricity prices, and renewable generations at different instances are uncertain. These uncertain quantities are predicted using an Auto-Regressive Moving Average

(ARMA) model and used in a model predictive control (MPC) framework to obtain the arbitrage decision at each instance. In numerical results we present the sensitivity analysis of storage performing arbitrage with varying ramping batteries and different ratio of selling and buying price of electricity.

8. Energy Storage in Madeira, Portugal: Co-optimizing for Arbitrage, Self-Sufficiency, Peak Shaving and Energy Backup [30] Energy storage applications are explored from a prosumer (consumers with generation) perspective for the island of Madeira in Portugal. These applications could also be relevant to other power networks. We formulate a convex co-optimization problem for performing arbitrage under zero feed-in tariff, increasing self-sufficiency by increasing self-consumption of locally generated renewable energy, provide peak shaving and act as a backup power source during anticipated and scheduled power outages. Using real data from Madeira we perform short and long timescale simulations in order to select end-user contract which maximizes their gains considering storage degradation based on operational cycles. We observe energy storage ramping capability decides peak shaving potential, fast ramping batteries can significantly reduce peak demand charge. The numerical experiment indicates that storage providing backup does not significantly reduce gains performing arbitrage and peak demand shaving. Furthermore, we also use AutoRegressive Moving Average (ARMA) forecasting along with Model Predic-tive Control (MPC) for real-time implementation of the proposed optimization problem in the presence of uncertainty.

9. Sensitivity to forecast errors in energy storage arbitrage for residential consumers [34] With the massive deployment of distributed energy resources, there has been an increase in the number of end consumers that own photovoltaic panels and storage systems. The optimal use of such storage when facing Time of Use (ToU) prices is directly related to the quality of the load and generation forecasts as well as the algorithm that controls the battery. The sensitivity of such control to different forecasts techniques is studied in this paper. It is shown that good and bad forecasts can result in losses in particularly bad days. Nevertheless, it is observed that performing Model Predictive Control with a simple forecast that is representative of the pasts can be profitable under different price and battery scenarios. We use real data from Pecan Street and ToU price levels with different buying and selling price for the numerical experiments.

10. Sizing and Profitability of Energy Storage for Prosumers in Madeira, Portugal [47] This paper proposes a framework to select the best-suited battery for co-optimizing for peak demand shaving, energy arbitrage and increase self-sufficiency in the context of power network in Madeira, Portugal. Feed-in-tariff for electricity network in Madeira is zero, which implies consumers with excess production should locally consume the excess generation rather than wasting it. Further, the power network operator applies a peak power contract for consumers which imposes an upper bound on the peak power seen by the power grid interfaced by energy meter. We investigate the value of storage in Madeira, using four different types of prosumers, categorized based on the relationship between their inelastic load and renewable generation. We observe that the marginal increase in the value of storage deteriorates with increase in size and ramping capabilities. We propose the use of profit per cycle per unit of battery capacity and expected payback period as indices for selecting the best-suited storage parameters to ensure profitability. This mechanism takes into account the consumption and generation patterns, profit, storage degradation, and cycle and calendar life of the battery. We also propose the inclusion of a friction coefficient in the original co-optimization formulation to increase the value of storage by reducing the operational cycles and eliminate low returning transactions.

11. Arbitrage with Power Factor Correction using Energy Storage [48] The importance of reactive power compensation for power factor (PF) correction will significantly increase with the large-scale integration of distributed generation interfaced via inverters producing only active power. In this work, we focus on co-optimizing energy storage for performing energy arbitrage as well as local power factor corrections. The joint optimization problem is non-convex, but can be solved efficiently using a McCormick relaxation along with penalty-based schemes. Using numerical simulations on real data and realistic storage profiles, we show that energy storage can correct PF locally without reducing arbitrage gains. It is observed that active and reactive power control is largely decoupled in nature for performing arbitrage and PF correction (PFC). Furthermore, we consider a stochastic online formulation of the problem with uncertain load, renewable and pricing profiles. We develop a model predictive control based storage control policy using ARMA forecast for the uncertainty. Using numerical simulations we observe that PFC is primarily governed by the size of the converter and therefore, look-ahead in time in the online setting does not affect PFC noticeably. However, arbitrage gains are more sensitive to uncertainty for batteries with faster ramp rates compared to slow ramping batteries.

12. A Utility Optimization Approach to Network Cache Design [11] In any caching system, the admission and eviction policies determine which contents are added and removed from a cache when a miss occurs. Usually, these policies are devised so as to mitigate staleness and increase the hit probability. Nonetheless, the utility of having a high hit probability can vary across contents. This occurs, for instance, when service level agreements must be met, or if certain contents are more difficult to obtain than others. In this paper, we propose utility-driven caching, where we associate with each content a utility, which is a function of the corresponding content hit probability. We formulate optimization problems where the objectives are to maximize the sum of utilities over all contents. These problems differ according to the stringency of the cache capacity constraint. Our framework enables us to reverse engineer classical replacement policies such as LRU and FIFO, by computing the utility functions that they maximize. We also develop online algorithms that can be used by service providers to implement various caching policies based on arbitrary utility functions.

13. Rapid Mixing of Dynamic Graphs with Local Evolution Rules [15] Dynamic graphs arise naturally in many contexts. In peer-to-peer networks, for instance, a participating peer may replace an existing connection with one neighbor by a new connection with a neighbor of that neighbor. Several such local rewiring rules have been proposed to ensure that peer-to-peer networks achieve good connectivity properties (e.g. high expansion) at equilibrium. However, the question of whether there exists such a rule that converges rapidly to equilibrium has remained open. In this work, we provide an affirmative answer: we exhibit a local rewiring rule that converges to equilibrium after each participating node has undergone only a number of changes that is at most poly-logarithmic in the system size. As a byproduct, we derive new results for random walks on graphs, bounding the spread of their law throughout the transient phase, i.e. prior to mixing. These rely on an extension of Cheeger's inequality, based on generalized isoperimetric constants, and may be of independent interest.

7.2. Reinforcement learning

14. On Matrix Momentum Stochastic Approximation and Applications to Q-learning [27] Stochastic approximation (SA) algorithms are recursive techniques used to obtain the roots of functions that can be expressed as expectations of a noisy parameterized family of functions. In this paper two new SA algorithms are introduced: 1) PolSA, an extension of Polyak's momentum technique with a specially designed matrix momentum, and 2) NeSA, which can either be regarded as a variant of Nesterov's acceleration method, or a simplification of PolSA. The rates of convergence of SA algorithms is well understood. Under special conditions, the mean square error of the parameter estimates is bounded by $\sigma^2/n + o(1/n)$, where $\sigma^2 \geq 0$ is an identifiable constant. If these conditions fail, the rate is typically sub-linear. There are two well known SA algorithms that ensure a linear rate, with minimal value of variance, σ^2: the Ruppert-Polyak averaging technique, and the stochastic Newton-Raphson (SNR) algorithm. It is demonstrated here that under mild technical assumptions, the PolSA algorithm also achieves this optimality criteria. This result is established via novel coupling arguments: It is shown that the parameter estimates obtained from the PolSA algorithm couple with those of the optimal variance (but computationally more expensive) SNR algorithm, at a rate $O(1/n^2)$. The newly proposed algorithms are extended to a reinforcement learning setting to obtain new Q-learning algorithms, and numerical results confirm the coupling of PolSA and SNR.

15. Zap Q-Learning - A User's Guide [28] There are two well known Stochastic Approximation techniques that are known to have optimal rate of convergence (measured in terms of asymptotic variance): the Stochastic Newton-Raphson (SNR) algorithm (a matrix gain algorithm that resembles the deterministic Newton-Raphson method), and the Ruppert-Polyak averaging technique. This paper surveys new applications of these concepts for Q-learning: (i)The Zap Q-Learning algorithm was introduced by the authors in a NIPS 2017 paper. It is based on a variant of SNR, designed to more closely mimic its deterministic cousin. The algorithm has optimal rate of convergence under general assumptions, and showed astonishingly quick convergence in numerical examples. These algorithms are surveyed and illustrated with numerical examples. A potential difficulty in implementation of the Zap-Q-Learning algorithm is the matrix inversion required in each iteration. (ii)Remedies are proposed based on stochastic approximation variants of two general deterministic techniques:

Polyak's momentum algorithms and Nesterov's acceleration technique. Provided the hyper-parameters are chosen with care, the performance of these algorithms can be comparable to the Zap algorithm, while computational complexity per iteration is far lower.

16. Zap Q-Learning With Nonlinear Function Approximation [44] The Zap stochastic approximation (SA) algorithm was introduced recently as a means to accelerate convergence in reinforcement learning algorithms. While numerical results were impressive, stability (in the sense of boundedness of parameter estimates) was established in only a few special cases. This class of algorithms is generalized in this paper, and stability is established under very general conditions. This general result can be applied to a wide range of algorithms found in reinforcement learning. Two classes are considered in this paper: (i)The natural generalization of Watkins' algorithm is not always stable in function approximation settings. Parameter estimates may diverge to infinity even in the *linear* function approximation setting with a simple finite state-action MDP. Under mild conditions, the Zap SA algorithm provides a stable algorithm, even in the case of *nonlinear* function approximation. (ii) The GQ algorithm of Maei et. al. 2010 is designed to address the stability challenge. Analysis is provided to explain why the algorithm may be very slow to converge in practice. The new Zap GQ algorithm is stable even for nonlinear function approximation.

17. Zap Q-Learning for Optimal Stopping Time Problems [43] The objective in this paper is to obtain fast converging reinforcement learning algorithms to approximate solutions to the problem of discounted cost optimal stopping in an irreducible, uniformly ergodic Markov chain, evolving on a compact subset of IR^n. We build on the dynamic programming approach taken by Tsitsikilis and Van Roy, wherein they propose a Q-learning algorithm to estimate the optimal state-action value function, which then defines an optimal stopping rule. We provide insights as to why the convergence rate of this algorithm can be slow, and propose a fast-converging alternative, the "Zap-Q-learning" algorithm, designed to achieve optimal rate of convergence. For the first time, we prove the convergence of the Zap-Q-learning algorithm under the assumption of linear function approximation setting. We use ODE analysis for the proof, and the optimal asymptotic variance property of the algorithm is reflected via fast convergence in a finance example.

7.3. Mathematics of wireless cellular networks

18. Performance analysis of cellular networks with opportunistic scheduling using queueing theory and stochastic geometry [6] Combining stochastic geometric approach with some classical results from queuing theory, in this paper we propose a comprehensive framework for the performance study of large cellular networks featuring opportunistic scheduling. Rapid and verifiable with respect to real data, our approach is particularly useful for network dimensioning and long term economic planning. It is based on a detailed network model combining an information-theoretic representation of the link layer, a queuing-theoretic representation of the users' scheduler, and a stochastic-geometric representation of the signal propagation and the network cells. It allows one to evaluate principal characteristics of the individual cells, such as loads (defined as the fraction of time the cell is not empty), the mean number of served users in the steady state, and the user throughput. A simplified Gaussian approximate model is also proposed to facilitate study of the spatial distribution of these metrics across the network. The analysis of both models requires only simulations of the point process of base stations and the shadowing field to estimate the expectations of some stochastic-geometric functionals not admitting explicit expressions. A key observation of our approach , bridging spatial and temporal analysis, relates the SINR distribution of the typical user to the load of the typical cell of the network. The former is a static characteristic of the network related to its spectral efficiency while the latter characterizes the performance of the (generalized) processor sharing queue serving the dynamic population of users of this cell.

19. Two-tier cellular networks for throughput maximization of static and mobile users [10] In small cell networks, high mobility of users results in frequent handoff and thus severely restricts the data rate for mobile users. To alleviate this problem, we propose to use heterogeneous, two-tier network structure where static users are served by both macro and micro base stations, whereas the mobile (i.e., moving) users are served only by macro base stations having larger cells; the idea is to prevent frequent data outage for mobile users due to handoff. We use the classical two-tier Poisson network model with different transmit powers, assume independent Poisson process of static users and doubly stochastic Poisson process of mobile users moving at a constant speed along infinite straight lines generated by a Poisson line process. Using stochastic geometry, we calculate the average downlink data rate of the typical static and mobile (i.e., moving) users, the latter accounted for handoff outage periods. We consider also the average throughput of these two types of users defined as their average data rates divided by the mean total number of users co-served by the same base station. We find that if the density of a homogeneous network and/or the speed of mobile users is high, it is advantageous to let the mobile users connect only to some optimal fraction of BSs to reduce the frequency of handoffs during which the connection is not assured. If a heterogeneous structure of the network is allowed, one can further jointly optimize the mean throughput of mobile and static users by appropriately tuning the powers of micro and macro base stations subject to some aggregate power constraint ensuring unchanged mean data rates of static users via the network equivalence property.

20. Location Aware Opportunistic Bandwidth Sharing between Static and Mobile Users with Stochastic Learning in Cellular Networks [9] We consider location-dependent opportunistic bandwidth sharing between static and mobile downlink users in a cellular network. Each cell has some fixed number of static users. Mobile users enter the cell, move inside the cell for some time and then leave the cell. In order to provide higher data rate to mobile users, we propose to provide higher bandwidth to the mobile users at favourable times and locations, and provide higher bandwidth to the static users in other times. We formulate the problem as a long run average reward Markov decision process (MDP) where the per-step reward is a linear combination of instantaneous data volumes received by static and mobile users, and find the optimal policy. The transition structure of this MDP is not known in general. To alleviate this issue, we propose a learning algorithm based on single timescale stochastic approximation. Also, noting that the unconstrained MDP can be used to solve a constrained problem, we provide a learning algorithm based on multi-timescale stochastic approximation. The results are extended to address the issue of fair bandwidth sharing between the two classes of users. Numerical results demonstrate performance improvement by our scheme, and also the trade-off between performance gain and fairness.

21. Per-Link Reliability and Rate Control: Two Facets of the SIR Meta Distribution [13] The meta distribution (MD) of the signal-to-interference ratio (SIR) provides fine-grained reliability performance in wireless networks modeled by point processes. In particular, for an ergodic point process, the SIR MD yields the distribution of the per-link reliability for a target SIR. Here we reveal that the SIR MD has a second important application, which is rate control. Specifically, we calculate the distribution of the SIR threshold (equivalently, the distribution of the transmission rate) that guarantees each link a target reliability and show its connection to the distribution of the per-link reliability. This connection also permits an approximate calculation of the SIR MD when only partial (local) information about the underlying point process is available.

22. Simple Approximations of the SIR Meta Distribution in General Cellular Networks [14] Compared to the standard success (coverage) probability , the meta distribution of the signal-to-interference ratio (SIR) provides much more fine-grained information about the network performance. We consider general heterogeneous cellular networks (HCNs) with base station tiers modeled by arbitrary stationary and ergodic non-Poisson point processes. The exact analysis of non-Poisson network models is notoriously difficult, even in terms of the standard success probability, let alone the meta distribution. Hence we propose a simple approach to approximate the SIR meta distribution for non-Poisson networks based on the ASAPPP ("approximate SIR analysis based on the Poisson point process") method. We prove that the asymptotic horizontal gap G_0 between its standard success probability and that for the Poisson point process exactly characterizes the gap between the bth moment of the conditional success probability, as the SIR threshold goes to 0. The gap G_0 allows two simple approximations of the meta distribution for general HCNs: 1) the per-tier approximation by applying the shift G_0 to each tier and 2) the effective gain approximation by directly shifting the meta distribution for the homogeneous independent Poisson network. Given the generality of the model considered and the fine-grained nature of the meta distribution, these approximations work surprisingly well.

23. Interference Queueing Networks [16] This work features networks of coupled processor sharing queues in the Euclidean space, where customers arrive according to independent Poisson point processes at every queue, are served, and then leave the network. The coupling is through service rates. In any given queue, this rate is inversely proportional the interference seen by this queue, which is determined by the load in neighboring queues, attenuated by some distance-based path-loss function. The main focus is on the infinite grid network and translation invariant path-loss case. The model is a discrete version of a spatial birth and death process where customers arrive to the Euclidean space according to Poisson rain and leave it when they have transferred an exponential file, assuming that the instantaneous rate of each transfer is determined through information theory by the signal to interference and noise ratio experienced by the user. The stability condition is identified. The minimal stationary regime is built using coupling from the past techniques. The mean queue size of this minimal stationary regime is determined in closed form using the rate conservation principle of Palm calculus. When the stability condition holds, for all bounded initial conditions, there is weak convergence to this minimal stationary regime; however, there exist translation invariant initial conditions for which all queue sizes converge to infinity.

24. Statistical learning of geometric characteristics of wireless networks [19] Motivated by the prediction of cell loads in cellular networks, we formulate the following new, fundamental problem of statistical learning of geometric marks of point processes: An unknown marking function, depending on the geometry of point patterns, produces characteristics (marks) of the points. One aims at learning this function from the examples of marked point patterns in order to predict the marks of new point patterns. To approximate (interpolate) the marking function, in our baseline approach, we build a statistical regression model of the marks with respect some local point distance representation. In a more advanced approach, we use a global data representation via the scattering moments of random measures, which build informative and stable to deformations data representation, already proven useful in image analysis and related application domains. In this case, the regression of the scattering moments of the marked point patterns with respect to the non-marked ones is combined with the numerical solution of the inverse problem, where the marks are recovered from the estimated scattering moments. Considering some simple, generic marks, often appearing in the modeling of wireless networks, such as the shot-noise values, nearest neighbour distance, and some characteristics of the Voronoi cells, we show that the scattering moments can capture similar geometry information as the baseline approach, and can reach even better performance, especially for non-local marking functions. Our results motivate further development of statistical learning tools for stochastic geometry and analysis of wireless networks, in particular to predict cell loads in cellular networks from the locations of base stations and traffic demand.

25. Determinantal thinning of point processes with network learning applications [21] A new type of dependent thinning for point processes in continuous space is proposed, which leverages the advantages of determinantal point processes defined on finite spaces and, as such, is particularly amenable to statistical, numerical, and simulation techniques. It gives a new point process that can serve as a network model exhibiting repulsion. The properties and functions of the new point process, such as moment measures, the Laplace functional, the void probabilities, as well as conditional (Palm) characteristics can be estimated accurately by simulating the underlying (non-thinned) point process, which can be taken, for example, to be Poisson. This is in contrast (and preference to) finite Gibbs point processes, which, instead of thinning, require weighting the Poisson realizations, involving usually intractable normalizing constants. Models based on determinantal point processes are also well suited for statistical (supervised) learning techniques, allowing the models to be fitted to observed network patterns with some particular geometric properties. We illustrate this approach by imitating with determinantal thinning the well-known Matérn II hard-core thinning, as well as a soft-core thinning depending on nearest-neighbour triangles. These two examples demonstrate how the proposed approach can lead to new, statistically optimized, probabilistic transmission scheduling schemes.

26. Analyzing LoRa long-range, low-power, wide-area networks using stochastic geometry [22] In this paper we present a simple, stochastic-geometric model of a wireless access network exploiting the LoRA (Long Range) protocol, which is a non-expensive technology allowing for long-range, single-hop connectivity for the Internet of Things. We assume a space-time Poisson model of packets transmitted by LoRA nodes to a fixed base station. Following previous studies of the impact of interference, we assume that a given packet is successfully received when no interfering packet arrives with similar power before the given packet payload phase. This is as a consequence of LoRa using different transmission rates for different link budgets (transmissions with smaller received powers use larger spreading factors) and LoRa intra-technology interference treatment. Using our model, we study the scaling of the packet reception probabilities per link budget as a function of the spatial density of nodes and their rate of transmissions. We consider both the parameter values recommended by the LoRa provider, as well as proposing LoRa tuning to improve the equality of performance for all link budgets. We also consider spatially non-homogeneous distributions of LoRa nodes. We show also how a fair comparison to non-slotted Aloha can be made within the same framework.

27. Reliability and Local Delay in Wireless Networks: Does Bandwidth Partitioning Help? [33] In a series of papers initiated through a collaboration with Nokia Bell Labs, we study the effect of bandwidth partitioning (BWP) on the reliability and delay performance in infrastructureless wireless networks. The reliability performance is characterized by the density of concurrent transmissions that satisfy a certain reliability (outage) constraint and the delay performance by so-called local delay, defined as the average number of time slots required to successfully transmit a packet. We concentrate on the ultrareliable regime where the target outage probability is close to 0. BWP has two conflicting effects: while the interference is reduced as the concurrent transmissions are divided over multiple frequency bands, the signal-to-interference ratio (SIR) requirement is increased due to smaller allocated bandwidth if the data rate is to be kept constant. Instead, if the SIR requirement is to be kept the same, BWP reduces the data rate and in turn increases the local delay. For these two approaches with adaptive and fixed SIR requirements, we derive closed-form expressions of the local delay and the maximum density of reliable transmissions in the ultrareliable regime. Our analysis shows that, in the ultrareliable regime, BWP leads to the reliability-delay tradeoff.

28. The Influence of Canyon Shadowing on Device-to-Device Connectivity in Urban Scenario [35] In this work, we use percolation theory to study the feasibility of large-scale connectivity of relay-augmented device-to-device (D2D) networks in an urban scenario, featuring a haphazard system of streets and canyon shadowing allowing only for line-of-sight (LOS) communications in a limited finite range. We use a homogeneous Poisson-Voronoi tessellation (PVT) model of streets with homogeneous Poisson users (devices) on its edges and independent Bernoulli relays on the vertices. Using this model, we demonstrated the existence of a minimal threshold for relays below which large-scale connectivity of the network is not possible, regardless of all other network parameters. Through simulations, we estimated this threshold to 71.3%. Moreover, if the mean street length is not larger than some threshold (predicted to 74.3% of the communication range; which might be the case in a typical urban scenario) then any (whatever small) density of users can be compensated by equipping more crossroads with relays. Above this latter threshold, good connectivity requires some minimal density of users, compensated by the relays in a way we make explicit. The existence of the above regimes brings interesting qualitative arguments to the discussion on the possible D2D deployment scenarios.

29. Relay-assisted Device-to-Device Networks: Connectivity and Uberization Opportunities [46] It has been shown that deploying device-to-device (D2D) networks in urban environments requires equipping a considerable proportion of crossroads with relays. This represents a necessary economic investment for an operator. In this work, we tackle the problem of the economic feasibility of such relay-assisted D2D networks. First, we propose a stochastic model taking into account a positive surface for streets and crossroads, thus allowing for a more realistic estimation of the minimal number of needed relays. Secondly, we introduce a cost model for the deployment of relays, allowing one to study operators' D2D deployment strategies. We investigate the example of an uberizing neo-operator willing to set up a network entirely relying on D2D and show that a return on the initial investment in relays is possible in a realistic period of time, even if the network is funded by a very low revenue per D2D user. Our results bring quantitative arguments to the discussion on possible uberization scenarios of telecommunications networks.

30. Continuum Line-of-Sight Percolation on Poisson-Voronoi Tessellations [45] In this work, we study a new model for continuum line-of-sight percolation in a random environment given by a Poisson-Voronoi tessellation. The edges of this tessellation are the support of a Cox point process, while the vertices are the support of a Bernoulli point process. Taking the superposition of these two processes, two points of are linked by an edge if and only if they are sufficiently close and located on the same edge of the supporting tessellation. We study the percolation of the random graph arising from this construction and prove that a subcritical phase as well as a supercritical phase exist under general assumptions. Our proofs are based on a renormalization argument with some notion of stabilization and asymptotic essential connectedness to investigate continuum percolation for Cox point processes. We also give numerical estimates of the critical parameters of the model. Our model can be seen as a good candidate for modelling telecommunications networks in a random environment with obstructive conditions for signal propagation.

7.4. High-dimensional statistical inference

31. Discrete Mean Field Games: Existence of Equilibria and Convergence [12] We consider mean field games with discrete state spaces (called discrete mean field games in the following) and we analyze these games in continuous and discrete time, over finite as well as infinite time horizons. We prove the existence of a mean field equilibrium assuming continuity of the cost and of the drift. These conditions are more general than the existing papers studying finite state space mean field games. Besides, we also study the convergence of the equilibria of N -player games to mean field equilibria in our four settings. On the one hand, we define a class of strategies in which any sequence of equilibria of the finite games converges weakly to a mean field equilibrium when the number of players goes to infinity. On the other hand, we exhibit equilibria outside this class that do not converge to mean field equilibria and for which the value of the game does not converge. In discrete time this non-convergence phenomenon implies that the Folk theorem does not scale to the mean field limit.

32. Modularity-based Sparse Soft Graph Clustering [32] Clustering is a central problem in machine learning for which graph-based approaches have proven their efficiency. In this paper, we study a relaxation of the modularity maxi-mization problem, well-known in the graph partitioning literature. A solution of this relaxation gives to each element of the dataset a probability to belong to a given cluster, whereas a solution of the standard modularity problem is a partition. We introduce an efficient optimization algorithm to solve this relaxation, that is both memory efficient and local. Furthermore, we prove that our method includes, as a special case, the Louvain optimization scheme, a state-of-the-art technique to solve the traditional modularity problem. Experiments on both synthetic and real-world data illustrate that our approach provides meaningful information on various types of data.

33. Phase Transitions, Optimal Errors and Optimality of Message-Passing in Generalized Linear Models [41] We consider generalized linear models where an unknown n-dimensional signal vector is observed through the successive application of a random matrix and a non-linear (possibly probabilistic) componentwise function. We consider the models in the high-dimensional limit, where the observation consists of m points, and $m/n \to \alpha$ where α stays finite in the limit $m, n \to \infty$. This situation is ubiquitous in applications ranging from supervised machine learning to signal processing. A substantial amount of work suggests that both the inference and learning tasks in these problems have sharp intrinsic limitations when the available data become too scarce or too noisy. Here, we provide rigorous asymptotic predictions for these thresholds through the proof of a simple expression for the mutual information between the observations and the signal. Thanks to this expression we also obtain as a consequence the optimal value of the generalization error in many statistical learning models of interest, such as the teacher-student binary perceptron, and introduce several new models with remarquable properties. We compute these thresholds (or "phase transitions") using ideas from statistical physics that are turned into rigorous methods thanks to a new powerful smart-path interpolation technique called the stochastic interpolation method, which has recently been introduced by two of the authors. Moreover we show that a polynomial-time algorithm refered to as generalized approximate message-passing reaches the optimal generalization performance for a large set of parameters in these problems. Our results clarify the difficulties and challenges one has to face when solving complex high-dimensional statistical problems.

34. Efficient inference in stochastic block models with vertex labels [18] We study the stochastic block model with two communities where vertices contain side information in the form of a vertex label. These vertex labels may have arbitrary label distributions, depending on the community memberships. We analyze a version of the popular belief propagation algorithm. We show that this algorithm achieves the highest accuracy possible whenever a certain function of the network parameters has a unique fixed point. When this function has multiple fixed points, the belief propagation algorithm may not perform optimally, where we conjecture that a non-polynomial time algorithm may perform better than BP. We show that increasing the information in the vertex labels may reduce the number of fixed points and hence lead to optimality of belief propagation.

35. Planting trees in graphs, and finding them back [36] In this paper we study detection and reconstruction of planted structures in Erdős-Rényi random graphs. Motivated by a problem of communication security, we focus on planted structures that consist in a tree graph. For planted line graphs, we establish the following phase diagram. In a low density region where the average degree λ of the initial graph is below some critical value $\lambda_c = 1$, detection and reconstruction go from impossible to easy as the line length K crosses some critical value $f(\lambda)\ln(n)$, where n is the number of nodes in the graph. In the high density region $\lambda > \lambda_c$, detection goes from impossible to easy as K goes from $o(\sqrt{n})$ to $\omega(\sqrt{n})$, and reconstruction remains impossible so long as $K = o(n)$. For D-ary trees of varying depth h and $2 \leq D \leq O(1)$, we identify a low-density region $\lambda < \lambda_D$, such that the following holds. There is a threshold $h* = g(D)\ln(\ln(n))$ with the following properties. Detection goes from feasible to impossible as h crosses $h*$. We also show that only partial reconstruction is feasible at best for $h \geq h*$. We conjecture a similar picture to hold for D-ary trees as for lines in the high-density region $\lambda > \lambda_D$, but confirm only the following part of this picture: Detection is easy for D-ary trees of size $\omega(\sqrt{n})$, while at best only partial reconstruction is feasible for D-ary trees of any size $o(n)$. These results are in contrast with the corresponding picture for detection and reconstruction of *low rank* planted structures, such as dense subgraphs and block communities: We observe a discrepancy between detection and reconstruction, the latter being impossible for a wide range of parameters where detection is easy. This property does not hold for previously studied low rank planted structures.

36. Robustness of spectral methods for community detection [37] This work is concerned with community detection. Specifically, we consider a random graph drawn according to the stochastic block model: its vertex set is partitioned into blocks, or communities, and edges are placed randomly and independently of each other with probability depending only on the communities of their two endpoints. In this context, our aim is to recover the community labels better than by random guess, based only on the observation of the graph.

In the sparse case, where edge probabilities are in $O(1/n)$, we introduce a new spectral method based on the distance matrix D, where $D_{ij} = 1$ iff the graph distance between i and j, noted $d(i,j)$ is equal to ℓ. We show that when $\ell \sim c\log(n)$ for carefully chosen c, the eigenvectors associated to the largest eigenvalues of D provide enough information to perform non-trivial community recovery with high probability, provided we are above the so-called Kesten-Stigum threshold. This yields an efficient algorithm for community detection, since computation of the matrix D can be done in $O(n^{1+\kappa})$ operations for a small constant κ.

We then study the sensitivity of the eigendecomposition of D when we allow an adversarial perturbation of the edges of G. We show that when the considered perturbation does not affect more than $O(n^{\varepsilon})$ vertices for some small $\varepsilon > 0$, the highest eigenvalues and their corresponding eigenvectors incur negligible perturbations, which allows us to still perform efficient recovery.

Our proposed spectral method therefore: i) is robust to larger perturbations than prior spectral methods, while semi-definite programming (or SDP) methods can tolerate yet larger perturbations; ii) achieves non-trivial detection down to the KS threshold, which is conjectured to be optimal and is beyond reach of existing SDP approaches; iii) is faster than SDP approaches.

7.5. Distributed optimization for machine learning

37. Optimal Convergence Rates for Convex Distributed Optimization in Networks [17] This work proposes a theoretical analysis of distributed optimization of convex functions using a network of computing units. We investigate this problem under two communication schemes (centralized and decentralized) and four classical regularity assumptions: Lipschitz continuity, strong convexity, smoothness, and a combination of strong convexity and smoothness. Under the decentralized communication scheme, we provide matching upper and lower bounds of complexity along with algorithms achieving this rate up to logarithmic constants.

For non-smooth objective functions, while the dominant term of the error is in $O(1/\sqrt{t})$, the structure of the communication network only impacts a second-order term in $O(1/t)$, where tt is time. In other words, the error due to limits in communication resources decreases at a fast rate even in the case of non-strongly convex objective functions. Such a convergence rate is achieved by the novel multi-step primal-dual (MSPD) algorithm. Under the centralized communication scheme, we show that the naive distribution of standard optimization algorithms is optimal for smooth objective functions, and provide a simple yet efficient algorithm called distributed randomized smoothing (DRS) based on a local smoothing of the objective function for non-smooth functions. We then show that DRS is within a $d^{1/4}$ multiplicative factor of the optimal convergence rate, where d is the underlying dimension.

38. Accelerated Decentralized Optimization with Local Updates for Smooth and Strongly Convex Objectives [31] In this paper, we study the problem of minimizing a sum of smooth and strongly convex functions split over the nodes of a network in a decentralized fashion. We propose the algorithm $ESDACD$, a decentralized accelerated algorithm that only requires local synchrony. Its rate depends on the condition number κ of the local functions as well as the network topology and delays. Under mild assumptions on the topology of the graph, $ESDACD$ takes a time $O((\tau_{\max} + \Delta_{\max})\sqrt{\kappa/\gamma}\ln(\epsilon^{-1}))$ to reach a precision ϵ where γ is the spectral gap of the graph, $\tau_{\max}$ the maximum communication delay and $\Delta_{\max}$ the maximum computation time. Therefore, it matches the rate of $SSDA$, which is optimal when $\tau_{\max} = \Omega(\Delta_{\max})$. Applying $ESDACD$ to quadratic local functions leads to an accelerated randomized gossip algorithm of rate $O(\sqrt{\theta_{\text{gossip}}/n})$ where θ_{gossip} is the rate of the standard randomized gossip. To the best of our knowledge, it is the first asynchronous gossip algorithm with a provably improved rate of convergence of the second moment of the error. We illustrate these results with experiments in idealized settings.

39. An Accelerated Decentralized Stochastic Proximal Algorithm for Finite Sums [49] Modern large-scale finite-sum optimization relies on two key aspects: distribution and stochastic updates. For smooth and strongly convex problems, existing decentralized algorithms are slower than modern accelerated variance-reduced stochastic algorithms when run on a single machine, and are therefore not efficient. Centralized algorithms are fast, but their scaling is limited by global aggregation steps that result in communication bottlenecks. In this work, we propose an efficient **A**ccelerated, **D**ecentralized stochastic algorithm for **F**inite**S**ums named ADFS, which uses local stochastic proximal updates and randomized pairwise communications between nodes. On machines, ADFS learns from samples in the same time it takes optimal algorithms to learn from samples on one machine. This scaling holds until a critical network size is reached, which depends on communication delays, on the number of samples , and on the network topology. We provide a theoretical analysis based on a novel augmented graph approach combined with a precise evaluation of synchronization times and an extension of the accelerated proximal coordinate gradient algorithm to arbitrary sampling. We illustrate the improvement of ADFS over state-of-the-art decentralized approaches with experiments.

7.6. Stochastic Geometry

40. On the Dimension of Unimodular Discrete Spaces, Part I: Definitions and Basic Properties [39] This work introduces two new notions of dimension, namely the *unimodular Minkowski and Hausdorff dimensions*, which are inspired from the classical analogous notions. These dimensions are defined for *unimodular discrete spaces*, introduced in this work, which provide a common generalization to stationary point processes under their Palm version and unimodular random rooted graphs. The use of unimodularity in the definitions of dimension is novel. Also, a toolbox of results is presented for the analysis of these dimensions. In particular, analogues of Billingsley's lemma and Frostman's lemma are presented. These lemmas are instrumental in deriving upper bounds on dimensions, whereas lower bounds are obtained from specific coverings. The notions of unimodular Hausdorff measure and unimodular dimension function are also introduced. This toolbox is used to connect the unimodular dimensions to various other notions such as growth rate, scaling limits, discrete dimension and amenability. It is also used to analyze the dimensions of a set of examples pertaining to point processes, branching processes, random graphs, random walks, and self-similar discrete random spaces.

41. On the Dimension of Unimodular Discrete Spaces, Part II: Relations with Growth Rate [40] The notions of unimodular Minkowski and Hausdorff dimensions are defined in [39] for unimodular random discrete metric spaces. This work is focused on the connections between these notions and the polynomial growth rate of the underlying space. It is shown that bounding the dimension is closely related to finding suitable equivariant weight functions (i.e., measures) on the underlying discrete space. The main results are unimodular versions of the mass distribution principle and Billingsley's lemma, which allow one to derive upper bounds on the unimodular Hausdorff dimension from the growth rate of suitable equivariant weight functions. Also, a unimodular version of Frostman's lemma is provided, which shows that the upper bound given by the unimodular Billingsley lemma is sharp. These results allow one to compute or bound both types of unimodular dimensions in a large set of examples in the theory of point processes, unimodular random graphs, and self-similarity. Further results of independent interest are also presented, like a version of the max-flow min-cut theorem for unimodular one-ended trees.

42. Doeblin trees [4] This work is centered on the random graph generated by a Doeblin-type coupling of discrete time processes on a countable state space whereby when two paths meet, they merge. This random graph is studied through a novel subgraph, called a bridge graph, generated by paths started in a fixed state at any time. The bridge graph is made into a unimodular network by marking it and selecting a root in a specified fashion. The unimodularity of this network is leveraged to discern global properties of the larger Doeblin graph. Bi-recurrence, i.e., recurrence both forwards and backwards in time, is introduced and shown to be a key property in uniquely distinguishing paths in the Doeblin graph, and also a decisive property for Markov chains indexed by $\mathbb{Z}$. Properties related to simulating the bridge graph are also studied.

43. The Stochastic Geometry of Unconstrained One-Bit Compression [5] A stationary stochastic geometric model is proposed for analyzing the data compression method used in one-bit compressed sensing. The data set is an unconstrained stationary set, for instance all of $\mathbb{R}^n$ or a stationary Poisson point process in $\mathbb{R}^n$. It is compressed using a stationary and isotropic Poisson hyperplane tessellation, assumed independent of the data. That is, each data point is compressed using one bit with respect to each hyperplane, which is the side of the hyperplane it lies on. This model allows one to determine how the intensity of the hyperplanes must scale with the dimension n to ensure sufficient separation of different data by the hyperplanes as well as sufficient proximity of the data compressed together. The results have direct implications in compressed sensing and in source coding.

44. Limit theory for geometric statistics of point processes having fast decay of correlations [7] We develop a limit theory (Laws of Large Numbers and Central Limit Theorems) for functionals of spatially correlated point processes. The "strength" of data correlation is captured and controlled by the speed of decay of the additive error in the asymptotic factorization the correlation functions, when the separation distance increases. In this way, the classical theory of Poisson and Bernoulli processes is extended to a larger class of data inputs, such as determinantal point processes with fast decreasing kernels, including the α-Ginibre ensembles, permanental point processes as well as the zero set of Gaussian entire functions. Both linear (U-statistics) and non-linear geometric statistics (such as clique counts, the number of Morse critical points, intrinsic volumes of the Boolean model, and total edge length of the k-nearest neighbor graph) are considered.

7.7. Information theory

45. Error Exponents for MAC Channelss [3] This work analyzes a class of Multiple Access Channels (MAC) where the sum of the dimensions of the transmitted signals matches that of the received signal. This channel is a classical object of information theory in the power constrained case. We first focus on the Poltyrev regime, namely the case without power constraint. Using point process techniques, we derive the capacity under general stationarity and ergodicity noise assumptions as well as a representation of the error probability. We use this to derive bounds on the error exponent in the Gaussian case. This also leads to new results on the power constrained error exponents.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. CRE with Orange

Two year contract titled *Taking into account the "massive MIMO" in the assessment of QoS and the dimensioning of 5G cellular networks* between Inria and Orange Labs started 2018. It is a part of a long-term collaboration between TREC/DYOGENE, represented by B. Błaszczyszyn and Orange Labs, represented by M. K. Karray on the development of analytic tools and methods allowing one to capture macroscopic relation between antennas roll-out, frequency allocation, volume of traffic carried on the network and quality of service parameters such as the average and the variation of bandwidth available to end users. This work addresses crucial technical and economical issues related to the operator core business, particularly related to the current evolution of the cellular network technology (4G⇒5G). The developed solutions are implemented by Orange Labs in the internal toolbox *CapRadio* (see 6.1.1) and used by the Direction of Regulatory Affairs of Orange.

8.1.2. Contract with EDF

Collaborative research in the area of demand dispatch of flexible loads. PI : A. Busic.

8.1.3. CIFRE with Orange

Contract with Orange started in 2017 and continued in 2018 for the co-advising by B. Błaszczyszyn of a PhD student of Orange, Quentin Le Gall.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Laboratory of Information, Networking and Communication Sciences (LINCS)

Dyogene participates in LINCS https://www.lincs.fr, a research centre co-founded by Inria, Institut Mines-Télécom, UPMC and Alcatel-Lucent Bell Labs (currently Nokia Bell Labs) dedicated to research and innovation in the domains of future information and communication networks, systems and services.

9.1.2. PGMO

Dyogene participates in the PGMO (Gaspard Monge Program for Optimization, operations research, and their interactions with data science) via the project a 2 year project "Distributed control of flexible loads" funded through the ICODE/IROE call. This is a collaborative project between University Paris-Sud (PI: Gilles Stoltz) and Inria (PI: Ana Busic).

9.2. National Initiatives

9.2.1. GdR GeoSto

Members of Dyogene participate in Research Group GeoSto (Groupement de recherche, GdR 3477) http://gdr-geostoch.math.cnrs.fr/ on Stochastic Geometry led by and David Coupier [Université de Valenciennes].

This is a collaboration framework for all French research teams working in the domain of spatial stochastic modeling, both on theory development and in applications.

9.2.2. GdR RO

Members of Dyogene participate in GdR-RO (Recherche Opérationelle; GdR CNRS 3002), http://gdrro.lip6.fr/, working group COSMOS (Stochastic optimization and control, modeling and simulation), lead by A. Busic and E. Hyon (LIP 6); http://gdrro.lip6.fr/?q=node/78

9.2.3. ANR JCJC PARI

Probabilistic Approach for Renewable Energy Integration: Virtual Storage from Flexible Loads. The project started in January 2017. PI — A. Bušić. This project is motivated by current and projected needs of a power grid with significant renewable energy integration. Renewable energy sources such as wind and solar have a high degree of unpredictability and time variation, which makes balancing demand and supply challenging. There is an increased need for ancillary services to smooth the volatility of renewable power. In the absence of large, expensive batteries, we may have to increase our inventory of responsive fossil-fuel generators, negating the environmental benefits of renewable energy. The proposed approach addresses this challenge by harnessing the inherent flexibility in demand of many types of loads. The objective of the project is to develop decentralized control for automated demand dispatch, that can be used by grid operators as ancillary service to regulate demand-supply balance at low cost. We call the resource obtained from these techniques virtual energy storage (VES). Our goal is to create the necessary ancillary services for the grid that are environmentally friendly, that have low cost and that do not impact the quality of service (QoS) for the consumers. Besides respecting the needs of the loads, the aim of the project is to design local control solutions that require minimal communications from the loads to the centralized entity. This is possible through a systems architecture that includes the following elements: i) local control at each load based on local measurements combined with a grid-level signal; ii) frequency decomposition of the regulation signal based on QoS and physical constraints for each class of loads.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. NEMO

NEMO, NEtwork MOtion https://cordis.europa.eu/project/id/788851, https://project.inria.fr/ercnemo is an ERC Advanced Grant (2019 – 2024, PI François Baccelli). It is an inter-disciplinary proposal centered on network dynamics. The inter-disciplinarity spans from communication engineering to mathematics, with an innovative interplay between the two. NEMO's aim is to introduce dynamics in stochastic geometry. General mathematical tools combining stochastic geometry, random graph theory, and the theory of dynamical systems will be developed. NEMO will leverage interactions of Inria with Ecole Normale Supérieure on the mathematical side, and with Nokia Bell Labs and Orange on the engineering side. In March 2019, an inaugural workshop *Processus ponctuels et graphes aléatoires unimodulaires*https://project.inria.fr/ercnemo/fr/presentation was organized at Inria Paris.

9.3.2. Collaborations with Major European Organizations

Partner: VITO (Belgium); https://vito.be/en.

Co-advising of PhD student I. Shilov. Started: Nov 2019. Topic: "Algorithmic Games and Distributed Learning for Peer-to-Peer Energy Trading". PhD scholarship by VITO.

9.4. International Initiatives

9.4.1. Inria International Partners

9.4.1.1. Informal International Partners

- University of Florida; Collaborations with Prof Sean Meyn (ECE), Associate Prof Prabir Barooah (MAE), and the PhD students: A. Devraj (ECE), A. Coffman (MAE), N. Cammardella (ECE), J. Mathias (ECE).
- Sharif University, Tehran; Collaborations with O. Mirsadeghi.
- UC Berkeley; Collaborations with V. Anantharam.
- Indian Statistical Institute (ISI), Bangalore; Collaborations with Yogeshwaran D.

9.4.2. Participation in Other International Programs

9.4.2.1. Indo-French Center of Applied Mathematics

IFCAM Project “Geometric statistics of stationary point processes” B. Błaszczyszyn and Yogeshwaran D. from Indian Statistical Institute (ISI), Bangalore, have got in 2018 the approval from Indo-French Centre for Applied Mathematics (IFCAM), for their joint project on “Geometric statistics of stationary point processes" for the period 2018–2021. Yogeshwaran D. was visiting Dyogene for two weeks in March and November 2019.

9.4.2.2. Microsoft Research-Inria collaboration

Microsoft Research-Inria collaboration: Laurent Massoulié heads the Microsoft Research-Inria Joint Centre, and also participates to the “Distributed Machine Learning” project of the Joint Centre, together with Francis Bach (Inria), Sébastien Bubeck and Lin Xiao (MSR Redmond), and PhD student Hadrien Hendrikx.

9.4.2.3. Inria International Chairs

IIC- MEYN Sean

Title: Distributed Control and Smart Grid

International Partner (Institution - Laboratory - Researcher):

University of Florida (United States) - Department of Electrical and Computer Engineering - Sean Meyn

Duration: 2019 – 2023

Start year: 2019

See also: https://www.inria.fr/sites/default/files/2019-12/HoldersChairesInt_EN.pdf

TOPIC: “Distributed Control and Smart Grid”

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Ali Khezeli [School of Mathematical Sciences, Tehran, Iran],
- Christian Hirsch [Bernoulli Institute, University of Groningen,
- David Métivier [Los Alamos National Laboratory, USA]
- Deepjyoti Deka [Los Alamos National Laboratory, USA]
- Guenter Last [Karlsruhe Institute of Technology, Germany],
- Hermann Thorisson [University of Islande],
- Holger Keeler [University of Melbourne, Australia] ,
- Hrvoje Pandžić [University of Zagreb, Croatia]
- Itai Benjamini [Weizmann Institute of Science, Rehovot, Israel],
- Joe Yukich [Lehigh University, Bethelem, PA, USA],
- Josu Doncel [University of the Basque Country, Spain],
- Lucas Pereira [Técnico Lisboa, Portugal]
- Miklós Abért [MTA Renyi Institute, Budapest, Hungary],
- Mir-Omid Haji-Mirsadeghi [Sharif University, Tehran, Iran],
- Natasa Dragovic [The University of Texas at Austin, TX, USA],
- Nelson Antunes [University of Faro, Portugal],
- Venkatachalam Anantharam [University of California, Berkeley, CA USA],
- Yogeshwaran D. [ISI, Bangalore, India],

9.5.1.1. Internships

- Bastien Dubail [École Normale Supérieure de Lyon],
- Emmanuel Kravitzch [Inria],
- Erwan Pichon [Inria].
- Ge Jin [Inria],
- Maxence Lefort [Inria].

9.5.2. Visits to International Teams

- C. Fricker: University of Faro, Portugal (one week).

9.5.2.1. Research Stays Abroad

- A.Busic: program participant (5 weeks in total) of “The mathematics of energy systems”, Isaac Newton Institute for Mathematical Sciences, Cambridge, UK. Spring 2019, https://www.newton.ac.uk/event/mes

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

In collaboration with M. Vojnovic (London School of Economics), A. Busic and L. Massoulié organized an international workshop on Machine Learning and User Decision Making https://ml-udm.github.io/ at the ENS on May 23-24, 2019.

10.1.1.1. Member of the Organizing Committees

Members of Dyogene (co-)organized the following events:

- ERC inaugural workshop *Processus ponctuels et graphes aléatoires unimodulaires*https://project.inria.fr/ercnemo/fr/presentation/,
- Scientific session *Modélisation et analyse des systèmes de vélo-partage* at RFTM2019 (2emes Rencontres Francophones Transport et Mobilité, 11-13/06/2019 Montréal); https://symposia.cirrelt.ca/RFTM2019/
- A. Busic: co-lead (with E. Hyon, LIP 6) of the research group COSMOS (Stochastic optimization and control, modeling and simulation) of the GDR-RO; http://gdrro.lip6.fr/?q=node/78. Includes scientific organization of a one-day workshops per year (approx. 50 participants); In 2019: 8th GDT COSMOS workshop: "Stochastic Optimization and Reinforcement Learning"; November 2019; http://gdrro.lip6.fr/?q=node/211
- A. Busic: one week workshop “Flexible operation and advanced control for energy systems” within INI Cambridge program “The mathematics of energy systems”, January 2019, Cambridge, UK, https://www.newton.ac.uk/event/mesw01

10.1.2. Journal

10.1.2.1. Reviewer - Reviewing Activities

All members of the team act as reviewers for numerous scientific journals.

10.1.3. Invited Talks

- Mathematics Colloquium at *Paris-Descartes*, November 2019, F. Baccelli (talk on high dimensional stochastic geometry);
- Mathematics Colloquium at *Karlsruhe Institute of Technology*, July 2019, F. Baccelli (talk on unimodular dimensions);
- Invited lecture at the workshop *Modern Applied Probability, in celebration of Sergey Foss' 65th birthday*, ICMS, Edinburgh, May 2019 F. Baccelli (talk on particle systems); https://www.icms.org.uk/sergey65.php
- Invited lecture at the workshop *Point processes in space, time, and beyond*, Aalborg University, May 2019 F. Baccelli (talk on random graphs); http://people.math.aau.dk/~rw/PointSpaceBeyond/
- Mathematics Colloquium at *Institut Elie Cartan de Lorraine*, April 2019 F. Baccelli (on unimodular random metric spaces);
- Colloquium at **CEA LETI**, Grenoble, January 2019, F. Baccelli (talk on wireless stochastic geometry);
- Training School on Machine Learning for Communications, ISEP, Paris, B. Błaszczyszyn; https://sites.google.com/view/mlc-training-school/
- "STAR workshop on random graphs", Groningen Netherlands, B. Błaszczyszyn; http://www.math.rug.nl/~hansen/STAR2019/
- invited talk in the Department of Statistics and Data Science Yale University, September 2019, M. Lelarge;
- Isaac Newton Institute for Mathematical Sciences, Cambridge, UK, April 2019, A. Busic (on grid balancing through distributed control of flexible loads) ; video: https://www.newton.ac.uk/seminar/20190502143015301
- Simons Institute, UC Berkeley, June 2019, A. Busic (on optimizing mean-field dynamics for distributed control of flexible loads)
- Journées Scientifiques Inria, Lyon, June 2019, A. Busic (Optimisation/contrôle dans des réseaux électriques)
- CWI – Inria workshop, September 2019, Ana Busic (Optimizing mean-field dynamics for flexible loads control in power systems)
- CMAP seminar, École Politechnique, December 2019, A. Busic (on optimal control in dynamic matching systems)

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Licence: B. Błaszczyszyn (Cours) Théorie de l'information et du codage 24 heqTD, L3, ENS Paris.
- Licence: A. Busic (Cours) and S. Samain (TD) Structures et algorithmes aléatoires 60heqTD, L3, ENS Paris.
- Licence: L. Massoulié (Cours) Social and Communication networks 60heqTD, L3, l'X.
- Master: B. Błaszczyszyn (Cours) Processus ponctuels, graphes aléatoires et géeométrie stochastique 39heqTD, M2 Probabilités et Modèles Aléatoires, UPMC.
- Master: A. Busic (Cours) and L. Stephan (TD) Modèles et algorithmes de réseaux 60heqTD, M1, ENS Paris.
- Master: A. Busic (Cours) Fondements de la modélisation des réseaux 18 heqTD, M2 MPRI.
- Master: M. Lelarge (Cours) Deep Learning Do it Yourself, M1, ENS Paris. X, X-HEC https://mlelarge.github.io/dataflowr-web/

- Master: M. Lelarge (Cours) Deep Learning Do it Yourself, M1, ENS Paris.
- Master: L. Massoulié (Cours) Inference in large random graphs, M2 Université d'Orsay.
- Summer school: M. Lelarge (Cours) HANDS-ON TOUR TO DEEP LEARNING WITH PYTORCH, https://mlelarge.github.io/dataflowr-web/cea_edf_inria.html.
- Invited mini-course: L. Massoulié at Journée spéciale Stat Math 2019 of Société Française de Statistiques, Institut Henri Poincaré.

10.2.2. Supervision

- PhD: Léo Miolane, "High dimensional statistics", defended in 2019, advised by M. Lelarge. [2]
- PhD: Md Umar Hashmi, "Decentralized control for renewable integration in smartgrids", defended in 2019, advised by A. Busic. [1]
- PhD in progress: Alexis Galland, Deep Learning on Graphs, since 2017, advised by M. Lelarge.
- PhD in progress: Quentin Le Gall "Crowd networking : modélisation de la connectivité D2D" since October 2017; PhD CIFRE co-advised by B. Błaszczyszyn and E. Cali (Orange).
- PhD in progress: Antoine Brochard "Signal processing for point processes and statistical learning for telecommunications", since September 2018; PhD CIFRE co-advised by B. Błaszczyszyn and Georgios Paschos (Huawei).
- PhD in progress: Sébastien Samain, "Monte Carlo methods for performance evaluation and reinforcement learning", since November 2016, advised by A. Busic,
- PhD in progress: Arnaud Cadas, "Dynamic matching models", since October 2017, supervised by A. Busic.
- PhD in progress: Michel Davydov, since September 2019, F. Baccelli.
- PhD in progress: Luca Ganassali, since September 2019.
- PhD in progress: Hadrien Hendrikx, since 2019
- PhD in progress: Sayeh Khaniha, form 2019, supervised by F. Baccelli.
- PhD in progress: Edouard Pineau, since 2019,
- PhD in progress: Bharath Roy, since 2019, supervised by F. Baccelli and B. Błaszczyszyn,
- PhD in progress: Ilia Shilov, since 2019, supervised by A. Busic,
- PhD in progress: Ludovic Stephan, since 2019.

10.2.3. Juries

- F. Baccelli: PhD reviewer of **Gourab GHATAK**, Télécom Paris; HDR reviewer of **Marios KOUNTOURIS**, Université Paris-Sud 1; HDR jury member of **Raphaël LACHIEZE-REY**, Université Paris-Descartes.
- B. Blaszczyszyn: PhD reviewer of: **Sanjoy Kumar JHAWAR**, Indian Institute of Science Bangalore, India; **Arnaud POINAS**, Université de Rennes 1; HDR reviewer of: **Marious KOUNTOURIS**, Université Paris-Sud 1; PhD jury member of : **Jalal RACHAD**, Télécom Paris.
- C. Fricker: PhD reviewer of: **Celine COMTE**, Université Paris-Saclay; PhD jury member of: **Hamza BEN AMMAR**, Université de Rennes 1.
- M. Lelarge: PhD jury member Xiaoyi MAI, Université Paris-Saclay.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] M. U. HASHMI.*Optimization and Control of Storage in Smart Grids*, PSL Research University, December 2019, https://tel.archives-ouvertes.fr/tel-02462786

[2] L. MIOLANE.*Fundamental limits of inference: A statistical physics approach*, Ecole normale supérieure - ENS PARIS ; Inria Paris, June 2019, https://hal.archives-ouvertes.fr/tel-02446988

Articles in International Peer-Reviewed Journal

[3] V. ANANTHARAM, F. BACCELLI.*Error Exponents for Dimension-Matched Vector Multiple Access Channels With Additive Noise*, in "IEEE Transactions on Information Theory", May 2019, vol. 65, n^o 5, p. 2808-2823 [*DOI :* 10.1109/TIT.2018.2875541], https://hal.inria.fr/hal-02422196

[4] F. BACCELLI, M.-O. HAJI-MIRSADEGHI, J. T. MURPHY.*Doeblin trees*, in "Electronic Journal of Probability", 2019, vol. 24, https://arxiv.org/abs/1811.10058 [*DOI :* 10.1214/19-EJP375], https://hal.inria.fr/hal-02415283

[5] F. BACCELLI, E. O'REILLY.*The stochastic geometry of unconstrained one-bit data compression*, in "Electronic Journal of Probability", 2019, vol. 24, https://arxiv.org/abs/1810.06095 [*DOI :* 10.1214/19-EJP389], https://hal.inria.fr/hal-02422200

[6] B. BŁASZCZYSZYN, M. K. KARRAY.*Performance analysis of cellular networks with opportunistic scheduling using queueing theory and stochastic geometry*, in "IEEE Transactions on Wireless Communications", September 2019, vol. 18, n^o 12, p. 5952-5966 [*DOI :* 10.1109/TWC.2019.2940938], https://hal.inria.fr/hal-01824986

[7] B. BŁASZCZYSZYN, D. YOGESHWARAN, J. E. YUKICH.*Limit theory for geometric statistics of point processes having fast decay of correlations*, in "Annals of Probability", 2019, vol. 47, n^o 2, p. 835-895, https://arxiv.org/abs/1606.03988 [*DOI :* 10.1214/18-AOP1273], https://hal.inria.fr/hal-01331939

[8] Y. CHABCHOUB, R. E. SIBAI, C. FRICKER.*Bike sharing systems: a new incentive rebalancing method based on spatial outliers detection*, in "International Journal of Space-Based and Situated Computing", 2019, https://hal.archives-ouvertes.fr/hal-02430574

[9] A. CHATTOPADHYAY, B. BŁASZCZYSZYN, E. ALTMAN.*Location Aware Opportunistic Bandwidth Sharing between Static and Mobile Users with Stochastic Learning in Cellular Networks*, in "IEEE Transactions on Mobile Computing", 2019, vol. 18, n^o 8, p. 1802-1815, https://arxiv.org/abs/1608.04260 - 16 Pages, 1 Figure, 1 Table [*DOI :* 10.1109/TMC.2018.2867037], https://hal.inria.fr/hal-01401007

[10] A. CHATTOPADHYAY, B. BŁASZCZYSZYN, E. ALTMAN.*Two-tier cellular networks for throughput maximization of static and mobile users*, in "IEEE Transactions on Wireless Communications", 2019, vol. 18, n^o 2, p. 997-1010, https://arxiv.org/abs/1605.07341 - title of the previous version: "Cell planning for mobility management in heterogeneous cellular networks", forthcoming [*DOI :* 10.1109/TWC.2018.2887386], https://hal.inria.fr/hal-01331936

[11] M. DEHGHAN, L. MASSOULIÉ, D. TOWSLEY, D. S. MENASCHE, Y. C. TAY.*A Utility Optimization Approach to Network Cache Design*, in "IEEE/ACM Transactions on Networking", June 2019, vol. 27, n^o 3, p. 1013-1027 [*DOI :* 10.1109/TNET.2019.2913677], https://hal.archives-ouvertes.fr/hal-02427851

[12] J. DONCEL, N. GAST, B. GAUJAL.*Discrete Mean Field Games: Existence of Equilibria and Convergence*, in "Journal of Dynamics and Games", 2019, vol. 6, n^o 3, p. 1-19, https://arxiv.org/abs/1909.01209 [*DOI :* 10.3934/JDG.2019016], https://hal.inria.fr/hal-01277098

[13] S. S. KALAMKAR, M. HAENGGI.*Per-Link Reliability and Rate Control: Two Facets of the SIR Meta Distribution*, in "IEEE wireless communications letters", August 2019, vol. 8, n^{o} 4, p. 1244-1247 [*DOI :* 10.1109/LWC.2019.2912614], https://hal.archives-ouvertes.fr/hal-02417332

[14] S. S. KALAMKAR, M. HAENGGI.*Simple Approximations of the SIR Meta Distribution in General Cellular Networks*, in "IEEE Transactions on Communications", August 2019, vol. 67, n^{o} 6, p. 4393-4406 [*DOI :* 10.1109/TCOMM.2019.2900676], https://hal.archives-ouvertes.fr/hal-02417353

[15] L. MASSOULIÉ, R. VARLOOT.*Rapid Mixing of Dynamic Graphs with Local Evolution Rules*, in "IEEE Transactions on Network Science and Engineering", July 2019, 1 [*DOI :* 10.1109/TNSE.2019.2926887], https://hal.archives-ouvertes.fr/hal-02427853

[16] A. SANKARARAMAN, F. BACCELLI, S. FOSS.*Interference queueing networks on grids*, in "Annals of Applied Probability", October 2019, vol. 29, n^{o} 5, p. 2929-2987, https://arxiv.org/abs/1710.09797 [*DOI :* 10.1214/19-AAP1470], https://hal.inria.fr/hal-02415268

[17] K. SCAMAN, F. BACH, S. BUBECK, Y. T. LEE, L. MASSOULIÉ.*Optimal Convergence Rates for Convex Distributed Optimization in Networks*, in "Journal of Machine Learning Research", 2019, vol. 20, p. 1 - 31, https://hal.inria.fr/hal-02440504

[18] C. STEGEHUIS, L. MASSOULIÉ.*Efficient inference in stochastic block models with vertex labels*, in "IEEE Transactions on Network Science and Engineering", April 2019, 1 [*DOI :* 10.1109/TNSE.2019.2913949], https://hal.archives-ouvertes.fr/hal-02427854

International Conferences with Proceedings

[19] A. BROCHARD, B. BŁASZCZYSZYN, S. MALLAT, S. ZHANG.*Statistical learning of geometric characteristics of wireless networks*, in "INFOCOM 2019 - IEEE International Conference on Computer Communications", Paris, France, Proc. of IEEE Infocom, IEEE, June 2019, https://arxiv.org/abs/1812.08265 [*DOI :* 10.1109/INFOCOM.2019.8737441], https://hal.inria.fr/hal-01962191

[20] A. BUŠIC, S. MEYN.*Distributed Control of Thermostatically Controlled Loads: Kullback-Leibler Optimal Control in Continuous Time*, in "58th IEEEConference on Decision and Control (CDC 2019)", Nice, France, Proceedings of the 58th IEEEConference on Decision and Control (CDC), 2019, https://hal.archives-ouvertes.fr/hal-02425979

[21] B. BŁASZCZYSZYN, H. P. KEELER.*Determinantal thinning of point processes with network learning applications*, in "WCNC 2019 - IEEE Wireless Communications and Networking Conference", Marrakech, Morocco, Proc WCNC, IEEE, 2019, https://arxiv.org/abs/1810.08672 - 8 pages, 6 figures. All code available online: https://github.com/hpaulkeeler/DetPoisson_MATLAB, https://hal.inria.fr/hal-01958924

[22] B. BŁASZCZYSZYN, P. MUHLETHALER.*Analyzing LoRa long-range, low-power, wide-area networks using stochastic geometry*, in "VALUETOOLS 2019 - 12th EAI International Conference on Performance Evaluation Methodologies and Tools", Palma de Mallorca, Spain, Proc. ACM VALUETOOLS, ACM, 2019, https://arxiv.org/abs/1811.01886 - 8 pages [*DOI :* 10.1145/3306309.3306327], https://hal.inria.fr/hal-01958939

[23] A. CADAS, A. BUŠIC, J. DONCEL.*Optimal Control of Dynamic Bipartite Matching Models*, in "VALUETOOLS 2019 - 12th EAI International Conference on Performance Evaluation Methodologies and Tools", Palma de Mallorca, Spain, VALUETOOLS 2019: Proceedings of the 12th EAI International Conference

on Performance Evaluation Methodologies and Tools, 2019, p. 39-46, https://arxiv.org/abs/1810.08541 [*DOI :* 10.1145/3306309.3306317], https://hal.archives-ouvertes.fr/hal-01968549

[24] N. CAMMARDELLA, A. BUŠIC, Y. JI, S. MEYN.*Kullback-Leibler-Quadratic Optimal Control of Flexible Power Demand*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, Proceedings of the 58th IEEE Conference on Decision and Control (CDC), 2019, https://hal.archives-ouvertes.fr/hal-02425974

[25] Y. CHEN, M. U. HASHMI, D. DEKA, M. CHERTKOV.*Stochastic Battery Operations using Deep Neural Networks*, in "IEEE ISGT NA 2019 - Innovative Smart Grid Technologies Conference", Washington, United States, February 2019, https://hal.archives-ouvertes.fr/hal-02070232

[26] A. R. COFFMAN, A. BUŠIC, P. BAROOAH.*Aggregate capacity for TCLs providing virtual energy storage with cycling constraints*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, Proceedings of the 58th IEEE Conference on Decision and Control (CDC), 2019, https://hal.archives-ouvertes.fr/hal-02425977

[27] A. M. DEVRAJ, A. BUŠIC, S. MEYN.*On Matrix Momentum Stochastic Approximation and Applications to Q-learning*, in "57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)", Monticello, IL, United States, Proceedings of 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2019, https://arxiv.org/abs/1809.06277 [*DOI :* 10.1109/ALLERTON.2019.8919828], https://hal.archives-ouvertes.fr/hal-01968558

[28] A. M. DEVRAJ, A. BUŠIC, S. MEYN.*Zap Q-Learning - A User's Guide*, in "ICC 2019 - Fifth Indian Control Conference", New Delhi, India, Proceedings of the Fifth Indian Control Conference (ICC 2019), IEEE, 2019, p. 10-15 [*DOI :* 10.1109/INDIANCC.2019.8715554], https://hal.archives-ouvertes.fr/hal-02429733

[29] M. U. HASHMI, A. MUKHOPADHYAY, A. BUŠIC, J. ELIAS, D. KIEDANSKI.*Optimal Storage Arbitrage under Net Metering using Linear Programming*, in "(SmartGridComm 2019 - IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids", Beijing, China, Proceedings of the 2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), IEEE, October 2019, p. 1-7 [*DOI :* 10.1109/SMARTGRIDCOMM.2019.8909753], https://hal.archives-ouvertes.fr/hal-02425626

[30] M. U. HASHMI, L. PEREIRA, A. BUŠIC.*Energy Storage in Madeira, Portugal: Co-optimizing for Arbitrage, Self-Sufficiency, Peak Shaving and Energy Backup*, in "13th IEEE PES PowerTech Conference", Milan, Italy, June 2019, https://hal.archives-ouvertes.fr/hal-02085839

[31] H. HENDRIKX, F. BACH, L. MASSOULIÉ.*Accelerated decentralized optimization with local updates for smooth and strongly convex objectives*, in "AISTATS 2019 - 22nd International Conference on Artificial Intelligence and Statistics", Naha, Okinawa, Japan, April 2019, https://hal.inria.fr/hal-01893568

[32] A. HOLLOCOU, T. BONALD, M. LELARGE.*Modularity-based Sparse Soft Graph Clustering*, in "AISTATS 2019 - 22nd International Conference on Artificial Intelligence and Statistics", Naha, Okinawa, Japan, April 2019, https://hal.archives-ouvertes.fr/hal-02005331

[33] S. S. KALAMKAR.*Reliability and Local Delay in Wireless Networks: Does Bandwidth Partitioning Help?*, in "GLOBECOM 2019 - IEEE Global Communications Conference", Waikoloa, Hawaii, United States, January 2019, https://hal.archives-ouvertes.fr/hal-02416870

[34] D. KIEDANSKI, M. U. HASHMI, A. BUŠIC, D. KOFMAN.*Sensitivity to forecast errors in energy storage arbitrage for residential consumers*, in "SmartGridComm 2019 - IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids", Beijing, China, October 2019, https://hal.telecom-paristech.fr/hal-02163114

[35] Q. LE GALL, B. BŁASZCZYSZYN, E. CALI, T. EN-NAJJARY.*The Influence of Canyon Shadowing on Device-to-Device Connectivity in Urban Scenario*, in "WCNC 2019 - IEEE Wireless Communications and Networking Conference", Marrakech, Morocco, Proc. of IEEE WCNC, IEEE, 2019, https://arxiv.org/abs/1810.03498 - 8 pages, 7 figures, 2 tables, https://hal.inria.fr/hal-01891024

[36] L. MASSOULIÉ, L. STEPHAN, D. TOWSLEY.*Planting trees in graphs, and finding them back*, in "COLT 2019 - Conference on Learning Theory", Phoenix, United States, June 2019, https://arxiv.org/abs/1811.01800 , https://hal.archives-ouvertes.fr/hal-02427845

[37] L. STEPHAN, L. MASSOULIÉ.*Robustness of spectral methods for community detection*, in "COLT 2019 - Conference on Learning Theory", Phoenix, United States, June 2019, https://arxiv.org/abs/1811.05808 , https://hal.archives-ouvertes.fr/hal-02427834

Scientific Books (or Scientific Book chapters)

[38] F. BACCELLI, B. BŁASZCZYSZYN, M. K. KARRAY.*Random Measures, Point Processes, and Stochastic Geometry*, January 2020, https://hal.inria.fr/hal-02460214

Other Publications

[39] F. BACCELLI, M.-O. HAJI-MIRSADEGHI, A. KHEZELI.*On the Dimension of Unimodular Discrete Spaces, Part I: Definitions and Basic Properties*, January 2019, https://arxiv.org/abs/1807.02980 - working paper or preprint, https://hal.inria.fr/hal-01976265

[40] F. BACCELLI, M.-O. HAJI-MIRSADEGHI, A. KHEZELI.*On the Dimension of Unimodular Discrete Spaces, Part II: Relations with Growth Rate*, January 2019, https://arxiv.org/abs/1808.02551 - working paper or preprint, https://hal.inria.fr/hal-01976257

[41] J. BARBIER, F. KRZAKALA, N. MACRIS, L. MIOLANE, L. ZDEBOROVÁ.*Phase Transitions, Optimal Errors and Optimality of Message-Passing in Generalized Linear Models*, November 2019, https://arxiv.org/abs/1708.03395 - 35 pages, 3 figures, https://hal-cea.archives-ouvertes.fr/cea-01614258

[42] B. BŁASZCZYSZYN, C. HIRSCH.*Optimal stationary markings*, January 2020, https://arxiv.org/abs/2001.08074 - 26 pages, https://hal.inria.fr/hal-02457091

[43] S. CHEN, A. M. DEVRAJ, A. BUŠIC, S. MEYN.*Zap Q-Learning for Optimal Stopping Time Problems*, 2019, https://arxiv.org/abs/1904.11538 - Accepted at the 2020 American Control Conference, https://hal.archives-ouvertes.fr/hal-02425983

[44] S. CHEN, A. M. DEVRAJ, A. BUŠIC, S. MEYN.*Zap Q-Learning With Nonlinear Function Approximation*, 2019, https://arxiv.org/abs/1910.05405 - working paper or preprint [*DOI :* 10.05405], https://hal.archives-ouvertes.fr/hal-02425985

[45] Q. L. GALL, B. BŁASZCZYSZYN, E. CALI, T. EN-NAJJARY.*Continuum Line-of-Sight Percolation on Poisson-Voronoi Tessellations*, July 2019, https://arxiv.org/abs/1904.10875 - 30 pages, 4 figures, https://hal.inria.fr/hal-02192469

[46] Q. L. GALL, B. BŁASZCZYSZYN, E. CALI, T. EN-NAJJARY.*Relay-assisted Device-to-Device Networks: Connectivity and Uberization Opportunities*, October 2019, https://arxiv.org/abs/1909.12867 - 7 pages, 5 figures, 3 tables, https://hal.archives-ouvertes.fr/hal-02320085

[47] M. U. HASHMI, J. CAVALEIRO, L. PEREIRA, A. BUŠIC.*Sizing and Profitability of Energy Storage for Prosumers in Madeira, Portugal*, December 2019, https://arxiv.org/abs/1911.10458 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02425980

[48] M. U. HASHMI, D. DEKA, A. BUŠIC, L. PEREIRA, S. BACKHAUS.*Arbitrage with Power Factor Correction using Energy Storage*, 2019, https://arxiv.org/abs/1903.06132 - 11 pages, 14 figures, https://hal.archives-ouvertes.fr/hal-02425982

[49] H. HENDRIKX, F. BACH, L. MASSOULIÉ.*An Accelerated Decentralized Stochastic Proximal Algorithm for Finite Sums*, September 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02280763

Project-Team EVA

Wireless Networking for Evolving & Adaptive Applications

RESEARCH CENTER
Paris

THEME
Networks and Telecommunications

Table of contents

Project-Team EVA

Creation of the Team: 2015 April 01, updated into Project-Team: 2016 May 01

Keywords:

Computer Science and Digital Science:

A1.2. - Networks
A1.2.1. - Dynamic reconfiguration
A1.2.2. - Supervision
A1.2.3. - Routing
A1.2.4. - QoS, performance evaluation
A1.2.5. - Internet of things
A1.2.6. - Sensor networks
A1.2.7. - Cyber-physical systems
A1.2.8. - Network security
A1.2.9. - Social Networks
A1.4. - Ubiquitous Systems
A1.6. - Green Computing
A2.3. - Embedded and cyber-physical systems
A2.3.1. - Embedded systems
A2.3.2. - Cyber-physical systems
A2.3.3. - Real-time systems
A3.4. - Machine learning and statistics
A3.4.1. - Supervised learning
A3.4.6. - Neural networks
A3.4.7. - Kernel methods
A4. - Security and privacy
A4.1. - Threat analysis
A4.1.1. - Malware analysis
A4.1.2. - Hardware attacks
A4.4. - Security of equipment and software
A4.5. - Formal methods for security
A4.6. - Authentication
A4.7. - Access control
A5.10. - Robotics
A5.10.6. - Swarm robotics
A5.10.8. - Cognitive robotics and systems
A6. - Modeling, simulation and control
A9.2. - Machine learning
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B5.1. - Factory of the future
B6. - IT and telecom

B6.2. - Network technologies
B6.2.1. - Wired technologies
B6.2.2. - Radio technology
B6.3.2. - Network protocols
B6.3.3. - Network Management
B6.3.4. - Social Networks
B6.4. - Internet of things
B6.6. - Embedded systems
B7. - Transport and logistics
B7.1.1. - Pedestrian traffic and crowds
B7.1.2. - Road traffic
B7.2. - Smart travel
B7.2.1. - Smart vehicles
B7.2.2. - Smart road
B8. - Smart Cities and Territories
B8.1. - Smart building/home
B8.1.1. - Energy for smart buildings
B8.1.2. - Sensor networks for smart buildings
B8.2. - Connected city
B8.4. - Security and personal assistance
B8.4.1. - Crisis management

1. Team, Visitors, External Collaborators

Research Scientists

Pascale Minet [Inria, Researcher, HDR]
Paul Muhlethaler [Team leader, Inria, Senior Researcher, HDR]
Malisa Vucinic [Inria,Starting Research Position]
Thomas Watteyne [Inria,Senior Researcher, HDR]

External Collaborators

Kris Pister [UC Berkeley, REALMS associate team]
Steven Glaser [UC Berkeley, REALMS associate team]
Xavi Vilajosana [UOC, Catalunya, MOU OpenMote]
Nadjib Achir [Habilite]
Selma Boumerdassi [CNAM]
Samia Bouzefrane [CNAM]
Philippe Jacquet [Bell Labs (Nokia) Nov 2019, HDR]
Anis Laouiti [Habilite]
Mohamed Elhadad [Vedecom]
Oyunchimeg Shagdar [Vedecom]

Technical Staff

Jonathan Munoz Soto [From April 2019 GeoBot, Engineer]
Trifun Savic [GeoBot, Engineer, from Feb 2019]
Keoma Brun-Laguna [Inria, ATT SmartMarina, Engineer]
Tengfei Chang [Inria, ADT 6TiSCH, Engineer]
Yasuyuki Tanaka [Inria, ADT DASMU, Engineer]

PhD Students

Mina Rady [CIFRE with Orange Labs, started January 2019]
Razanne Abu Aisheh [CIFRE with Nokia Bell Labs, started December 2019]
Amar Abane [CNAM, from Mar 2019 until Jul 2019]
Iman Hmedoush [Inria]
Jonathan Munoz Soto [Inria, until March 2019]
Mamoudou Sangare [Univerity Conakry, from October 2017 until december 2020]
Abdallah Sobehy [Univ Paris-Saclay]
Fouzi Boukhalfa [Vedecom from October 2018 until September 2021]

Post-Doctoral Fellow

Timothy Claeys [Inria, from Dec 2019]

Visiting Scientists

Elsa Nicol [CEO & Co-Founder, Wattson Elements, Falco]
Olivier Laporte [Customer Support Europe & Co-Founder, Wattson Elements, Falco]
Martina Brachmann [RISE, Sweden, Postdoc, Nov 2019]
Ana Laura Diedrichs [UTN, Argentina, PhD student, Oct-Nov 2019]
Leila Azouz [ENSI Tunis, Professor, Oct 2019]
Randy Hare [GeoBot project, Jun-Aug 2019]
Diego Dujovne [Universidad Diego Portales, Chile, July 2019]
Branko Kerkez [U. Michigan, USA, Professor, May 2019]
Mikolaj Chwalisz [TU Berlin, PhD student, May 2019]

2. Overall Objectives

2.1. Overall Objectives

It is forecast that the vast majority of Internet connections will be wireless. The EVA project grasps this opportunity and focuses on wireless communication. EVA tackles challenges related to providing efficient communication in wireless networks and, more generally, in all networks that are not already organized when set up, and consequently need to evolve and spontaneously find a match between application requirements and the environment. These networks can use opportunistic and/or collaborative communication schemes. They can evolve through optimization and self-learning techniques. Every effort is made to ensure that the results provided by EVA have the greatest possible impact through standardization. The miniaturization and ubiquitous nature of computing devices has opened the way to the deployment of a new generation of wireless (sensor) networks. These networks are central to the work in EVA, as EVA focuses on such crucial issues as power conservation, connectivity, determinism, reliability and latency. Wireless Sensor Network (WSN) deployments are also to be a new key subject, especially for emergency situations (e.g. after a disaster). Industrial process automation and environmental monitoring are considered in greater depth.

3. Research Program

3.1. Pitch

Designing Tomorrow's Internet of (Important) Things

Inria-EVA is a leading research team in low-power wireless communications. The team pushes the limits of low-power wireless mesh networking by applying them to critical applications such as industrial control loops, with harsh reliability, scalability, security and energy constraints. Grounded in real-world use cases and experimentation, EVA co-chairs the IETF 6TiSCH and LAKE standardization working groups, co-leads Berkeley's OpenWSN project and works extensively with Analog Devices' SmartMesh IP networks. Inria-EVA is the birthplace of the Wattson Elements startup and the Falco solution. The team is associated with Prof. Glaser's (UC Berkeley) and Prof. Kerkez (U. Michigan) through the REALMS associate research team, and with OpenMote through a long-standing Memorandum of Understanding.

3.2. Physical Layer

We study how advanced physical layers can be used in low-power wireless networks. For instance, collaborative techniques such as multiple antennas (e.g. Massive MIMO technology) can improve communication efficiency. The core idea is to use massive network densification by drastically increasing the number of sensors in a given area in a Time Division Duplex (TDD) mode with time reversal. The first period allows the sensors to estimate the channel state and, after time reversal, the second period is to transmit the data sensed. Other techniques, such as interference cancellation, are also possible.

3.3. Wireless Access

Medium sharing in wireless systems has received substantial attention throughout the last decade. HiPERCOM2 has provided models to compare TDMA and CSMA. HiPERCOM2 has also studied how network nodes must be positioned to optimize the global throughput.

EVA pursues modeling tasks to compare access protocols, including multi-carrier access, adaptive CSMA (particularly in VANETs), as well as directional and multiple antennas. There is a strong need for determinism in industrial networks. The EVA team focuses particularly on scheduled medium access in the context of deterministic industrial networks; this involves optimizing the joint time slot and channel assignment. Distributed approaches are considered, and the EVA team determines their limits in terms of reliability, latency and throughput. Furthermore, adaptivity to application or environment changes are taken into account.

3.4. Coexistence of Wireless Technologies

Wireless technologies such as cellular, low-power mesh networks, (Low-Power) WiFi, and Bluetooth (low-energy) can reasonably claim to fit the requirements of the IoT. Each, however, uses different trade-offs between reliability, energy consumption and throughput. The EVA team studies the limits of each technology, and will develop clear criteria to evaluate which technology is best suited to a particular set of constraints.

Coexistence between these different technologies (or different deployments of the same technology in a common radio space) is a valid point of concern.

The EVA team aims at studying such coexistence, and, where necessary, propose techniques to improve it. Where applicable, the techniques will be put forward for standardization. Multiple technologies can also function in a symbiotic way.

For example, to improve the quality of experience provided to end users, a wireless mesh network can transport sensor and actuator data in place of a cellular network, when and where cellular connectivity is poor.

The EVA team studies how and when different technologies can complement one another. A specific example of a collaborative approach is Cognitive Radio Sensor Networks (CRSN).

3.5. Energy-Efficiency and Determinism

Reducing the energy consumption of low-power wireless devices remains a challenging task. The overall energy budget of a system can be reduced by using less power-hungry chips, and significant research is being done in that direction. That being said, power consumption is mostly influenced by the algorithms and protocols used in low-power wireless devices, since they influence the duty-cycle of the radio.

EVA will search for energy-efficient mechanisms in low-power wireless networks. One new requirement concerns the ability to predict energy consumption with a high degree of accuracy. Scheduled communication, such as the one used in the IEEE 802.15.4 TSCH (Time Slotted CHannel Hopping) standard, and by IETF 6TiSCH, allows for a very accurate prediction of the energy consumption of a chip. Power conservation will be a key issue in EVA.

To tackle this issue and match link-layer resources to application needs, EVA's 5-year research program dealing with Energy-Efficiency and Determinism centers around 3 studies:

- Performance Bounds of a TSCH network. We propose to study a low-power wireless TSCH network as a Networked Control System (NCS), and use results from the NCS literature. A large number of publications on NCS, although dealing with wireless systems, consider wireless links to have perfect reliability, and do not consider packet loss. Results from these papers can not therefore be applied directly to TSCH networks. Instead of following a purely mathematical approach to model the network, we propose to use a non-conventional approach and build an empirical model of a TSCH network.
- Distributed Scheduling in TSCH networks. Distributed scheduling is attractive due to its scalability and reactivity, but might result in a sub-optimal schedule. We continue this research by designing a distributed solution based on control theory, and verify how this solution can satisfy service level agreements in a dynamic environment.

3.6. Network Deployment

Since sensor networks are very often built to monitor geographical areas, sensor deployment is a key issue. The deployment of the network must ensure full/partial, permanent/intermittent coverage and connectivity. This technical issue leads to geometrical problems which are unusual in the networking domain.

We can identify two scenarios. In the first one, sensors are deployed over a given area to guarantee full coverage and connectivity, while minimizing the number of sensor nodes. In the second one, a network is re-deployed to improve its performance, possibly by increasing the number of points of interest covered, and by ensuring connectivity. EVA will investigate these two scenarios, as well as centralized and distributed approaches. The work starts with simple 2D models and will be enriched to take into account more realistic environment: obstacles, walls, 3D, fading.

3.7. Data Gathering and Dissemination

A large number of WSN applications mostly do data gathering (a.k.a "convergecast"). These applications usually require small delays for the data to reach the gateway node, requiring time consistency across gathered data. This time consistency is usually achieved by a short gathering period.

In many real WSN deployments, the channel used by the WSN usually encounters perturbations such as jamming, external interferences or noise caused by external sources (e.g. a polluting source such as a radar) or other coexisting wireless networks (e.g. WiFi, Bluetooth). Commercial sensor nodes can communicate on multiple frequencies as specified in the IEEE 802.15.4 standard. This reality has given birth to the multichannel communication paradigm in WSNs.

Multichannel WSNs significantly expand the capability of single-channel WSNs by allowing parallel transmissions, and avoiding congestion on channels or performance degradation caused by interfering devices.

In EVA, we will focus on raw data convergecast in multichannel low-power wireless networks. In this context, we are interested in centralized/distributed algorithms that jointly optimize the channel and time slot assignment used in a data gathering frame. The limits in terms of reliability, latency and bandwidth will be evaluated. Adaptivity to additional traffic demands will be improved.

3.8. Self-Learning Networks

To adapt to varying conditions in the environment and application requirements, the EVA team investigate self-learning networks. Machine learning approaches, based on experts and forecasters, are investigated to predict the quality of the wireless links in a WSN. This allows the routing protocol to avoid using links exhibiting poor quality and to change the route before a link failure. Additional applications include where to place the aggregation function in data gathering. In a content delivery network (CDN), it is very useful to predict popularity, expressed by the number of requests per day, for a multimedia content. The most popular contents are cached near the end-users to maximize the hit ratio of end-users' requests. Thus the satisfaction degree of end-users is maximized and the network overhead is minimized.

3.9. Internet of Things Security

Existing Internet threats might steal our digital information. Tomorrow's threats could disrupt power plants, home security systems, hospitals. The Internet of Things is bridging our digital security with personal safety. Popular magazines are full of stories of hacked devices (e.g. drone attack on Philips Hue), IoT botnets (e.g. Mirai), and inherent insecurity.

Why has the IoT industry failed to adopt the available computer security techniques and best practices? Our experience from research, industry collaborations, and the standards bodies has shown that the main challenges are:

1. The circumvention of the available technical solutions due to their inefficiency.
2. The lack of a user interface for configuring the product in the field resulting in default parameters being (re)used.
3. Poorly tested software, often lacking secure software upgrade mechanisms.

Our research goal is to contribute to a more secure IoT, by proposing technical solutions to these challenges for low-end IoT devices with immediate industrial applicability and transfer potential. We complement the existing techniques with the missing pieces to move towards truly usable and secure IoT systems.

4. Application Domains

4.1. Industrial Process Automation

Wireless networks have become ubiquitous and are an integral part of our daily lives. These networks are present in many application domains; the most important are detailed in this section.

Networks in industrial process automation typically perform **monitoring and control** tasks. Wired industrial communication networks, such as HART [0], have been around for decades and, being wired, are highly reliable. Network administrators tempted to "go wireless" expect the same reliability. Reliable process automation networks – especially when used for control – often impose stringent latency requirements. Deterministic wireless networks can be used in critical systems such as control loops, however, the unreliable nature of the wireless medium, coupled with their large scale and "ad-hoc" nature raise some of the most important challenges for low-power wireless research over the next 5-10 years.

Through the involvement of team members in standardization activities, protocols and techniques are proposed for the standardization process with a view to becoming the *de-facto* standard for wireless industrial process automation. Besides producing top level research publications and standardization activities, EVA intends this activity to foster further collaborations with industrial partners.

[0] Highway Addressable Remote Transducer

4.2. Environmental Monitoring

Today, outdoor WSNs are used to monitor vast rural or semi-rural areas and may be used to detect fires. Another example is detecting fires in outdoor fuel depots, where the delivery of alarm messages to a monitoring station in an upper-bounded time is of prime importance. Other applications consist in monitoring the snow melting process in mountains, tracking the quality of water in cities, registering the height of water in pipes to foresee flooding, etc. These applications lead to a vast number of technical issues: deployment strategies to ensure suitable coverage and good network connectivity, energy efficiency, reliability and latency, etc.

We work on such applications in an associate team ”REALMS” comprising members from EVA, the university of Berkeley and the university of Michigan.

4.3. The Internet of Things

The general agreement is that the Internet of Things (IoT) is composed of small, often battery-powered objects which measure and interact with the physical world, and encompasses smart home applications, wearables, smart city and smart plant applications.

It is absolutely essential to (1) clearly understand the limits and capabilities of the IoT, and (2) develop technologies which enable user expectation to be met.

The EVA team is dedicated to understanding and contributing to the IoT. In particular, the team maintains a good understanding of the different technologies at play (Bluetooth, IEEE 802.15.4, WiFi, cellular), and their trade-offs. Through scientific publications and other contributions, EVA helps establish which technology best fits which application.

4.4. Military, Energy and Aerospace

Through the HIPERCOM project, EVA has developed cutting-edge expertise in using wireless networks for military, energy and aerospace applications. Wireless networks are a key enabling technology in the application domains, as they allow physical processes to be instrumented (e.g. the structural health of an airplane) at a granularity not achievable by its wired counterpart. Using wireless technology in these domains does however raise many technical challenges, including end-to-end latency, energy-efficiency, reliability and Quality of Service (QoS). Mobility is often an additional constraint in energy and military applications. Achieving scalability is of paramount importance for tactical military networks, and, albeit to a lesser degree, for power plants. EVA will work in this domain.

Smart cities share the constraint of mobility (both pedestrian and vehicular) with tactical military networks. Vehicular Ad-hoc NETworks (VANETs) will play an important role in the development of smarter cities.

The coexistence of different networks operating in the same radio spectrum can cause interference that should be avoided. Cognitive radio provides secondary users with the frequency channels that are temporarily unused (or unassigned) by primary users. Such opportunistic behavior can also be applied to urban wireless sensor networks. Smart cities raise the problem of transmitting, gathering, processing and storing big data. Another issue is to provide the right information at the place where it is most needed.

4.5. Emergency Applications

In an “emergency” application, heterogeneous nodes of a wireless network cooperate to recover from a disruptive event in a timely fashion, thereby possibly saving human lives. These wireless networks can be rapidly deployed and are useful to assess damage and take initial decisions. Their primary goal is to maintain connectivity with the humans or mobile robots (possibly in a hostile environment) in charge of network deployment. The deployment should ensure the coverage of particular points or areas of interest. The wireless network has to cope with pedestrian mobility and robot/vehicle mobility. The environment, initially unknown, is progressively discovered and may contain numerous obstacles that should be avoided. The nodes of the wireless network are usually battery-powered. Since they are placed by a robot or a human, their weight is very limited. The protocols supported by these nodes should be energy-efficient to maximize network lifetime. In such a challenging environment, sensor nodes should be replaced before their batteries are depleted. It is therefore important to be able to accurately determine the battery lifetime of these nodes, enabling predictive maintenance.

4.6. Types of Wireless Networks

The EVA team will distinguish between opportunistic communication (which takes advantage of a favorable state) and collaborative communication (several entities collaborate to reach a common objective). Furthermore, determinism can be required to schedule medium access and node activity, and to predict energy consumption.

In the EVA project, we will propose **self-adaptive wireless networks** whose evolution is based on:

- optimization to minimize a single or multiple objective functions under some constraints (e.g. interference, or energy consumption in the routing process).
- machine learning to be able to predict a future state based on past states (e.g. link quality in a wireless sensor network) and to identify tendencies.

The types of wireless networks encountered in the application domains can be classified in the following categories.

4.6.1. Wireless Sensor and Mesh Networks

Standardization activities at the IETF have defined an “upper stack” allowing low-power mesh networks to be seamlessly integrated in the Internet (6LoWPAN), form multi-hop topologies (RPL), and interact with other devices like regular web servers (CoAP).

Major research challenges in sensor networks are mostly related to (predictable) power conservation and efficient multi-hop routing. Applications such as monitoring of mobile targets, and the generalization of smart phone devices and wearables, have introduced the need for WSN communication protocols to cope with node mobility and intermittent connectivity.

Extending WSN technology to new application spaces (e.g. security, sports, hostile environments) could also assist communication by seamless exchanges of information between individuals, between individuals and machines, or between machines, leading to the Internet of Things.

4.6.2. Deterministic Low-Power Networks

Wired sensor networks have been used for decades to automate production processes in industrial applications, through standards such as HART. Because of the unreliable nature of the wireless medium, a wireless version of such industrial networks was long considered infeasible.

In 2012, the publication of the IEEE 802.15.4e standard triggered a revolutionary trend in low-power mesh networking: merging the performance of industrial networks, with the ease-of-integration of IP-enabled networks. This integration process is spearheaded by the IETF 6TiSCH working group, created in 2013. A 6TiSCH network implements the IEEE 802.15.4e TSCH protocol, as well as IETF standards such as 6LoWPAN, RPL and CoAP. A 6TiSCH network is synchronized, and a communication schedule orchestrates all communication in the network. Deployments of pre-6TiSCH networks have shown that they can achieve over 99.999% end-to-end reliability, and a decade of battery lifetime.

The communication schedule of a 6TiSCH network can be built and maintained using a centralized, distributed, or hybrid scheduling approach. While the mechanisms for managing that schedule are being standardized by the IETF, which scheduling approach to use, and the associated limits in terms of reliability, throughput and power consumption remain entirely open research questions. Contributing to answering these questions is an important research direction for the EVA team.

4.6.3. MANETs and VANETs

In contrast to routing, other domains in MANETs such as medium access, multi-carrier transmission, quality of service, and quality of experience have received less attention. The establishment of research contracts for EVA in the field of MANETs is expected to remain substantial. MANETs will remain a key application domain for EVA with users such as the military, firefighters, emergency services and NGOs.

Vehicular Ad hoc Networks (VANETs) are arguably one of the most promising applications for MANETs. These networks primarily aim at improving road safety. Radio spectrum has been ring-fenced for VANETs worldwide, especially for safety applications. International standardization bodies are working on building efficient standards to govern vehicle-to-vehicle or vehicle-to-infrastructure communication.

4.6.4. Cellular and Device-to-Device Networks

We propose to initially focus this activity on spectrum sensing. For efficient spectrum sensing, the first step is to discover the links (sub-carriers) on which nodes may initiate communications. In Device-to-Device (D2D) networks, one difficulty is scalability.

For link sensing, we will study and design new random access schemes for D2D networks, starting from active signaling. This will assume the availability of a control channel devoted to D2D neighbor discovery. It is therefore naturally coupled with cognitive radio algorithms (allocating such resources): coordination of link discovery through eNode-B information exchanges can yield further spectrum usage optimization.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Startup Falco wins the Innovation Competition at the Paris Nautic Show (200.000 visitors), December 2019
- Startup Falco wins the Favorite Startup Pitch battle at MassChallenge, Boston, November 2019
- Startup Falco wins the Amplify Pitch battle, October 2019
- Startup Falco awardee of the prestigous Netva “Deeptech North America” program, June 2019
- Startup Falco awardee of the prestigous WILCO/WILCOMET accelerator program
- Amar Abane and **Paul Muhlethaler** receive the best student paper award at PEMWN 2019 for the paper “Modeling and Improving Named Data Networking over IEEE 802.15.4”
- Abdallah Sobehy, under the supervision of **Paul Muhlethaler** and Eric Renault from EVA, secured the first place in the Indoors Positioning Competition held during the IEEE’s Communication Theory Workshop 2019, Selfoss, Iceland. The objective was to localize a transmitter using Channel State Information (CSI) received at a Massive MIMO antenna which is one of the main drivers of the 5G. The contesters were provided with training data to develop algorithms that predict the transmitter’s position from CSI. On the competition day, the teams were given 2000 CSI readings to predict the corresponding positions using the developed algorithms. The evaluation criterion is the Mean Square Error (MSE) of the predicted positions. The proposed method relies on CSI preprocessing and Deep Learning (Multi-Layer Perceptron Neural Network). With an MSE of 2.3 cm, the proposed solution clinched the first place among 8 teams from top universities around the world such as: University of Toronto (Canada), Ruhr University Bochum (Germany). Heriot-Watt University (England), University of Padova (Italy), IMdea networks institute (Spain), Aalborg University (Denmark), and Yuan Ze University (Taiwan).
- Abdellah Sobehy receives the “Le premier prix du jury” during the fifth edition of‘ “la journée doctorants de Samovar”

5.1.2. Transfer

- Internet Engineering Steering Group (IESG) approval of draft-ietf-6tisch-minimal-security to be published as RFC
- Malisa Vucinic named co-chair of the IETF LAKE standardization working group
- Thomas Watteyne co-chairs IETF working group 6TiSCH
- Thomas Watteyne co-chairs IETF design team of fragment forwarding
- Thomas Watteyne part of the IETF IoT directorate
- Startup Falco had a booth at the Paris Nautic Show (Dec 2019, 200.000 visitors)
- Startup Falco had a booth at the Cap d'Agde Nautic Show (Oct-Nov 2019, 50.000 visitors)
- Startup Falco selected to join the Parisian Incubator Agoranov, April 2019

6. New Software and Platforms

6.1. OpenWSN

KEYWORDS: Internet of things - 6TiSCH - 6LoWPAN - CoAP

FUNCTIONAL DESCRIPTION: OpenWSN is an open-source implementation of a fully standards-based protocol stack for the Internet of Things. It has become the de-facto implementation of the IEEE802.15.4e TSCH standard, has a vibrant community of academic and industrial users, and is the reference implementation of the work we do in the IETF 6TiSCH standardization working group.

- Partner: University of California Berkeley
- Contact: Thomas Watteyne
- URL: http://www.openwsn.org/

6.2. 6TiSCH Simulator

High-level simulator of a 6TiSCH network

KEYWORDS: Network simulator - 6TiSCH

FUNCTIONAL DESCRIPTION: The simulator is written in Python. While it doesn't provide a cycle-accurate emulation, it does implement the functional behavior of a node running the full 6TiSCH protocol stack. This includes RPL, 6LoWPAN, CoAP and 6P. The implementation work tracks the progress of the standardization process at the IETF.

- Contact: Malisa Vucinic

6.3. Argus

KEYWORDS: Cloud - Low-Power WIreless - Sniffer

FUNCTIONAL DESCRIPTION: There are three piece to the Argus:

The Argus Probe is the program which attaches to your low-power wireless sniffer and forwards its traffic to the Argus Broker.

The Argus Broker sits somewhere in the cloud. Based on MQTT, it connect Argus Probes with Argus Clients based on a pub-sub architecture.

Several Argus Clients can the started at the same time. It is a program which subscribes to the Argus Broker and displays the frames in Wireshark.

- Contact: Remy Leone

6.4. SolSystem

Sensor Object Library System

KEYWORDS: Low-Power WIreless - Back-End System - SmartMesh IP

FUNCTIONAL DESCRIPTION: The source code is composed of the definition of the SOL structure (https://github.com/realms-team/sol), the code that runs on the manager (https://github.com/realms-team/solmanager, written in Python) and the code that runs on the server receiving the data (https://github.com/realms-team/solserver, written in Python)

- Contact: Keoma Brun-Laguna
- URL: http://www.solsystem.io/

6.5. 6TiSCH Wireshark Dissector

KEYWORDS: 6TiSCH - Wireshark

FUNCTIONAL DESCRIPTION: Implementation on the dissectors is done through an open-source repository, stable code is regularly contributed back to the main Wireshark code base.

- Contact: Jonathan Muñoz

6.6. F-Interop

Remote Conformance and Interoperability Tests for the Internet of Thing

KEYWORDS: Interoperability - Iot - Conformance testing - Standardization

- Partners: UPMC - IMEC - ETSI - EANTC - Mandat International - Digital Catapult - University of Luxembourg - Device Gateway
- Contact: Remy Leone

6.7. Mercator

KEYWORDS: Deployment - Low-Power WIreless - Testbeds - Connectivity

FUNCTIONAL DESCRIPTION: The firmware is written as part of the OpenWSN project. Scripts and analysis tools are written in Python.

- Contact: Keoma Brun-Laguna

7. New Results

7.1. Falco startup launched!

Participants: Elsa Nicol, Keoma Brun-Laguna, Thomas Watteyne.

The Falco startup (https://wefalco.com/) was launched on 14 January 2019. During 2019, it developed a complete technical solution (PCB, assembly, networking, back-end) and completed a large market development campaign. Falco was selected to join the Parisian Incubator Agoranov. It was then awarded the prestigous Netva "Deeptech North America" program, and won the Favorite Startup Pitch battle at MassChallenge, Boston, as well as the Amplify Pitch battle. It had a booth at the Cap d'Agde and Paris Nautic shows. On 14 December 2019, Falco wins the Innovation Competition at the Paris Nautic Show.

7.2. 6TiSCH Standardization

Participants: Malisa Vucinic, Jonathan Muñoz, Tengfei Chang, Yasuyuki Tanaka, Thomas Watteyne.

The standardization work at 6TiSCH remains a strong federator of the work done in the team. In 2019, the working group finalized the work on the draft-ietf-6tisch-minimal-security and draft-ietf-6tisch-architecture specification, which are both in the editor queue. The draft-ietf-6tisch-msf has also passed the working group last call. This standardization work has resulted in several papers on 6TiSCH, including a tutorial [9], [11], [12] fragmentation in 6TiSCH [8], implementation details [17], simulating 6TiSCH [24], experimental approaches [21], [26], localization [25], multi-PHY extensions [10]. The HDR of Thomas Watteyne [2] reports on the work on 6TiSCH over the past years. Some work has started on implementing 6TiSCH on single-chip micro-motes [19], [23], [7], [18].

7.3. 6TiSCH Security

Participants: Malisa Vucinic, Thomas Watteyne.

The security work of Inria-EVA is a continuation of the efforts started during the H2020 ARMOUR project. The work focused on stabilizing the “Minimal Security” solution that has now been approved to be published as an RFC [13]. The solution that is standardized enables secure network access and configuration of 6TiSCH devices under the assumption that they have been provisioned with a secret key. Ongoing work extends this solution to support true zero-configuration network setup, under the assumption that the devices have been provisioned with certificates at manufacturing time.

7.4. 6TiSCH Benchmarking

Participants: Malisa Vucinic, Tengfei Chang, Yasuyuki Tanaka, Thomas Watteyne.

With the pure 6TiSCH standardizes coming to an end, the focus of the group is moving towards benchmarking how well it works. This has resulted in the following action. Although seemingly different, they all contribute to the overall goal of better understanding (the performance of) 6TiSCH.

We have built and put online the OpenTestbed, a collection of 80 OpenMote B boards deployed in 20 “pods”. These allow us to test the performance of the OpenWSN firmware in a realistic setting. You can access its management interface at http://testbed.openwsn.org/.

A tool complementary to the testbed is the 6TiSCH simulator (https://bitbucket.org/6tisch/simulator) which Yatsuyiki Tanaka is leading. The simulator now represents exactly the behavior of the 6TiSCH protocol stack, and has been a catalyst for benchmarking activities around 6TiSCH.

Beyond Inria, the benchmarking activity around 6TiSCH is a hot topic, with projects such as the 6TiSCH Open Data Action [26] (SODA, http://www.soda.ucg.ac.me/), the IoT Benchmarks Initiative (https://www.iotbench.ethz.ch/), and the Computer and Networking Experimental Research using Testbeds (CNERT) workshop at INFOCOM, all of which Inria-EVA is very involved in.

7.5. LAKE Standardization

Participants: Malisa Vucinic, Timothy Claeys, Thomas Watteyne.

In October 2019, a new working group was formed in the IETF with the goal of standardizing a lightweight authenticated key exchange protocol for IoT use cases. The group is co-chaired by Malisa Vucinic of Inria-EVA. Through our work in 6TiSCH and the requirements for the follow up work of the “Minimal Security Framework for 6TiSCH”, we directly contributed to the creation of this working group whose expected output is the key exchange protocol for the IoT. The document we lead in the LAKE working group [37] compiles the requirements for a lightweight authenticated key exchange protocol for OSCORE. OSCORE (RFC8613) is a lightweight communication security protocol providing end-to-end security on application layer for constrained IoT settings. It is expected to be deployed with standards and frameworks using CoAP such as 6TiSCH, LPWAN, OMA Specworks LwM2M, Fairhair Alliance and Open Connectivity Foundation.

7.6. IoT and Low-Power Wireless Meshed Networks

More than 50 billion devices will be connected in 2020. This huge infrastructure of devices, which is managed by highly developed technologies, is called the Internet of Things (IoT). The IoT provides advanced services, and brings economic and societal benefits. This is the reason why engineers and researchers in both industry and scientific communities are interested in this area. The Internet of Things enables the interconnection of smart physical and virtual objects, managed by highly developed technologies. Low-Power Wireless Meshed Network is an essential part of this paradigm. It uses smart, autonomous and usually limited capacity devices in order to sense and monitor their environment.

7.6.1. Centralized or Distributed Scheduling for IEEE 802.15.4e TSCH networks

Participants: Yasuyuki Tanaka, Pascale Minet, Thomas Watteyne, Malisa Vucinic, Tengfei Chang, Keoma Brun-Laguna.

The wireless TSCH (Time Slotted Channel Hopping) network specified in the e amendment of the IEEE 802.15.4 standard has many appealing properties. Its schedule of multichannel slotted data transmissions ensures the absence of collisions. Because there is no retransmission due to collisions, communication is faster. Since the devices save energy each time they do not take part in a transmission, the power autonomy of nodes is prolonged. Furthermore, channel hopping mitigates multipath fading and interferences.

All communication in a TSCH network is orchestrated by the communication schedule it is using. The scheduling algorithm used hence drives the latency and capacity of the network, and the power consumption of the nodes. To increase the flexibility and the self-organizing capacities required by IoT, the networks have to be able to adapt to changes. These changes may concern the application itself, the network topology by adding or removing devices, the traffic generated by increasing or decreasing the device sampling frequency, for instance. That is why flexibility of the schedule ruling all network communications is needed. We have designed a number of scheduling algorithms for TSCH networks, answering different needs. For instance, the centralized Load-based scheduler that assigns cells per flow, starting with the flow originating from the most loaded node has proved optimal for many configurations. Simulations with the 6TiSCH simulator showed that it gets latencies close to the optimal. They also highlighted that end-to-end latencies are positively impacted by message prioritization (i.e. each node transmits the oldest message first) at high loads, and negatively impacted by unreliable links, as presented at GlobeCom 2019 [30].

Among the distributed scheduling algorithms proposed in the literature, many rely on assumptions that may be violated by real deployments. This violation usually leads to conflicting transmissions of application data, decreasing the reliability and increasing the latency of data delivery. Others require a processing complexity that cannot be provided by sensor nodes of limited capabilities. Still others are unable to adapt quickly to traffic or topology changes, or are valid only for small traffic loads. We have designed MSF and YSF, two distributed scheduling algorithms that are adaptive and compliant with the standardized protocols used in the 6TiSCH working group at IETF. The Minimal Scheduling Function (MSF) is a distributed scheduling algorithm in which neighbor nodes locally negotiate adding and removing cells. MSF was evaluated by simulation and experimentation, before becoming the default scheduling algorithm of the IETF 6TiSCH working group, and now an official standard. We also designed LLSF, a scheduling algorithm focused on low latency communication. We proposed a full-featured 6TiSCH scheduling function called YSF, that autonomously takes into account all the aspects of network dynamics, including the network formation phase and parent switches. YSF aims at minimizing latency and maximizing reliability for data gathering applications. Simulation results obtained with the 6TiSCH simulator show that YSF yields lower end-to-end latency and higher end-to-end reliability than MSF, regardless of the network topology. Unlike other top-down scheduling functions, YSF does not rely on any assumption regarding network topology or traffic load, and is therefore more robust in real network deployments. An intensive simulation campaign made with the 6TiSCH simulator has provided comparative performance results. Our proposal outperforms MSF, the 6TiSCH Minimal Scheduling Function, in terms of end-to-end latency and end-to-end packet delivery ratio.

Furthermore we published additional research on computing the upper bounds on the end-to-end latency, finding the best trade-off between latency and network lifetime.

7.6.2. Modeling and Improving Named Data Networking over IEEE 802.15.4

Participants: Amar Abane, Samia Bouzefrane (Cnam), Paul Muhlethaler.

Enabling Named Data Networking (NDN) in real world Internet of Things (IoT) deployments becomes essential to benefit from Information Centric Networking (ICN) features in current IoT systems. One objective of the model is to show that caching can attenuate the number of transmissions generated by broadcast to achieve a reasonable overhead while keeping the data dissemination power of NDN. To design realistic NDN-based communication solutions for IoT, revisiting mainstream technologies such as low-power wireless standards may be the key. We explore the NDN forwarding over IEEE 802.15.4 by modeling a broadcast-based forwarding [27]. Based on the observations, we adapt the Carrier-Sense Multiple Access (CSMA) algorithm of 802.15.4 to improve NDN wireless forwarding while reducing broadcast effects in terms of packet redundancy, round-trip time and energy consumption. As future work, we aim to explore more complex CSMA adaptations for lightweight forwarding to make the most of NDN and design a general-purpose Named-Data CSMA.

7.6.3. Evaluation of LORA with stochastic geometry

Participants: Bartek Blaszczyszyn (Dyogene), Paul Muhlethaler.

We present a simple, stochastic-geometric model of a wireless access network exploiting the LoRA (Long Range) protocol, which is a non-expensive technology allowing for long-range, single-hop connectivity for the Internet of Things. We assume a space-time Poisson model of packets transmitted by LoRA nodes to a fixed base station. Following previous studies of the impact of interference, we assume that a given packet is successfully received when no interfering packet arrives with similar power before the given packet payload phase, see [16]. This is as a consequence of LoRa using different transmission rates for different link budgets (transmissions with smaller received powers use larger spreading factors) and LoRa intra-technology interference treatment. Using our model, we study the scaling of the packet reception probabilities per link budget as a function of the spatial density of nodes and their rate of transmissions. We consider both the parameter values recommended by the LoRa provider, as well as proposing LoRa tuning to improve the equality of performance for all link budgets. We also consider spatially non-homogeneous distributions of LoRa nodes. We show how a fair comparison to non-slotted Aloha can be made within the same framework.

7.6.4. Position Certainty Propagation: A location service for MANETs

Participants: Abdallah Sobehy, Paul Muhlethaler, Eric Renault (Telecom Sud-Paris).

A location method based on triangulation (via Channel State Information (CSI) based localization method is proposed [6]. A known method of triangulation is adopted to deduce the location of a node from 3 reference nodes (anchor nodes). We propose an optimized energy-aware and low computational solution, requiring 3-GPS equipped nodes (anchor nodes) in the network. Moreover, the computations are lightweight and can be implemented distributively among nodes. Knowing the maximum range of communication for all nodes and distances between 1-hop neighbors, each node localizes itself and shares its location with the network in an efficient manner. We simulate our proposed algorithm on a NS-3 simulator, and compare our solution with state-of-the-art methods. Our method is capable of localizing more nodes i.e. $\simeq 90\%$ of nodes in a network with an average degree $\simeq 10$.

7.7. Industry 4.0 and Low-Power Wireless Meshed Networks

The Internet of Things (IoT) connects tiny electronic devices able to measure a physical value (temperature, humidity, etc.) and/or to actuate on the physical world (pump, valve, etc). Due to their cost and ease of deployment, battery-powered wireless IoT networks are rapidly being adopted.

The promise of wireless communication is to offer wire-like connectivity. Major improvements have been made in that direction, but many challenges remain as industrial applications have strong operational requirements. This section of the IoT application is called Industrial IoT (IIoT).

By the year 2020, it is expected that the number of connected objects will exceed several billion devices. These objects will be present in everyday life for a smarter home and city as well as in future smart factories that will revolutionize the industry organization. This is actually the expected fourth industrial revolution, better known as Industry 4.0. In which, the Internet of Things (IoT) is considered as a key enabler for this major transformation. The IoT will allow more intelligent monitoring and self-organizing capabilities than traditional factories. As a consequence, the production process will be more efficient and flexible with products of higher quality.

To produce better quality products and improve monitoring in Industry 4.0, strong requirements in terms of latency, robustness and power autonomy have to be met by the networks supporting the Industry 4.0 applications.

7.7.1. Reliability for the Industrial Internet of Things (IIoT) and Industry 4.0

Participants: Yasuyuki Tanaka, Pascale Minet, Keoma Brun-Laguna, Thomas Watteyne.

The main IIoT requirement is reliability. Every bit of information that is transmitted in the network must not be lost. Current off-the-shelf solutions offer over 99.999% reliability.

To provide the end-to-end reliability targeted by industrial applications, we investigate an approach based on message retransmissions (on the same path). We propose two methods to compute the maximum number of transmissions per message and per link required to achieve the targeted end-to-end reliability. The MFair method is very easy to compute and provides the same reliability over each link composing the path, by means of different maximum numbers of transmissions, whereas the MOpt method minimizes the total number of transmissions necessary for a message to reach the sink. MOpt provides a better reliability and a longer lifetime than MFair, which provides a shorter average end-to-end latency. This study [5] was published in the Sensors journal in 2019.

7.8. Machine Learning applied to Networking

7.8.1. Machine Learning for energy-efficient and QoS-aware Data Centers

Participants: Ruben Milocco (Comahue University, Argentina, Invited Professor), Pascale Minet, Eric Renault (Telecom Sud-Paris), Selma Boumerdassi (Cnam).

To limit global warming, all industrial sectors must make effort to reduce their carbon footprint. Information and Communication Technologies (ICTs) alone generate 2% of global CO_2 emissions every year. Due to the rapid growth in Internet services, data centers have the largest carbon footprint of all ICTs. According to ARCEP (the French telecommunications regulator), Internet data traffic multiplied by 4.5 between 2011 and 2016. In order to support such a growth and maintain this traffic, data centers'energy consumption needs to be optimized.

We determine whether resource allocation in DCs can satisfy the three following requirements: 1) meet user requirements (e.g. short response times), 2) keep the data center efficient, and 3) reduce the carbon footprint.

An efficient way to reduce the energy consumption in a DC is to turn off servers that are not used for a minimum duration. The high dynamicity of the jobs submitted to the DC requires periodically adjusting the number of active servers to meet job requests. This is called Dynamic Capacity Provisioning. This provisioning can be based on prediction. In such a case, a proactive management of the DC is performed. The goal of this study is to provide a methodology to evaluate the energy cost reduction brought by proactive management, while keeping a high level of user satisfaction.

The state-of-the art shows that appropriate proactive management improves the cost, either by improving QoS or saving energy. As a consequence, there is great interest in studying different proactive strategies based on predictions of either the energy or the resources needed to serve CPU and memory requests. The cost depends on 1) the proactive strategy used, 2) the workload requested by jobs and 3) the prediction used. The problem complexity explains why, despite its importance, the maximum cost savings have not been evaluated in theoretical studies.

We propose a method to compute the upper bound of the relative cost savings obtained by proactive management compared to a purely reactive management based on the Last Value. With this method, it becomes possible to quantitatively compare the efficiency of two predictors.

We also show how to apply this method to a real DC and how to select the value of the DC parameters to get the maximum cost savings. Two types of predictors are studied: linear predictors, represented by the ARMA model, and nonlinear predictors obtained by maximizing the conditional probability of the next sample, given the past. They are both applied to the publicly available Google dataset collected over a period of 29 days. We evaluate the largest benefit that can be obtained with those two predictors. Some of these results have been presented at HPCS 2019 [20].

7.8.2. *Machine Learning applied to IoT networks*

Participants: Miguel Landry Foko Sindjoung (Phd Student, Dschang University, Cameroon, Inria Internship), Pascale Minet.

Knowledge of link quality in IoT networks allows a more accurate selection of wireless links to build the routes used for data gathering. The number of re-transmissions is decreased, leading to shorter end-to-end latency, better end-to-end reliability and a longer network lifetime.

We propose to predict link quality by means of machine learning techniques applied on two metrics: the Received Signal Strength Indicator (RSSI) and the Packet Delivery Ratio (PDR). These two metrics were selected because RSSI is a hardware metric that is easily obtained and PDR takes into account packets that are not successfully received, unlike RSSI.

The data set used in this study was collected from a TSCH network deployed in the Grenoble testbed consisting of 50 nodes operating on 16 channels. Data collected by Mercator include 108659 measurements of PDR and average RSSI. We train the model over the training set and predict the link quality on the channel considered for the samples in the validation set. By comparing the predicted values with the real values, the confusion matrix is computed by evaluating the number of true-positive, true-negative, false-positive and false-negative for the link and channel considered.

Whatever the link quality estimator used, RSSI, PDR or both, the Random Forest (RF) classifier model outperforms the other models studied: Linear Regression, Linear Support Vector Machine, Support Vector Machine.

Since using Bad links that have been predicted Good strongly penalizes network performance in terms of end-to-end latency, end-to-end reliability and network lifetime, the joint use of PDR and RSSI improves the accuracy of link quality prediction. Hence, we recommend using the Random Forest classifier applied on both PDR and RSSI metrics. This work has been presented at the PEMWN 2019 conference [33].

7.9. Machine Learnig applied to Smart Farming

Participants: Jamal Ammouri (Internship Cnam), Malika Boudiaf (Ummto, Tizi-Ouzou, Algeria), Samia Bouzefrane (Cnam), Pascale Minet, Meziane Yacoub (Cnam).

Intelligent Farming System (IFS) is made possible by the use of 4 elements: sensors and actuators, the Internet of Things (IoT), edge/cloud processing, and machine learning.

Soil degradation and a hot climate explain the poor yield of olive groves in North Algeria. Edaphic, climatic and geographical data were collected from 10 olive groves over several years and analyzed by means of Self-Organizing Maps (SOMs). SOM is a non-supervised neural network that projects high-dimensional data onto a low-dimension discrete space, called a topological map, such that close data are mapped onto nearby locations on the map. In the paper [28] presented at the PEMWN 2019 conference, we have shown how to use self-organizing maps to determine olive grove clusters with similar features, characterize each cluster and show the temporal evolution of each olive grove. With the SOM, it becomes possible to alert the farmer when some specific action needs to be done in the case of hydric stress, NPK stress, pest/disease attack. As a result, the nutritional quality of the oil produced is improved. SOM can be integrated in the Intelligent Farming System (IFS) to boost conservation agriculture.

This work requires a strong collaboration with agronomists. Malika Boudiaf (Laboratoire Ressources Naturelles, UMMTO, Tizi-Ouzou, Algeria) provided the data set and gave us many explanations about soil conservation. Meziane Yacoub (Cnam) is an expert in SOMs. Jamal Ammouri (Cnam) was co-advised by Samia Bouzefrane, Pascale Minet and Meziane Yacoub.

7.10. Protocols and Models for Wireless Networks - Application to VANETs

7.10.1. Connection-less IoT - Protocol and models

Participants: Iman Hemdoush, Cédric Adjih, Paul Mühlethaler.

The goal is to construct some next-generation access protocols, for the IoT (or alternately for vehicular networks). One starting point are methods from the family of Non-Orthogonal Multiple Access (NOMA), where multiple transmissions can "collide" but can still be recovered - with sophisticated multiple access protocols (MAC) that take the physical layer/channel into account. One such example is the family of the Coded Slotted Aloha methods.Another direction is represented by some vehicular communications where vehicles communicate directly with each other without necessarily going through the infrastructure. This is also true more generally in any wireless network where the control is relaxed (such as in unlicensed IoT networks like LoRa). One observation is that in such distributed scenarios, explicit or implicit forms of signaling (with sensing, messaging, etc.), can be used for designing sophisticated protocols - including using machine learning techniques.

During this study, some of the following tools should be used: protocol/algorithm design (ensuring properties by construction), simulations (ns-2, ns-3, matlab, ...) on detailed or simplified network models, mathematical modeling (stochastic geometry, etc...) ; machine-learning techniques or modeling as code-on-graphs.

The first result we have obtained concerns Irregular Repetition Slotted Aloha (IRSA) which is a modern method of random access for packet networks that is based on repeating transmitted packets, and on successive interference cancellation at the receiver. In classical idealized settings of slotted random access protocols (where slotted ALOHA achieves $1/e$), it has been shown that IRSA could asymptotically achieve the maximal throughput of 1 packet per slot. Additionally, IRSA had previously been studied for many different variants and settings, including the case where the receiver is equipped with “multiple-packet reception” (MPR) capability. We extensively revisit the case of IRSA with MPR. We present a method to compute optimal IRSA degree distributions with a given maximum degree n. A tighter bound for the load threshold (G/K) was proven, showing that plain K-IRSA cannot reach the asymptotic known bound $G/K = 1$ for $K > 1$, and we prove a new, lower bound for its performance. Numerical results illustrate that optimal degree distributions can approach this bound. Second, we analyze the error floor behavior of K-IRSA and provide an insightful approximation of the packet loss rate at low loads, and show its excellent performance. Third, we show how to formulate the search for the appropriate parameters of IRSA as an optimization problem, and how to solve it efficiently. By doing that for a comprehensive set of parameters, and by providing this work with simulations, we give numerical results that shed light on the performance of IRSA with MPR. A final open question is: what is the impact of introducing more structure in the slot selection (like Spatially Coupled Coded Slotted Aloha) and how best to do so?

7.10.2. Indoor positionning using Channel State Information (CSI) from a MIMO antenna

Participants: Abdallah Sobehy, Paul Muhlethaler, Eric Renault (Telecom Sud-Paris).

The channel status information is used for locating a node by applying machine learning [35] techniques. We propose a novel lightweight deep learning solution to the indoor positioning problem based on noise and dimensionality reduction of MIMO Channel State Information (CSI): real and imaginary parts of the signal received. Based on preliminary data analysis, the magnitude of the CSI is selected as the input feature for a Multilayer Perceptron (MLP) neural network. Polynomial regression is then applied to batches of data points to filter noise and reduce input dimensionality by a factor of 14. The MLP’s hyper-parameters are empirically tuned to achieve the highest accuracy. The method is applied to a CSI dataset estimated at an 8 x 2 MIMO antenna that is published by the organizers of the Communication Theory Workshop Indoor

Positioning Competition. The proposed solution is compared with a state-of-the-art method presented by the authors who designed the MIMO antenna that is used to generate the data-set. Our method yields a mean error which is 8 times less than that of its counterpart. We conclude that the arithmetic mean and standard deviation misrepresent the results since the errors follow a log- normal distribution. The mean of the log error distribution of our method translates to a mean error as low as 1.5 cm. We have shown that, using a K-nearest neighbor learning method an even better, indoor positioning is achieved. The input feature is the magnitude component of CSI which is pre-processed to reduce noise and allow for a quicker search. The Euclidean distance between CSI is the criterion chosen for measuring the closeness between samples. The proposed method is compared with three other methods, all based on deep learning approaches and tested with the same data-set. The K-nearest neighbor method presented in this paper achieves a Mean Square Error (MSE) of 2.4 cm, which outperforms its counterparts.

7.10.3. Predicting Vehicles Positions using Roadside Units: a Machine-Learning Approach

Participants: Samia Bouzefrane (Cnam), Soumya Banerjee (Birla Institute Of Technology, Mesra), Paul Mühlethaler, Mamoudou Sangare.

We study positioning systems using Vehicular Ad Hoc Networks (VANETs) to predict the position of vehicles. We use the reception power of the packets received by the Road Side Units (RSUs) and sent by the vehicles on the roads. In fact, the reception power is strongly influenced by the distance between a vehicle and a RSU. We have already used and compared three widely recognized techniques : K Nearest Neighbors (KNN), Support Vector Machine (SVM) and Random Forest. We have studied these techniques in various configurations and discuss their respective advantages and drawbacks. We revisit the positioning problem VANETs but we also consider Neural Networks (NN) to predict the position [22]. The neural scheme we have tested in this paper consists of one hidden layer with three neurons. To boost this technique we use an ensemble neural network with 50 elements built with a bagging algorithm. The numerical experiments presented in this contribution confirm that a precise prediction can only be obtained when there is a main direct path of propagation. The prediction is altered when the training is incomplete or less precise but the precision remains acceptable. In contrast, with Rayleigh fading, the accuracy obtained is much less striking. We observe that the Neural Network is nearly always the best approach. With a direct path the ranking is: Neural Network, Random Forest, KNN and SVM except in the case when we have no measurement in [30m; 105m] where the ranking is Neural Network, Random Forest, SVM and KNN. When there is no direct path, the ranking is SVM, NN, RF and KNN but the difference in performance between SVM and NN is small.

7.10.4. Combining random access TDMA scheduling strategies for vehicular ad hoc networks

Participants: Fouzi Boukhalfa, Mohamed Hadded (Vedecom), Paul Mühlethaler, Oyunchimeg Shagdar (Vedecom).

This work is based on Fouzi Boukhalfa's PhD which started in October 2018, [29],[15]. The idea is to combine TDMA protocols with random access techniques to benefit from the advantages of both techniques. Fouzi Boukhalfa proposes to combine the DTMAC protocol introduced by Mohamed Hadded with a generalization of CSMA. This generalized CSMA uses active signaling; the idea is to send signaling bursts in order to select a unique transmitter. The protocol that Fouzi Boukhalfa obtains reduces the access and merging collisions of DTMAC but can also propose access with low latency for emergency traffic. The idea is that vehicles access their slots reserved with DTMAC but the transmission slots encompass a special section at the beginning with active signaling. The transmission of the signaling burst, during a mini-slot, is organized according to a random binary key. A '1' in the key means that a signaling burst will be transmitted, while a '0' means that the vehicle senses the channel on this mini-slot to potentially find the transmission of a signaling burst by another vehicle. Fouzi Boukhalfa shows that if we use a random key to transmit the signaling burst it very significantly decreases the collision rate (both merging and access collisions) and that emergency traffic can have a very small access delay. Fouzi Boukhalfa builds an analytical model which thoroughly confirms the simulation result. This model can encompass detection error in the selection process of the signaling bursts. It is shown that with a reasonable error rate the performance is only marginaly affected.

7.10.5. Forecasting traffic accidents in VANETs

Participants: Samia Bouzefrane (Cnam), Soumya Banerjee (Birla Institute Of Technology, Mesra), Paul Mühlethaler, Mamoudou Sangare.

Road traffic accidents have become a major cause of death. With increasing urbanization and populations, the volume of vehicles has increased exponentially. As a result, traffic accident forecasting and the identification of the accident prone areas can help reduce the risk of traffic accidents and improve the overall life expectancy.

Conventional traffic forecasting techniques use either a Gaussian Mixture Model (GMM) or a Support Vector Classifier (SVC) to model accident features. A GMM on the one hand requires large amount of data and is computationally inexpensive, SVC on the other hand performs well with less data but is computationally expensive. We present a prediction model that combines the two approaches for the purpose of forecasting traffic accidents. A hybrid approach is proposed, which incorporates the advantages of both the generative (GMM) and the discriminant model (SVC). Raw feature samples are divided into three categories: those representing accidents with no injuries, accidents with non incapacitating injuries and those with incapacitating injuries. The output or the accident severity class was divided into three major categories namely: no injury in the accident, non-incapacitating injury in the accident and an incapacitating injury in the accident. A hybrid classifier is proposed which combines the descriptive strength of the baseline Gaussian mixture model (GMM) with the high performance classification capabilities of the support vector classifier (SVC). A new approach is introduced using the mean vectors obtained from the GMM model as input to the SVC. The model was supported with data pre-processing and re-sampling to convert the data points into suitable form and avoid any kind of biasing in the results. Feature importance ranking was also performed to choose relevant attributes with respect to accident severity. This hybrid model successfully takes advantage of both models and obtained a better accuracy than the baseline GMM model. The radial basis kernel outperforms the linear kernel by achieving an accuracy of 85.53%. Data analytics performed including the area under the receiver operating characteristics curve (AUC-ROC) and area under the precision/recall curve(AUC-PR) indicate the successful application of this model in traffic accident forecasting. Experimental results show that the proposed model can significantly improve the performance of accident prediction. Improvements of up to 24% are reported in the accuracy as compared to the baseline statistical model (GMM). The data about circumstances of personal injury in road accidents, the types of vehicles involved and the consequential casualties were obtained from data.govt.uk.

Although a significant improvement in accuracy has been observed, this study has several limitations. The first concerns the dataset used. This research is based on a road traffic accident dataset from the year of 2017 which contains very few data samples for the no injury and non-incapacitating injury types of accident. The data was unbalanced not just with respect to the output class but also with respect to the sub features of various attributes. Moreover, aggregating the accident severity into just three categories limits the scope of the study and the results obtained. The greater the number of severity classes, the less is the amount of extra training data required to feed in the SVC to avoid overfitting. Thus, datasets with sufficient records corresponding to each class are desirable and must be used for further study.
The second limitation concerns the dependence of the SVC model on parameters and attribute selection. In this study, the performance of SVC relies heavily on the feature selection results and the mean vectors obtained from the GMM. In order to improve the accuracy of the support vector classifier, other approaches like particle swarm optimization (PSO), ant colony optimization, genetic algorithms etc. could be used for effective parameter selection. In addition to this, more kernels like the polynomial kernel and the sigmoid kernel could be tested to improve future model performances.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

Participants: Razanne Abu Aisheh, Mina Rady, Thomas Watteyne.

Razanne Abu Aisheh is doing her PhD under a CIFRE agreement between Inria and Nokia Bell Labs. Mina Rady is doing his PhD under a CIFRE agreement between Inria and Orange Labs.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. Inria Project Labs, Exploratory Research Actions and Technological Development Actions

- IPL SPARTA
- ATT SmartMarina, 2019. Help transfer the technology of the SmartMarina project to startup Falco. Keoma Brun-Laguna is lead.
- ADT 6TiSCH, 2018-2020. Benchmark the performance of 6TiSCH under realistic scenarios, through experimentaion using the OpenTestbed. Tengfei Chang is lead.
- ADT DASMU (Distributed Adaptive Scheduling for MUltichannel Wireless Sensor Networks), 2018-2019. DASMU focuses on a distributed scheduling algorithm which relies on realistic assumptions, does not require complex computation, is valid for any traffic load, is adaptive and compliant with the standardized protocols used in the 6TiSCH working group at IETF. First results have been obtained and an intensive simulation campaign made with the 6TiSCH simulator has provided comparative performance results. Our proposal, called YSF, outperforms MSF, the 6TiSCH Minimal Scheduling Function, in terms of end-to-end latency and end-to-end packet delivery ratio. Thanks to this ADT, Yasuyuki Tanaka has joined the EVA team for two years.

9.1.2. ANR

- The GeoBot FUI project (https://geobot.fr/) is one of the most innovative, challenging and fun projects around wireless localization in the world today. It applies true innovation to a real-world problem, with a clear target application (and customer) in mind. The GeoBot partners are building a small robot (think of a matchbox-sized RC car) that will be inserted into a gas pipe, and move around it to map the location of the different underground pipes. Such mapping is necessary to prevent gas-related accidents, for example during construction. At the end of the project, this solution will be commercialized and used to map the network of gas pipe in France, before being used in worldwide. Each partner is in charge of a different aspect of the problem: robotics, analysis of the inertial data, visualization, etc. Inria is in charge of the wireless part. We will be equipping the robot with a wireless chip(set) in order to (1) communicate with the robot as it moves about in the pipes while standing on the surface, and (2) discover the relative location of the robot w.r.t. a person on the surface. Inria is evaluating different wireless technologies, benchmarking around ranging accuracy and capabilities to communicate. We start from off-the-shelf kits from different vendors and build a custom board, benchmark it, and integrate it with the other partners of the project.

9.1.3. Other collaborations

- EVA has a collaboration with Orange Labs. **Thomas Watteyne** supervises the PhD of Mina Rady, which happens under a CIFRE agreement with Orange Labs.
- EVA has a collaboration with Vedecom. **Paul Muhlethaler** supervises Fouzi Boukhalfa's PhD funded by Vedecom. This PhD aims at studying low latency and high reliability vehicle-to-vehicle communication to improve roads safety.
- EVA has an ongoing collaboration with SODEAL company, which runs the Cap d'Agde marina, as part of Falco startup.
- EVA has an ongoing collaboration with SELOR company, which runs the Lorient marinas, as part of the Falco startup.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

The H2020 following project is ongoing:

- H2020 SPARTA, Jan 2019 – December 2020.

9.2.2. Collaborations in European Programs, Except FP7 & H2020

Inria-EVA has collaboration in 2018 with ETSI (the European Telecommunications Standards Institute) to organize the F-Interop 6TiSCH 2 Interop Event on 2-4 February 2018 in Paris.

9.3. International Initiatives

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

9.3.1.1. REALMS

- Title: Real-Time Real-World Monitoring Systems
- International Partner (Institution - Laboratory - Researcher):
  - University of California Berkeley (United States) - Civil and Environmental Engineering - Steven Glaser
  - University of Michigan (United States) - Civil and Environmental Engineering - Branko Kerkez
- Start year: 2015
- See also: http://glaser.berkeley.edu et http://www-personal.umich.edu/~bkerkez/
- The Internet of Things revolution prompted the development of new products and standards; The IEEE 802.15.4e (2012) standard introduced the Time Synchronized Channel Hoping (TSCH) which can provide end-to-end reliability of 99.999 % and an energy autonomy of many years. This exceptional performance prompted the IETF to create the 6TISCH working group to standardize the integration of TSCH networks in the Internet. While the first experimental data have highlighted the great robustness of these networks, there is no data of a real network, accessible in real time, on a large scale and over a long period. Such data is needed to better model network performance and produce better products and standards. The teams of Professors Glaser and Kerkez are successfully deploying such networks to study mountain hydrology, monitor water quality and manage rainwater in urban environments. A model is missing to assist in the deployment and operation of these networks, as well as to monitor an operational network.

9.3.2. Inria International Partners

9.3.2.1. Declared Inria International Partners

Inria-EVA has a long-standing Memorandum of Understanding with the OpenMote company (http://www.openmote.com/), which runs until 2020. OpenMote emerged as a spin-off of the OpenWSN project, co-led by **Thomas Watteyne** and Prof. Xavier Vilajosana, Professor at the Open University of Catalonia and Chief Technical Officer at OpenMote.

The collaboration has been ongoing since 2012 and at the time of writing has resulted in:

- Joint academic publications, including 7 journal articles, 1 letter, 1 book chapter, 5 conference papers, 2 tutorials and invited talks.
- Joint standardization activities, in particular in the IETF 6TiSCH working group, co-chaired by **Thomas Watteyne** and for which Prof. Xavier Vilajosana is a key contributor. This activity has resulted in the joint participation in 12 IETF face-to-face meetings, joint participation in over 100 audioconferences, co-authorship of 3 Internet-Drafts and joint organization of 2 interop events.
- Joint software development, as both institutions closely collaborate in the maintenance, development, promotion and research along the OpenWSN project, including the development of the protocol stack, the integration of novel hardware technologies, the support to the community and the participation in standardization activities and interoperability events.

This MOU is NOT a commitment of funds by any party.

9.3.2.2. Informal International Partners

The Inria-EVA team collaborates extensively with Prof. Pister's group at UC Berkeley on the OpenWSN and Smart Dust projects. This activity translated into several members of the Pister team visiting Inria-EVA and vice-versa in 2018.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

1. **Martina Brachmann (RISE, Sweden)** (November 2019) working on TSCH on the RISE: Current Research and Future Directions in Networked Embedded Systems with Thomas Watteyne, Malisa Vucinic, Tengfei Chang
2. **Ana Laura Diedrichs (UTN, Argentina)** (Oct-Nov 2019) working on WirelessWine with Thomas Watteyne, Keoma Brun-Laguna
3. **Prof Leila Seidane Azouz** 1-30 October 2019 working with **Pascale Minet** and **Paul Muhlethaler** on wireless networks.
4. **Prof Ruben Milocco** visited EVA from 1-30 October 2019 working with **Pascale Minet** on evaluation of data center performance and **Paul Muhlethaler** on wireless network relaying.
5. **Prof. Diego Dujovne (UDP, Chile)** (July 2019) working on WirelessWine with Thomas Watteyne
6. **Prof. Branko Kerkez (U. Michigan)** (May 2019) working on REALMS associate team with Thomas Watteyne
7. **Mikolaj Chwalisz (TU Berlin)** (May 2019) working on Towards efficient coexistence of IEEE 802.15.4e TSCH and IEEE 802.11 Collaboration with Tengfei Chang, Thomas Watteyne

9.4.1.1. Internships

1. Amy Hane, Intern, from Sep 2019 until Dec 2019
2. Camilo Andres Lopez Lopez, Intern, from May 2019 until Aug 2019
3. Ba Hai Le, Intern, Apr-Aug 2019
4. Victor Kenichi Nascimento Kobayashi,Intern, from May 2019 until Aug 2019
5. Sharut Gupta, Intern, from May 2019 until July 2019
6. Miguel Landry Foko Sindjoung, Intern, from Mar 2019 until Jun 2019.

9.4.1.2. Research Stays Abroad

- **Thomas Watteyne** spent the month of August 2019 at UC Berkeley, working with Prof. Glaser on the SnowHow project, and with Prof. Pister on Smart Dust and OpenWSN.
- Tengfei Chang spent June 2019 in California working with Prof. Pister working on Smart Dust UC Berkeley.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organization

10.1.1.1. General Chair, Scientific Chair

- **Thomas Watteyne** is TPC co-chair of the 6th International Workshop on Computer and Networking Experimental Research using Testbeds (CNERT), held in conjunction with IEEE INFOCOM, Paris, France, 29 April-2 May, 2019. More information at https://infocom2019.ieee-infocom.org/cnert-computer-and-networking-experimental-research-using-testbeds.

- **Paul Muhlethaler** was general co-chair with Eric Renault of the second conference of application of Machine Learning for Networks (MLN 2019 3-5 December 2019), a conference hosted by Inria Paris. Two tutorials and two keynotes were given:
  - *Cloud Solution Architect (Artificial Intelligence and Machine Learning)* by Franck Gaillard (Microsoft, France) [tutorial 1].
  - *Towards intelligent mobile networks* by Marie Line Alberi Morel (NOKIA Bell Lab, France) and Kamal Singh (Télécom Saint-Etienne / University Jean Monnet, France) [tutorial 2].
  - *From Learning to Reasoning : A topos perspective* by Jean-Claude Belfiore (Huawei, France) [keynote 1].
  - *Paradigm shifts in communication networks: cloud-native, automation and AI/ML* by Alberto Conte (NOKIA Bell Labs, France) [keynote 2]

  Thirty two technical presentations have been given during the three days of the MLN 2019 conference.
- **Pascale Minet** was general co-chair with Leila Saidane from ENSI (Tunisia) of the PEMWN 2019 conference, the 8th IFIP/IEEE international conference on Performance Evaluation and Modeling of Wired and Wireless Networks, technically co-sponsored by IFIP WG6.2 and IEEE ComSoc. This conference was held in Paris (Inria Research Center), the 26th, 27th and 28th of November 2018. Three tutorials were given:
  - *Information Centric Networks: Toward an Internet of Named Objects* by Andrea Araldo, Telecom Sud-Paris, France.
  - *Shannon versus Turing, the limit of Artificial Intelligence* by Philippe Jacquet, Inria, Paris-Saclay, France.
  - *5G: Key technologies, standard and evolutions* by Marceau Coupechoux, Telecom Paris-Tech, France.

  Fourteen technical presentations have been given during the three days of the PEMWN 2019 conference.

10.1.1.2. Member of the Organizing Committees

- **Paul Muhlethaler** organized the DGA Inria workshop on the topic ”Au dela des communications quantiques, les réseaux quantique” in May 2019.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- **Paul Muhlethaler** was a Steering committee member of IWVSC 2019.
- **Pascale Minet** was chair of the Technical Program Committee of PEMWN 2019, 8th International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks, November 2019.

10.1.2.2. Member of Conference Program Committees

- **Pascale Minet**
  - CoRes 2019, 4emes Rencontres Francophones sur la Conception de Protocoles, l’Evaluation de Performance et l’Expérimentation des Réseaux de Communication, May 2019.
  - EUSPN 2019, 10th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN), November 2019.,
  - GlobeCom 2019, IEEE Global Communications Conference, December 2019.
  - IWVSC 2019, 3rd International Workshop on Vehicular Adhoc Networks for Smart Cities, November 2019,

  - MLN 2019, Machine Learning for Networking, December 2019,
  - MSPN 2019, 5th International Conference on Mobile, Secure and Programmable Networking, June 2019,
  - PECCS 2019, 9th International conference on Pervasive and Embedded Computing and Communication Systems, September 2019,
  - VTC 2019, 88th IEEE Vehicular Technology Conference, September 2019,
  - Wireless Days 2019, IFIP/IEEE Wireless Days, April 2019,
  - WiSEE 2019, 7th IEEE International Conference on Wireless for Space and Extreme Environments, October 2019.
- **Paul Muhlethaler**:
  - ISCC 2019, 30 June -28 July 2019, Barcelona Spain,
  - MLN 2019, Machine Learning for Networking, December 2019, Paris.
  - PEMWN 2019, 8th International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks, 26 - 28 November 2019, Paris,
  - Wireless Days, IFIP/IEEE Wireless Days 24 - 26 April 2019, Manchester

10.1.3. Journal

10.1.3.1. Reviewer - Reviewing Activities

- **Paul Muhlethaler**
  - Reviewer Ad Hoc Networks Journal (Elsevier),
  - Reviewer Annals of Telecommunications,
  - Reviewer International Journal of Distributed Sensor Networks. Hindawi,
  - Reviewer IEEE Transactions on Information Theory,
  - Reviewer IEEE Transactions on Vehicular Technology,
  - Reviewer IEEE Transactions on Wireless Communications,
  - Reviewer MDPI Sensors
- **Pascale Minet**
  - Acta Astronautica,
  - Ad Hoc Networks,
  - Annals of Telecommunications,
  - Computer Communications,
  - Computer Networks,
  - Engineering Applications of Artificial Intelligence,
  - Future Internet,
  - IEEE Access,
  - IEEE Internet of Things,
  - IEEE Transactions on Mobile Computing,
  - IEEE Transactions on Industrial Informatics,
  - International Journal of Advanced Intelligence Paradigms,
  - International Journal of Communication Systems,
  - Sensors Journal,
  - Wireless Networks.
- Nadjib Achir

- Reviewer Sensor Networks (MDPI)
- Reviewer Wireless Communications and Mobile Computing (Wiley)
- Reviewer Internet of Things Journal (IEEE)
- Reviewer Ad Hoc Networks Journal (Elsevier)

- Selma Boumerdassi
  - Reviewer Ad Hoc Networks Journal (Elsevier);
  - Reviewer The journal of Future Generation Computer Systems (Elsevier).
- Samia Bouzefrane
  - The International Journal of Computer and Telecommunications Networking (Elsevier),
  - The IEEE Transactions on Mobile Computing,
  - The Information and Software Technology Journal (Elsevier)
  - The Springer Multimedia Tools and Application Journal
  - The ACM Transaction on Internet Technology
  - The Concurrency and Computation Practice and Experience Journal
  - the Journal of Systems and Software (Elsevier)

10.1.4. Invited Talks

- **Thomas Watteyne**
  - SmartMesh IP. Captronic/Arrow mesh networking day. Montpellier, France, 13 June 2019.
  - OpenTestbed; Poor Man's IoT Testbed. TILECS Workshop. Grenoble, France. 4 July 2019.
  - SmartMesh IP. Captronic/Arrow mesh networking day. Toulouse, France, 18 April 2019.
- **Pascale Minet**
  - "Estimation and Prediction of Link Quality in Low-Power Wireless Mesh Networks" at the Machine Learning for Networking (MLN) conference, December 2019, Paris.

10.1.5. Leadership within the Scientific Community

Thomas Watteyne co-chairs the IETF 6TiSCH standardization group. **Malisa Vucinic** co-chairs the IETF LAKE standardization group.

10.1.6. Scientific Expertise

- **Thomas Watteyne** regularly consults with major player in the (Industrial) IoT space.
- **Pascale Minet** was also reviewer of a proposal submitted to the Call for Generic Projects issued by the ANR.

10.1.7. Research Administration

- **Thomas Watteyne** is member of the Inria-Paris "Commission de Developpement Technologique", since 2018, where we ensure Inria project teams get sufficient engineering resources to change the world.
- **Paul Muhlethaler** is member of the Inria-Paris "Comite de Centre", since 2016 (alternate of Michel Kern).
- **Thomas Watteyne** is member of the Inria-Paris "Comite de Centre", since 2016, where we work on making sure Inria-Paris will always remain one of the greatest places to work at!
- **Pascale Minet** is member of the Inria-Paris "Commission d'Emplois Scientifiques" (Delegations, Postdocs, and PhD scholarship) since 2018.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- **Thomas Watteyne** teaches 6-week course on IoT, with associated hands-on labs. Undergraduate level. ENSTA ParisTech. Together with Tarak Arbi and Dominique Barthel. Spring 2019.
- **Thomas Watteyne**teaches 1/2-day crash course on the Industrial IoT, Telecom ParisTech. Graduate level. 27 September 2019.
- **Thomas Watteyne**teaches 1-day hands-on course on IIoT, and support of subsequent projects, MsC level, University College London, 6 February 2019.

10.2.2. Supervision

- PhD : Fouzi Boukhalfa, Low and high reliability access in Vehicular Ad-hoc NETworks. Sorbonne University. **Paul Muhlethaler**.
- PhD (ongoing) Razanne Abu Aisheh, Robotics, Sorbonne University. **Thomas Watteyne**,
- PhD (ongoing) Mina Rady, Heterogeneous architectures for the IoT, Sorbonne University. **Thomas Watteyne** and **Paul Muhlethaler**, under a CIFRE agreement with Orange Labs, Meylan, France.
- PhD (defended) Jonathan Munoz, Time slotted systems for long range communications, Sorbonne University, **Thomas Watteyne** and **Paul Muhlethaler**.
- PhD (in progress) Amar Abane, Name Data Networks in the Internet of Things, Cnam. Samia Bouzefrane and **Paul Muhlethaler**.
- PhD (in progress) Abdallah Soheby, Etude et evaluation de la dissemination des informations dans la 5G. Eric Renault and **Paul Muhlethaler**.
- PhD (in progress) Mamoudou Sangara, Utilisation de techniques de Machine Learning dans les reseaux VANETs. Samia Bouzefrane and **Paul Muhlethaler**.
- PhD (in progress) Iman Hmedoush, Connection protocols for the 5G IoT. Cedric Adjih and **Paul Muhlethaler**.

10.2.3. Juries

- HdR:
  - Hichem Sedjelmaci “Attacks Detection and Prediction Frameworks Based on AI Techniques for Wireless and Mobile Networks” Univeristy de Burgundy, October 2019 **Paul Muhlethaler** reviewer.
  - Katia Jaffres-Runser, "Modélisations et Optimisation Multicritere de Réseaux", Institut National Polytechnique de Toulouse, July 2019, **Pascale Minet** reviewer.
- PhD:
  - Vasileios Kotsiou, “Reliable Communications for the Industrial Internet of Things”, University of Strasbourg, France, **Thomas Watteyne** examiner.
  - Meriem Smache, “Deterministic Network Security for the Industrial Internet of Things (IIoT)”, CEA and École des Mines de Saint-Étienne, France, **Thomas Watteyne** examiner.
  - Esteban Municio, “Towards Scalable and End-to-End Programmable Industrial Internet of Things”, University of Antwerpen , Belgium, **Thomas Watteyne** examiner.
  - Amel Arfaoui. “Game based for security of Intenet of things” Univeristy de Burgundy, 15 October 2019 **Paul Muhlethaler** examiner
  - Amar Abane. “Une Architecture Réaliste pour l’IOT” Cnam and University of Tizi Ouzou 2 December 2019 **Paul Muhlethaler** examiner

  - Karen Boulos, "BBU-RRH Association Optimization in Cloud-Radio Access Networks", University of Paris-Saclay, July 2019, **Pascale Minet** reviewer.
  - Jinpeng Wang, "Impact of Mobility and Deployment in Confined Spaces on Low Power and Lossy Network", University of Clermont-Auvergne, July 2019, **Pascale Minet** examiner.
  - Yoann Desmouceaux, "Network Layer Protocols for Data Center Scalability", University of Paris-Saclay, April 2019, **Pascale Minet** examiner.

10.3. Popularization

10.3.1. Dans les Medias

- **Thomas Watteyne**
  - Falco : Des capteurs pour connecter ports et plaisanciers. BoatIndustry, 15 November 2019
  - Agde : le port connecté, une réalité. Midi Libre, 31 October 2019.
  - Gestion portuaire : les ports de Lorient et Cap d'Agde entrent dans l'ère du port connecté, Actu Maritime, 1 August 2019.
  - Port connecté : un nouveau cap franchi grâce à Falco (an interview of our startup's CEO Elsa Nicol), inria.fr, 18 February 2019.

10.3.2. Creation of media or tools for science outreach

A video concerning the research on protocols for VANETs by Mohamed Hadded has been made by Inria.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] J. M. Munoz Soto.*km-scale Industrial Networking*, UPMC - Sorbonne University, March 2019, https://hal.inria.fr/tel-02285706

[2] T. Watteyne.*The Internet of (Important) Things*, Sorbonne Universite, May 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02135069

Articles in International Peer-Reviewed Journal

[3] A. Abane, M. Daoui, S. Bouzefrane, P. Muhlethaler.*A lightweight forwarding strategy for Named Data Networking in low-end IoT*, in "Journal of Network and Computer Applications", 2019, vol. 148, 102445 [*DOI :* 10.1016/j.jnca.2019.102445], https://hal.archives-ouvertes.fr/hal-02334088

[4] A. Abane, M. Daoui, S. Bouzefrane, P. Muhlethaler.*NDN-over-ZigBee: A ZigBee support for Named Data Networking*, in "Future Generation Computer Systems", April 2019, vol. 93, p. 792-798 [*DOI :* 10.1016/j.future.2017.09.053], https://hal.archives-ouvertes.fr/hal-02411296

[5] P. Minet, Y. Tanaka.*Optimal Number of Message Transmissions for Probabilistic Guarantee in the IoT*, in "Sensors", September 2019, https://hal.archives-ouvertes.fr/hal-02433637

[6] A. SOBEHY, E. RENAULT, P. MUHLETHALER.*Position certainty propagation: a localization service for Ad-Hoc Networks*, in "Computers", January 2019, vol. 8, n^o^ 1, Article 6 [*DOI :* 10.3390/COMPUTERS8010006], https://hal.archives-ouvertes.fr/hal-02178389

[7] I. SUCIU, F. MAKSIMOVIC, B. WHEELER, D. BURNETT, O. KHAN, T. WATTEYNE, X. VILAJOSANA, K. PISTER.*Dynamic Channel Calibration on a Crystal-Free Mote-on-a-Chip*, in "IEEE Access", August 2019 [*DOI :* 10.1109/ACCESS.2019.2937689], https://hal.inria.fr/hal-02420914

[8] Y. TANAKA, P. MINET, T. WATTEYNE.*6LoWPAN Fragment Forwarding*, in "IEEE Communications Standards Magazine", July 2019, vol. 3, n^o^ 1, p. 35-39, https://hal.inria.fr/hal-02061838

[9] P. THUBERT, X. V. GUILLEN, T. WATTEYNE.*6TiSCH, enabling IPv6 over 802.15.4 TSCH*, in "IEEE Communications Standards Magazine", March 2019, https://hal.inria.fr/hal-02420677

[10] P. TUSET-PEIRÓ, F. VAZQUEZ-GALLEGO, J. M. MUNOZ SOTO, T. WATTEYNE, J. ALONSO-ZARATE, X. VILAJOSANA.*Experimental Interference Robustness Evaluation of IEEE 802.15.4-2015 OQPSK-DSSS and SUN-OFDM Physical Layers for Industrial Communications*, in "Sensors", September 2019, https://hal.inria.fr/hal-02420946

[11] X. VILAJOSANA, T. WATTEYNE, T. CHANG, M. VUČINIĆ, S. DUQUENNOY, P. THUBERT.*IETF 6TiSCH: A Tutorial*, in "Communications Surveys and Tutorials, IEEE Communications Society", September 2019, forthcoming, https://hal.inria.fr/hal-02420974

[12] X. VILAJOSANA, T. WATTEYNE, M. VUČINIĆ, T. CHANG, K. PISTER.*6TiSCH: Industrial Performance for IPv6 Internet of Things Networks*, in "Proceedings of the IEEE", June 2019, vol. 107, n^o^ 6, p. 1153 - 1165 [*DOI :* 10.1109/JPROC.2019.2906404], https://hal.inria.fr/hal-02266569

[13] M. VUČINIĆ, J. SIMON, K. PISTER, M. RICHARDSON.*Constrained Join Protocol (CoJP) for 6TiSCH*, in "Internet Engineering Task Force", 2019, forthcoming, https://hal.inria.fr/hal-02407587

[14] Z. ZHANG, S. D. GLASER, T. WATTEYNE, S. MALEK.*Long-term Monitoring of the Sierra Nevada Snowpack Using Wireless Sensor Networks*, in "IEEE internet of things journal", June 2020, https://hal.inria.fr/hal-01388391

International Conferences with Proceedings

[15] F. BOUKHALFA, M. HADDED, P. MUHLETHALER, O. SHAGDAR.*An Analytical Model for Performance Analysis of an Active Signaling-based TDMA MAC Protocol for Vehicular Networks*, in "VTC-FALL 2019 - IEEE Vehicular Technology Conference", Honolulu, Hawaii, United States, September 2019, https://hal.archives-ouvertes.fr/hal-02178355

[16] B. BŁASZCZYSZYN, P. MUHLETHALER.*Analyzing LoRa long-range, low-power, wide-area networks using stochastic geometry*, in "VALUETOOLS 2019 - 12th EAI International Conference on Performance Evaluation Methodologies and Tools", Palma de Mallorca, Spain, Proc. ACM VALUETOOLS, ACM, 2019, https://arxiv.org/abs/1811.01886 - 8 pages [*DOI :* 10.1145/3306309.3306327], https://hal.inria.fr/hal-01958939

[17] T. CHANG, T. WATTEYNE, X. VILAJOSANA.*Competition: OpenWSN, a Development Environment for 6TiSCH*, in "International Conference on Embedded Wireless Systems and Networks (EWSN)", Beijing, China, February 2019, https://hal.inria.fr/hal-01968644

[18] T. CHANG, T. WATTEYNE, B. WHEELER, F. MAKSIMOVIC, O. KHAN, S. MESRI, L. LEE, K. PISTER, I. SUCIU, X. VILAJOSANA.*Demo: 6TiSCH on SCμM, Running a Synchronized Protocol Stack without Crystals*, in "EWSN 2020 - ACM International Conference on Embedded Wireless Systems and Networks", Lyon, France, February 2020, https://hal.inria.fr/hal-02436262

[19] F. MAKSIMOVIC, B. WHEELER, D. BURNETT, O. KHAN, S. MESRI, I. SUCIU, L. LEE, A. MORENO, A. SUNDARARAJAN, B. ZHOU, R. ZOLL, A. NG, T. CHANG, X. VILAJOSANA, T. WATTEYNE, A. NIKNEJAD, K. PISTER.*A Crystal-Free Single-Chip Micro Mote with Integrated 802.15.4 Compatible Transceiver, sub-mW BLE Compatible Beacon Transmitter, and Cortex M0*, in "2019 VLSI - IEEE Symposium on VLSI Technology & Circuits", Kyoto, Japan, June 2019 [*DOI :* 10.23919/VLSIC.2019.8777971], https://hal.inria.fr/hal-02420868

[20] R. MILOCCO, P. MINET, E. RENAULT, S. BOUMERDASSI.*Cost reduction bounds of proactive management based on request prediction*, in "HPCS 2019: 17th International Conference on High Performance Computing & Simulation", Dublin, Ireland, 2019 International Conference on High Performance Computing & Simulation (HPCS), IEEE Computer Society, 2019, p. 1-10, https://hal.archives-ouvertes.fr/hal-02432652

[21] J. M. MUNOZ SOTO, F. RINCON, T. CHANG, X. VILAJOSANA, B. VERMEULEN, T. WALCARIUS, W. VAN DE MEERSSCHE, T. WATTEYNE.*OpenTestBed: Poor Man's IoT Testbed*, in "IEEE INFOCOM - CNERT : Workshop on Computer and Networking Experimental Research using Testbeds", Paris, France, April 2019, https://hal.inria.fr/hal-02266558

[22] M. SANGARE, D.-V. N. NGUYEN, S. BANERJEE, P. MUHLETHALER, S. BOUZEFRANE.*Comparing different Machine-Learning techniques to predict Vehicles' Positions using the received Signal Strength of periodic messages*, in "WMNC 2019. 12th IFIP Wireless and Mobile Networking Conference", Paris, France, September 2019, https://hal.archives-ouvertes.fr/hal-02178360

[23] I. SUCIU, F. MAKSIMOVIC, D. BURNETT, O. KHAN, B. WHEELER, A. SUNDARARAJAN, T. WATTEYNE, X. VILAJOSANA, K. PISTER.*Experimental Clock Calibration on a Crystal-Free Mote-on-a-Chip*, in "IEEE INFOCOM - CNERT 2019 : International Workshop on Computer and Networking Experimental Research using Testbeds", Paris, France, April 2019, https://hal.inria.fr/hal-02266563

[24] Y. TANAKA, K. BRUN-LAGUNA, T. WATTEYNE.*Demo: Simulating a 6TiSCH Network using Connectivity Traces from Testbeds*, in "CNERT - IEEE INFOCOM Workshop on Computer and Networking Experimental Research using Testbeds", Paris, France, April 2019, https://hal.inria.fr/hal-02266552

[25] Y. TANAKA, H. LE, V. KOBAYASHI, C. LOPEZ, T. WATTEYNE, M. RADY.*Demo: Blink - Room-Level Localization Using SmartMesh IP*, in "EWSN 2020 - ACM International Conference on Embedded Wireless Systems and Networks", Lyon, France, February 2020, https://hal.inria.fr/hal-02436263

[26] M. VUČINIĆ, B. SKRBIC, E. KOCAN, M. PEJANOVIC-DJURISIC, T. WATTEYNE.*OpenBenchmark: Repeatable and Reproducible Internet of Things Experimentation on Testbeds*, in "CNERT - IEEE INFOCOM Workshop on Computer and Networking Experimental Research using Testbeds", Paris, France, April 2019, https://hal.inria.fr/hal-02266556

Conferences without Proceedings

[27] A. ABANE, P. MUHLETHALER, S. BOUZEFRANE, A. BATTOU.*Modeling and Improving Named Data Networking over IEEE 802.15.4*, in "IFIP/IEEE PEMWN 2019 : IFIP/IEEE International Conference on

Performance Evaluation and Modeling in Wired and Wireless Networks", Paris, France, November 2019, https://hal.archives-ouvertes.fr/hal-02333795

[28] J. AMMOURI, P. MINET, M. BOUDIAF, S. BOUZEFRANE, M. YACOUB.*Self-Organizing Maps Applied to Soil Conservation in Mediterranean Olive Groves*, in "PEMWN2019 - 8th IFIP/IEEE International Conference on Performance Evaluation and Modeling inWired andWireless Networks", Paris, France, November 2019, https://hal.archives-ouvertes.fr/hal-02432848

[29] F. BOUKHALFA, M. HADDED, P. MUHLETHALER, O. SHAGDAR.*An active signaling mechanism to reduce access collisions in a distributed TDMA based MAC protocol for vehicular networks*, in "AINA-2019 - Advanced Information Networking and Applications", Matsue, Japan, March 2019, https://hal.archives-ouvertes.fr/hal-02177650

[30] K. BRUN-LAGUNA, P. MINET, Y. TANAKA.*Optimized Scheduling for Time-Critical Industrial IoT*, in "GLOBECOM 2019 - 2019 IEEE Global Communications Conference", Waikoka, Hawai, United States, December 2019, https://hal.archives-ouvertes.fr/hal-02432579

[31] M. HADDED, K. TOUMI, A. LAOUITI, P. MUHLETHALER.*A trust framework for centralized TDMA scheduling mechanism in Vehicular Ad hoc Networks*, in "WTS 2019: Wireless Telecommunications Symposium", New York, United States, April 2019, https://hal.archives-ouvertes.fr/hal-02454384

[32] R. MILOCCO, P. MUHLETHALER, S. BOUMERDASSI.*Energy-efficient relay selection over fading channels*, in "CCNC 2019 - 16th IEEE Annual Consumer Communications & Networking Conference", Las Vegas, United States, IEEE, January 2019, p. 1-2 [*DOI :* 10.1109/CCNC.2019.8651688], https://hal.archives-ouvertes.fr/hal-02294729

[33] M. L. F. SINDJOUNG, P. MINET.*Wireless Link Quality Prediction in IoT Networks*, in "PEMWN 2019 - 8th IFIP/IEEE International Conference on Performance Evaluation and Modeling inWired andWireless Networks", Paris, France, November 2019, https://hal.archives-ouvertes.fr/hal-02432805

[34] A. SOBEHY, E. RENAULT, P. MUHLETHALER.*CSI based indoor localization using Ensemble Neural Networks*, in "MLN 2019 : 2nd IFIP International Conference on Machine Learning for Networking", Paris, France, December 2019, https://hal.archives-ouvertes.fr/hal-02334588

[35] A. SOBEHY, E. RENAULT, P. MUHLETHALER.*NDR: Noise and Dimensionality Reduction of CSI for indoor positioning using deep learning*, in "GLOBECOM 2019: IEEE Global Communications Conference", Hawaii, United States, IEEE Computer Society, December 2019, https://hal.archives-ouvertes.fr/hal-02314929

Scientific Books (or Scientific Book chapters)

[36] *Machine Learning for Networking: first international conference, MLN 2018, Paris, France, November 27-29, 2018, revised selected papers*, Lecture Notes in Computer Science (LNCS), Springer, Paris, France, 2019, vol. 11407, 388 [*DOI :* 10.1007/978-3-030-19945-6], https://hal.archives-ouvertes.fr/hal-02445117

Other Publications

[37] M. VUČINIĆ, G. SELANDER, J. MATTSSON, D. GARCIA-CARRILLO.*Requirements for a Lightweight AKE for OSCORE*, December 2019, draft-ietf-lake-reqs-00, https://hal.inria.fr/hal-02407621

Project-Team GANG

Networks, Graphs and Algorithms

IN COLLABORATION WITH: Institut de Recherche en Informatique Fondamentale

IN PARTNERSHIP WITH:
CNRS

Université Denis Diderot (Paris 7)

RESEARCH CENTER
Paris

THEME
Networks and Telecommunications

Table of contents

Project-Team GANG

Creation of the Project-Team: 2007 July 01

Keywords:

Computer Science and Digital Science:

A1.2. - Networks
A1.2.3. - Routing
A1.2.9. - Social Networks
A1.3. - Distributed Systems
A3.5. - Social networks
A3.5.1. - Analysis of large graphs
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A7.1. - Algorithms
A7.1.3. - Graph algorithms
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.7. - Graph theory
A8.8. - Network science

Other Research Topics and Application Domains:

B1.1.6. - Evolutionnary biology
B1.1.10. - Systems and synthetic biology
B6.3.2. - Network protocols
B6.3.4. - Social Networks
B7.2. - Smart travel

1. Team, Visitors, External Collaborators

Research Scientists

Laurent Viennot [Team leader, Inria, Senior Researcher, HDR]
Pierre Fraigniaud [CNRS, Senior Researcher, HDR]
Amos Korman [CNRS, Researcher, HDR]
Adrian Kosowski [Inria, Researcher, until Sep 2019, HDR]

Faculty Members

Michel Habib [Paris University, Professor emeritus, HDR]
Yacine Boufkhad [Univ Denis Diderot, Associate Professor]
Pierre Charbit [Univ Denis Diderot, Associate Professor]
Fabien de Montgolfier [Univ Denis Diderot, Associate Professor]
Hugues Fauconnier [Univ Denis Diderot, Associate Professor, HDR]
Carole Gallet Delporte [Univ Denis Diderot, Professor, HDR]

PhD Students

Mengchuan Zou [Inria, PhD Student, until Nov 2019]
Simon Collet [ERC BDA, PhD Student, until Nov 2019]
Brieuc Guinard [ERC BDA, PhD Student]
Zeinab Nehaiï [Université Paris Diderot, PhD Student]

Administrative Assistants

Christine Anocq [Inria, Administrative Assistant]
Nelly Maloisel [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

GANG focuses on algorithm design for large scale networks using structural properties of these networks. Application domains include the development of optimized protocols for large dynamic networks such as mobile networks or overlay networks over Internet. This includes for instance peer-to-peer applications, or the navigability of social networks. GANG tools come from recent advances in the field of graph algorithms, both in centralized and distributed settings. In particular, this includes graph decomposition and geometric properties (such as low doubling dimension, low dimension embedding, etc.). Today, the management of large networks, Internet being the reference, is best effort. However, the demand for mobility (ad hoc networks, wireless connectivity, etc.) and for dynamicity (node churn, fault tolerance, etc.) is increasing. In this distributed setting, it becomes necessary to design a new generation of algorithms and protocols to face the challenge of large scale mobility and dynamicity. In the mean time, recent and sophisticated theoretical results have emerged, offering interesting new tracks for managing large networks. These results concern centralized and decentralized algorithms for solving key problems in communication networks, including routing, but also information retrieval, localization, or load balancing. They are mainly based on structural properties observed in most of real networks: approximate topology with low dimension metric spaces, low treewidth, low doubling dimension, graph minor freeness, etc. In addition, graph decomposition techniques have recently progressed. The scientific community has now tools for optimizing network management. First striking results include designing overlay networks for peer-to-peer systems and understanding the navigability of large social networks.

3. Research Program

3.1. Graph and Combinatorial Algorithms

We focus on two approaches for designing algorithms for large graphs: decomposing the graph and relying on simple graph traversals.

3.1.1. Graph Search

We more deeply study multi-sweep graph searches. In this domain a graph search only yields a total ordering of the vertices which can be used by the subsequent graph searches. This technique can be used on huge graphs and does not need extra memory. We have already obtained preliminary results in this direction and many well-known graph algorithms can be put in this framework. The idea behind this approach is that each sweep discovers some structure of the graph. At the end of the process either we have found the underlying structure (for example an interval representation for an interval graph) or an approximation of it (for example in hard discrete optimization problems). We envision applications to exact computations of centers in huge graphs, to underlying combinatorial optimization problems, but also to networks arising in biology.

3.1.2. Graph Decomposition

In order to summarize a graph into a more compact and more human-readable form, we introduced the hub-laminar decomposition. It is suitable for graphs that are dominated by long isometric cycles or shortest paths, called laminar, which meet only at their extremities, called hubs. Computing this decomposition is NP-hard but a canonical approximation may be computed under some hypotheses on the distances between hubs. It provides a distance labelling for the decomposable graphs. We also investigated the case were the decomposition is reduced to a single cycle, yielding the problem of fining the longest isometric cycle, which is NP-complete and for which a first approximation algorithm was proposed in ENTCS.

3.1.3. Graph Exploration

In the course of graph exploration, a mobile agent is expected to regularly visit all the nodes of an unknown network, trying to discover all its nodes as quickly as possible. Our research focuses on the design and analysis of agent-based algorithms for exploration-type problems, which operate efficiently in a dynamic network environment, and satisfy imposed constraints on local computational resources, performance, and resilience. Our recent contributions in this area concern the design of fast deterministic algorithms for teams of agents operating in parallel in a graph, with limited or no persistent state information available at nodes. We plan further studies to better understand the impact of memory constraints and of the availability of true randomness on efficiency of the graph exploration process.

3.2. Distributed Computing

The distributed computing community can be viewed as a union of two sub-communities. This is also true in our team. Although they have interactions, they are disjoint enough not to leverage each other's results. At a high level, one is mostly interested in timing issues (clock drifts, link delays, crashes, etc.) while the other one is mostly interested in spatial issues (network structure, memory requirements, etc.). Indeed, one sub-community is mostly focusing on the combined impact of asynchronism and faults on distributed computation, while the other addresses the impact of network structural properties on distributed computation. Both communities address various forms of computational complexity, through the analysis of different concepts. This includes, e.g., failure detectors and wait-free hierarchy for the former community and compact labeling schemes, and computing with advice for the latter community. We have an ambitious project to achieve the reconciliation between the two communities by focusing on the same class of problems, the yes/no-problems, and establishing the scientific foundations for building up a consistent theory of computability and complexity for distributed computing. The main question addressed is therefore: is the absence of globally coherent computational complexity theories covering more than fragments of distributed computing, inherent to the field? One issue is obviously the types of problems located at the core of distributed computing. Tasks like consensus, leader election, and broadcasting are of very different nature. They are not *yes-no* problems, neither are they minimization problems. Coloring and Minimal Spanning Tree are optimization problems but we are often more interested in constructing an optimal solution than in verifying the correctness of a given solution. Still, it makes full sense to analyze the *yes-no* problems corresponding to checking the validity of the output of tasks. Another issue is the power of individual computation. The FLP impossibility result as well as Linial's lower bound hold independently of the individual computational power of the involved computing entities. For instance, the individual power of solving NP-hard problems in constant time would not help overcoming these limits, which are inherent to the fact that computation is distributed. A third issue is the abundance of models for distributed computing frameworks, from shared memory to message passing, spanning all kinds of specific network structures (complete graphs, unit-disk graphs, etc.) and/or timing constraints (from complete synchronism to full asynchronism). There are however models, typically the wait-free model and the LOCAL model, which, though they do not claim to reflect accurately real distributed computing systems, enable focusing on some core issues. Our ongoing research program is to carry many important notions of Distributed Computing into a *standard* computational complexity.

3.3. Network Algorithms and Analysis

Based on our scientific expertise in both graph algorithms and distributed algorithms, we plan to analyze the behavior of various networks such as future Internet, social networks, overlay networks resulting from distributed applications or online social networks.

3.3.1. Information Dissemination

One of the key aspects of networks resides in the dissemination of information among the nodes. We aim at analyzing various procedures of information propagation from dedicated algorithms to simple distributed schemes such as flooding. We also consider various models, e.g. where noise can alter information as it propagates or where memory of nodes is limited.

3.3.2. Routing Paradigms

We try to explore new routing paradigms such as greedy routing in social networks for example. We are also interested in content centric networking where routing is based on content name rather than content address. One of our targets is multiple path routing: how to design forwarding tables providing multiple disjoint paths to the destination?

3.3.3. Beyond Peer-to-Peer

Based on our past experience of peer-to-peer application design, we would like to broaden the spectrum of distributed applications where new efficient algorithms can be designed and their analysis can be performed. We especially target online social networks as we see them as collaborative tools for exchanging information. A basic question resides in making the right connections for gathering filtered and accurate information with sufficient coverage.

3.3.4. SAT and Forwarding Information Verification

As forwarding tables of networks grow and are sometimes manually modified, the problem of verifying them becomes critical and has recently gained interest. Some problems that arise in network verification such as loop detection for example, may be naturally encoded as Boolean Satisfiability problems. Beside theoretical interest in complexity proofs, this encoding allows one to solve these problems by taking advantage of efficient Satisfiability testing solvers. Indeed, SAT solvers have proved to be very efficient in solving problems coming from various areas (Circuit Verification, Dependency and Conflicts in Software distributions...) and encoded in Conjunctive Normal Form. To test an approach using SAT solvers in network verification, one needs to collect data sets from a real network and to develop good models for generating realistic networks. The technique of encoding and the solvers themselves need to be adapted to this kind of problems. All this represents a rich experimental field of future research.

3.3.5. Network Analysis

Finally, we are interested in analyzing the structural properties of practical networks. This can include diameter computation or ranking of nodes. As we mostly consider large networks, we are often interested in efficient heuristics. Ideally, we target heuristics that give exact answers and are reasonably fast in practice although short computation time is not guaranteed for all networks. We have already designed such heuristics for diameter computation; understanding the structural properties that enable short computation time in practice is still an open question.

3.3.6. Network Parameter

Betweenness centrality is a graph parameter that has been successfully applied to network analysis. In the context of computer networks, it was considered for various objectives, ranging from routing to service placement. However, as observed by Maccari et al. [INFOCOM 2018], research on betweenness centrality for improving protocols was hampered by the lack of a usable, fully distributed algorithm for computing this parameter. In [21], we resolved this issue by designing an efficient algorithm for computing betweenness centrality, which can be implemented by minimal modifications to any distance-vector routing protocol based on Bellman-Ford. The convergence time of our implementation is shown to be proportional to the diameter of the network.

4. Application Domains

4.1. Large scale networks

Application domains include evaluating Internet performances, the design of new peer-to-peer applications, enabling large scale networks, and developping tools for transportation networks.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

Amos Korman won the 2020 prize of innovations in distributed computing.

Pierre Fraigniaud and his co-authors obtained the best paper award at SSS 2019:
BEST PAPERS AWARDS :
[19]
A. CASTAÑEDA, P. FRAIGNIAUD, A. PAZ, S. RAJSBAUM, M. ROY, C. TRAVERS.*Synchronous t-Resilient Consensus in Arbitrary Graphs*, in "SSS 2019 - 21st International Symposium on Stabilization, Safety, and Security of Distributed Systems", Pisa, Italy, October 2019 [*DOI :* 10.1007/978-3-030-34992-9_5], https://hal.inria.fr/hal-02433524

6. New Software and Platforms

6.1. big-graph-tools

KEYWORD: Graph algorithmics

FUNCTIONAL DESCRIPTION: Gang is developping a software for big graph manipulation. A preliminary library offering diameter and skeleton computation. This library was used to compute the diameters of the worldwide road network (200M edges) and the largest strongly connected component of the Twitter follower-followee graph (23G edges).

- Contact: Laurent Viennot
- URL: https://who.rocq.inria.fr/Laurent.Viennot/dev/big-graph-tools/

6.2. GRPH

The high performance graph library for Java

KEYWORDS: Graph - Graph algorithmics - Java

FUNCTIONAL DESCRIPTION: Grph is an open-source Java library for the manipulation of graphs. Its design objectives are to make it portable, simple to use/extend, computationally/memory efficient, and, according to its initial motivation: useful in the context of graph experimentation and network simulation. Grph also has the particularity to come with tools like an evolutionary computation engine, a bridge to linear programming solvers, a framework for distributed computing, etc.

Grph offers a very general model of graphs. Unlike other graph libraries which impose the user to first decide if he wants to deal with directed, undirected, hyper (or not) graphs, the model offered by Grph is unified in a general class that supports mixed graphs made of undirected and directed simple and hyper edges. Grph achieves great efficiency through the use of multiple code optimization techniques such as multi-core parallelism, caching, adequate data structures, use of primitive objects, exploitation of low-level processor caches, on-the-fly compilation of specific C/C++ code, etc. Grph attempts to access the Internet in order to check if a new version is available and to report who is using it (login name and hostname). This has no impact whatsoever on performance and security.

- Participants: Aurélien Lancin, David Coudert, Issam Tahiri, Luc Hogie and Nathann Cohen
- Contact: Luc Hogie
- URL: http://www.i3s.unice.fr/~hogie/grph/

7. New Results

7.1. Graph and Combinatorial Algorithms

7.1.1. Fast Diameter Computation within Split Graphs

When can we compute the diameter of a graph in quasi linear time? In [22], we address this question for the class of *split graphs*, that we observe to be the hardest instances for deciding whether the diameter is at most two. We stress that although the diameter of a non-complete split graph can only be either 2 or 3, under the Strong Exponential-Time Hypothesis (SETH) we cannot compute the diameter of a split graph in less than quadratic time. Therefore it is worth to study the complexity of diameter computation on *subclasses* of split graphs, in order to better understand the complexity border. Specifically, we consider the split graphs with bounded *clique-interval number* and their complements, with the former being a natural variation of the concept of interval number for split graphs that we introduce in this paper. We first discuss the relations between the clique-interval number and other graph invariants such as the classic interval number of graphs, the treewidth, the *VC-dimension* and the *stabbing number* of a related hypergraph. Then, in part based on these above relations, we almost completely settle the complexity of diameter computation on these subclasses of split graphs:

- For the k-clique-interval split graphs, we can compute their diameter in truly subquadratic time if $k = \mathcal{O}(1)$, and even in quasi linear time if $k = o(\log n)$ and in addition a corresponding ordering is given. However, under SETH this cannot be done in truly subquadratic time for any $k = \omega(\log n)$.
- For the *complements* of k-clique-interval split graphs, we can compute their diameter in truly subquadratic time if $k = \mathcal{O}(1)$, and even in time $\mathcal{O}(km)$ if a corresponding ordering is given. Again this latter result is optimal under SETH up to polylogarithmic factors.

Our findings raise the question whether a k-clique interval ordering can always be computed in quasi linear time. We prove that it is the case for $k = 1$ and for some subclasses such as bounded-treewidth split graphs, threshold graphs and comparability split graphs. Finally, we prove that some important subclasses of split graphs – including the ones mentioned above – have a bounded clique-interval number.

7.1.2. Diameter computation on H-minor free graphs and graphs of bounded (distance) VC-dimension

Under the Strong Exponential-Time Hypothesis, the diameter of general unweighted graphs cannot be computed in truly subquadratic time. Nevertheless there are several graph classes for which this can be done such as bounded-treewidth graphs, interval graphs and planar graphs, to name a few. We propose to study unweighted graphs of constant *distance VC-dimension* as a broad generalization of many such classes – where the distance VC-dimension of a graph G is defined as the VC-dimension of its ball hypergraph: whose hyperedges are the balls of all possible radii and centers in G. In particular for any fixed H, the class of H-minor free graphs has distance VC-dimension at most $|V(H)| - 1$. In [23], we show the following.

- Our first main result is a Monte Carlo algorithm that on graphs of distance VC-dimension at most d, for any fixed k, either computes the diameter or concludes that it is larger than k in time $\tilde{\mathcal{O}}(k \cdot mn^{1-\varepsilon_d})$, where $\varepsilon_d \in (0; 1)$ only depends on d. We thus obtain a *truly subquadratic-time parameterized* algorithm for computing the diameter on such graphs.
- Then as a byproduct of our approach, we get the first truly subquadratic-time randomized algorithm for *constant* diameter computation on all the *nowhere dense* graph classes. The latter classes include all proper minor-closed graph classes, bounded-degree graphs and graphs of bounded expansion.
- Finally, we show how to remove the dependency on k for *any* graph class that excludes a fixed graph H as a minor. More generally, our techniques apply to any graph with constant distance VC-dimension and *polynomial expansion* (or equivalently having strongly sublinear balanced separators). As a result for all such graphs one obtains a truly subquadratic-time randomized algorithm for computing their diameter.

We note that all our results also hold for *radius* computation. Our approach is based on the work of Chazelle and Welzl who proved the existence of spanning paths with strongly sublinear *stabbing number* for every hypergraph of constant VC-dimension. We show how to compute such paths efficiently by combining known algorithms for the stabbing number problem with a clever use of ε-nets, region decomposition and other partition techniques.

7.1.3. Approximation of eccentricites and distance using δ-hyperbolicity

In [9], we show that the eccentricities of all vertices of a δ-hyperbolic graph $G = (V, E)$ can be computed in linear time with an additive one-sided error of at most $c \cdot \delta$, i.e., after a linear time preprocessing, for every vertex v of G one can compute in $O(1)$ time an estimate $\overline{ecc_G(v)}$ of its eccentricity $ecc_G(v) := max\{d_G(u, v) : u \in V\}$ such that $ecc_G(v) \leq \overline{ecc_G(v)} \leq ecc_G(v) + c \cdot \delta$ for a small constant c. We prove that every δ-hyperbolic graph G has a shortest path tree T, constructible in linear time, such that for every vertex v of G, $ecc_G(v) \leq ecc_T(v) \leq ecc_G(v) + c \cdot \delta$, where $ecc_T(v) := max\{d_T(u, v) : u \in V\}$. These results are based on an interesting monotonicity property of the eccentricity function of hyperbolic graphs: the closer a vertex is to the center of G, the smaller its eccentricity is. We also show that the distance matrix of G with an additive one-sided error of at most $c' \cdot \delta$ can be computed in $O(|V|^2 log^2 |V|)$ time, where $c' < c$ is a small constant. Recent empirical studies show that many real world graphs (including Internet application networks, web networks, collaboration networks, social networks, biological networks, and others) have small hyperbolicity. So, we analyze the performance of our algorithms for approximating eccentricities and distance matrix on a number of real-world networks. Our experimental results show that the obtained estimates are even better than the theoretical bounds.

7.1.4. Graph and Hypergraph Decompositions

In [26], we study modular decomposition of hypergraphs and propose some polynomial algorithms to this aim. We also study several notions of approximation of modular decomposition of graphs, by relaxing the definition of modules introducing a tolerance (ϵ edges can miss) this will be presented at CALDAM 2020, Hyderabad. Both topics can be seen as the search for new models of regularity in discrete structures, as in particular bipartite graphs. In both references our polynomial algorithms have to be improved before being applied on real-world data.

7.2. Distributed Computing

7.2.1. Distributed Interactive Proofs

In a distributed locally-checkable proof, we are interested in checking the legality of a given network configuration with respect to some Boolean predicate. To do so, the network enlists the help of a *prover* — a computationally-unbounded oracle that aims at convincing the network that its state is legal, by providing the nodes with certificates that form a distributed proof of legality. The nodes then verify the proof by examining their certificate, their local neighborhood and the certificates of their neighbors.

In [24], we examine the power of a *randomized* form of locally-checkable proof, called *distributed Merlin-Arthur protocols*, or dMA for short. In a dMA protocol, the prover assigns each node a short certificate, and the nodes then exchange *random messages* with their neighbors. We show that while there exist problems for which dMA protocols are more efficient than protocols that do not use randomness, for several natural problems, including Leader Election, Diameter, Symmetry, and Counting Distinct Elements, dMA protocols are no more efficient than standard nondeterministic protocols. This is in contrast with Arthur-Merlin (dAM) protocols and Randomized Proof Labeling Schemes (RPLS), which are known to provide improvements in certificate size, at least for some of the aforementioned properties.

The study of interactive proofs in the context of distributed network computing is a novel topic, recently introduced by Kol, Oshman, and Saxena [PODC 2018]. In the spirit of sequential interactive proofs theory, we study in [20] the power of distributed interactive proofs. This is achieved via a series of results establishing trade-offs between various parameters impacting the power of interactive proofs, including the number of interactions, the certificate size, the communication complexity, and the form of randomness used. Our results also connect distributed interactive proofs with the established field of distributed verification. In general, our results contribute to providing structure to the landscape of distributed interactive proofs.

7.2.2. Topological Approach of Network Computing

More than two decades ago, combinatorial topology was shown to be useful for analyzing distributed fault-tolerant algorithms in shared memory systems and in message passing systems. In [18], we show that combinatorial topology can also be useful for analyzing distributed algorithms in networks of arbitrary structure. To illustrate this, we analyze consensus, set-agreement, and approximate agreement in networks, and derive lower bounds for these problems under classical computational settings, such as the LOCAL model and dynamic networks.

In [19], we study the number of rounds needed to solve consensus in a synchronous network G where at most t nodes may fail by crashing. This problem has been thoroughly studied when G is a complete graph, but very little is known when G is arbitrary. We define a notion of radius that considers all ways in which t nodes may crash, and present an algorithm that solves consensus in radius rounds. Then we derive a lower bound showing that our algorithm is optimal for vertex-transitive graphs, among oblivious algorithms.

7.2.3. Making Local Algorithms Wait-Free

When considering distributed computing, reliable message-passing synchronous systems on the one side, and asynchronous failure-prone shared-memory systems on the other side, remain two quite independently studied ends of the reliability/asynchrony spectrum. The concept of locality of a computation is central to the first one, while the concept of wait-freedom is central to the second one. In [8], we propose a new DECOUPLED model in an attempt to reconcile these two worlds. It consists of a synchronous and reliable communication graph of nodes, and on top a set of asynchronous crash-prone processes, each attached to a communication node. To illustrate the DECOUPLED model, the paper presents an asynchronous 3-coloring algorithm for the processes of a ring. From the processes point of view, the algorithm is wait-free. From a locality point of view, each process uses information only from processes at distance $O(log^{*}n)$ from it. This local wait-free algorithm is based on an extension of the classical Cole and Vishkin's vertex coloring algorithm in which the processes are not required to start simultaneously.

In [31], we show that, for any task T associated to a locally checkable labeling (lcl), if T is solvable in t rounds by a deterministic algorithm in the local model, then T remains solvable by a deterministic algorithm in $O(t)$ rounds in an asynchronous variant of the local model whenever $t = O(polylogn)$.

7.2.4. Towards Synthesis of Distributed Algorithms with SMT Solvers

In [32], we consider the problem of synthesizing distributed algorithms working on a specific execution context. We show it is possible to use the linear time temporal logic in order to both specify the correctness of algorithms and their execution contexts. We then provide a method allowing to reduce the synthesis problem of finite state algorithms to some model-checking problems. We finally apply our technique to automatically generate algorithms for consensus and epsilon-agreement in the case of two processes using the SMT solver Z3.

7.2.5. On Weakest Failure Detector

Failure detectors are devices (objects) that provide the processes with information on failures. They were introduced to enrich asynchronous systems so that it becomes possible to solve problems (or implement concurrent objects) that are otherwise impossible to solve in pure asynchronous systems where processes are prone to crash failures. The most famous failure detector (which is called "eventual leader" and denoted Ω) is the weakest failure detector which allows consensus to be solved in n-process asynchronous systems where up to $t = n - 1$ processes may crash in the read/write communication model, and up to $t < n/2$ processes may crash in the message-passing communication model.

When looking at the mutual exclusion problem (or equivalently the construction of a lock object), while the weakest failure detectors are known for both asynchronous message-passing systems and read/write systems in which up to $t < n$ processes may crash, for the starvation-freedom progress condition, it is not yet known for weaker deadlock-freedom progress condition in read/write systems. In [34], we extend the previous results, namely, it presents the weakest failure detector that allows mutual exclusion to be solved in asynchronous n-process read/write systems where any number of processes may crash, whatever the progress condition (deadlock-freedom or starvation-freedom).

In read/read/write communication model, and in the message-passing communication model, all correct processes are supposed to participate in a consensus instance and in particular the eventual leader.

In [33], we considers the case where some subset of processes that do not crash (not predefined in advance) are allowed not to participate in a consensus instance. In this context Ω cannot be used to solve consensus as it could elect as eventual leader a non-participating process. This paper presents the weakest failure detector that allows correct processes not to participate in a consensus instance.This failure detector, denoted Ω^*, is a variant of Ω. The paper presents also an Ω^*-based consensus algorithm for the asynchronous read/write model, in which any number of processes may crash, and not all the correct processes are required to participate.

7.2.6. Multi-Round Cooperative Search Games with Multiple Players

We study search in the context of competing agents. The setting we consider combines game-theoretic concepts with notions related to parallel computing. Assume that a treasure is placed in one of M boxes according to a known distribution and that k searchers are searching for it in parallel during T rounds. In [27], we study the question of how to incentivize selfish players so that group performance would be maximized. Here, this is measured by the *success probability*, namely, the probability that at least one player finds the treasure. We focus on *congestion policies*$C(\ell)$ that specify the reward that a player receives if it is one of ℓ players that (simultaneously) find the treasure for the first time. Our main technical contribution is proving that the *exclusive policy*, in which $C(1) = 1$ and $C(\ell) = 0$ for $\ell > 1$, yields a *price of anarchy* of $\left(1 - (1 - 1/k)^k\right)^{-1}$, and that this is the best possible price among all symmetric reward mechanisms. For this policy we also have an explicit description of a symmetric equilibrium, which is in some sense unique, and moreover enjoys the best success probability among all symmetric profiles. For general congestion policies, we show how to polynomially find, for any $\theta > 0$, a symmetric multiplicative $(1 + \theta)(1 + C(k))$-equilibrium.

Together with an appropriate reward policy, a central entity can suggest players to play a particular profile at equilibrium. As our main conceptual contribution, we advocate the use of symmetric equilibria for such purposes. Besides being fair, we argue that symmetric equilibria can also become highly robust to crashes of players. Indeed, in many cases, despite the fact that some small fraction of players crash (or refuse to participate), symmetric equilibria remain efficient in terms of their group performances and, at the same time, serve as approximate equilibria. We show that this principle holds for a class of games, which we call *monotonously scalable* games. This applies in particular to our search game, assuming the natural *sharing policy*, in which $C(\ell) = 1/\ell$. For the exclusive policy, this general result does not hold, but we show that the symmetric equilibrium is nevertheless robust under mild assumptions.

7.3. Models and Algorithms for Networks

7.3.1. Exploiting Hopsets: Improved Distance Oracles for Graphs of Constant Highway Dimension and Beyond

For fixed $h \geq 2$, we consider in [25] the task of adding to a graph G a set of weighted shortcut edges on the same vertex set, such that the length of a shortest h-hop path between any pair of vertices in the augmented graph is exactly the same as the original distance between these vertices in G. A set of shortcut edges with this property is called an *exact h-hopset* and may be applied in processing distance queries on graph G. In particular, a 2-hopset directly corresponds to a distributed distance oracle known as a *hub labeling*. In this work, we explore centralized distance oracles based on 3-hopsets and display their advantages in several practical scenarios. In particular, for graphs of constant highway dimension, and more generally for graphs of constant skeleton dimension, we show that 3-hopsets require *exponentially* fewer shortcuts per node than any previously described distance oracle, and also offer a speedup in query time when compared to simple oracles based on a direct application of 2-hopsets. Finally, we consider the problem of computing minimum-size h-hopset (for any $h \geq 2$) for a given graph G, showing a polylogarithmic-factor approximation for the case of unique shortest path graphs. When $h = 3$, for a given bound on the space used by the distance oracle, we provide a construction of hopset achieving polylog approximation both for space and query time compared to the optimal 3-hopset oracle given the space bound.

7.3.2. Hardness of exact distance queries in sparse graphs through hub labeling

A *distance labeling scheme* is an assignment of bit-labels to the vertices of an undirected, unweighted graph such that the distance between any pair of vertices can be decoded solely from their labels. An important class of distance labeling schemes is that of *hub labelings*, where a node $v \in G$ stores its distance to the so-called hubs $S_v \subseteq V$, chosen so that for any $u, v \in V$ there is $w \in S_u \cap S_v$ belonging to some shortest uv path. Notice that for most existing graph classes, the best distance labelling constructions existing use at some point a hub labeling scheme at least as a key building block.

In [28], our interest lies in hub labelings of sparse graphs, i.e., those with $|E(G)| = O(n)$, for which we show a lowerbound of $\frac{n}{2^{O(\sqrt{\log n})}}$ for the average size of the hubsets. Additionally, we show a hub-labeling construction for sparse graphs of average size $O(\frac{n}{RS(n)^c})$ for some $0 < c < 1$, where $RS(n)$ is the so-called Ruzsa-Szemerédi function, linked to structure of induced matchings in dense graphs. This implies that further improving the lower bound on hub labeling size to $\frac{n}{2^{(\log n)^{o(1)}}}$ would require a breakthrough in the study of lower bounds on $RS(n)$, which have resisted substantial improvement in the last 70 years.

For general distance labeling of sparse graphs, we show a lowerbound of $\frac{1}{2^{\Theta(\sqrt{\log n})}} SumIndex(n)$, where $SumIndex(n)$ is the communication complexity of the Sum-Index problem over Z_n. Our results suggest that the best achievable hub-label size and distance-label size in sparse graphs may be $\Theta(\frac{n}{2^{(\log n)^c}})$ for some $0 < c < 1$.

7.3.3. Fast Public Transit Routing with Unrestricted Walking through Hub Labeling

In [30], we propose a novel technique for answering routing queries in public transportation networks that allows unrestricted walking. We consider several types of queries: earliest arrival time, Pareto-optimal journeys regarding arrival time, number of transfers and walking time, and profile, i.e. finding all Pareto-optimal journeys regarding travel time and arrival time in a given time interval. Our techniques uses hub labeling to represent unlimited foot transfers and can be adapted to both classical algorithms RAPTOR and CSA. We obtain significant speedup compared to the state-of-the-art approach based on contraction hierarchies. A research report version is deposited on HAL with number hal-02161283.

7.3.4. Independent Lazy Better-Response Dynamics on Network Games

In [29], we study an *independent* best-response dynamics on network games in which the nodes (players) decide to revise their strategies independently with some probability. We are interested in the *convergence time* to the equilibrium as a function of this probability, the degree of the network, and the potential of the underlying games.

7.3.5. A Comparative Study of Neural Network Compression

There has recently been an increasing desire to evaluate neural networks locally on computationally-limited devices in order to exploit their recent effectiveness for several applications; such effectiveness has nevertheless come together with a considerable increase in the size of modern neural networks, which constitute a major downside in several of the aforementioned computationally-limited settings. There has thus been a demand of compression techniques for neural networks. Several proposal in this direction have been made, which famously include hashing-based methods and pruning-based ones. However, the evaluation of the efficacy of these techniques has so far been heterogeneous, with no clear evidence in favor of any of them over the others. In [36], we address this latter issue by providing a comparative study. While most previous studies test the capability of a technique in reducing the number of parameters of state-of-the-art networks , we follow [CWT

+ 15] in evaluating their performance on basic architectures on the MNIST dataset and variants of it, which allows for a clearer analysis of some aspects of their behavior. To the best of our knowledge, we are the first to directly compare famous approaches such as HashedNet, Optimal Brain Damage (OBD), and magnitude-based pruning with L1 and L2 regularization among them and against equivalent-size feed-forward neural networks with simple (fully-connected) and structural (convolutional) neural networks. Rather surprisingly, our experiments show that (iterative) pruning-based methods are substantially better than the HashedNet architecture, whose compression doesn't appear advantageous to a carefully chosen convolutional network. We also show that, when the compression level is high, the famous OBD pruning heuristics deteriorates to the point of being less efficient than simple magnitude-based techniques.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR DESCARTES

Participants: Carole Gallet Delporte, Hugues Fauconnier, Pierre Fraigniaud, Adrian Kosowski, Laurent Viennot.

Cyril Gavoille (U. Bordeaux) leads this project that grants 1 Post-Doc. H. Fauconnier is the local coordinator (This project began in October 2016).

Despite the practical interests of reusable frameworks for implementing specific distributed services, many of these frameworks still lack solid theoretical bases, and only provide partial solutions for a narrow range of services. We argue that this is mainly due to the lack of a generic framework that is able to unify the large body of fundamental knowledge on distributed computation that has been acquired over the last 40 years. The DESCARTES project aims at bridging this gap, by developing a systematic model of distributed computation that organizes the functionalities of a distributed computing system into reusable modular constructs assembled via well-defined mechanisms that maintain sound theoretical guarantees on the resulting system. DESCARTES arises from the strong belief that distributed computing is now mature enough to resolve the tension between the social needs for distributed computing systems, and the lack of a fundamentally sound and systematic way to realize these systems.

8.1.2. ANR MultiMod

Participants: Adrian Kosowski, Laurent Viennot.

David Coudert (Sophia Antipolis) leads this project. L. Viennot coordinates locally. The project began in 2018.

The MultiMod project aims at enhancing the mobility of citizens in urban areas by providing them, through a unique interface enabling to express their preferences, the most convenient transportation means to reach their destinations. Indeed, the increasing involvement of actors and authorities in the deployment of more responsible and cost-effective logistics and the progress made in the field of digital technology have made possible to create synergies in the creation of innovative services for improving the mobility in cities. However, users are faced with a number of solutions that coexist at different scales, providing complementary information for the mobility of users, but that make very complex to find the most convenient itinerary at a given time for a specific user. In this context, MultiMod aims at improving the mobility of citizens in urban areas by proposing contextualized services, linking users, to facilitate multimodal transport by combining, with flexibility, all available modes (planned/dynamic carpooling, public transport (PT), car-sharing, bicycle, etc.).

We consider the use of carpooling in metropolitan areas, and so for short journeys. Such usage enables itineraries that are not possible with PT, allows for opening up areas with low PT coverage by bringing users near PT (last miles), and for faster travel-time when existing PT itineraries are too complex or with too low frequency (e.g., one bus per hour). In this context, the application must help the driver and the passenger as much as possible. In particular, the application must propose the meeting-point, indicate the driver the detour duration, and indicate the passenger how to reach this meeting-point using PT. Here, the time taken by drivers and passengers to agree becomes a critical issue and so the application must provide all needed information to quickly take a decision (i.e., in one click).

In addition, the era of Smart City gathers many emerging concepts, driven by innovative technological players, which enables the exploitation of real-time data (e.g., delay of a bus, traffic jam) made available by the various actors (e.g., communities in the framework of Open Data projects, users via their mobile terminals, traffic supervision authorities). In the MultiMod project, we will use these rich sources of data to propose itineraries that are feasible at query-time. Our findings will enable the design of a mobility companion able not only to guide the user along her journey, including when and how to change of transportation mean, but also to propose itinerary changes when the current one exceeds a threshold delay. The main originality of this project is thus to address the problem of computing itineraries in large-scale networks combining PT, carpooling and real-time data, and to satisfy the preferences of users. We envision that the outcome of this project will significantly improve the daily life of citizens.

The targeted metropolitan area for validating our solutions is Ile-de-France. Indeed, Instant-System is currently developing the new application “Vianavigo lab” which will replace the current “Vianavigo” application for the PT network of Ile-de-France. Our findings will therefore be tested at scale and eventually be integrated and deployed in production servers and mobile applications. The smaller networks of Bordeaux and Nice will be used to perform preliminary evaluations since Instant System already operates applications in these cities (Boogi Nice, Boogi Bordeaux). An important remark is that new features and algorithms can contractually be deployed in production every 4 months, thus enabling Instant System to measure and challenge the results of the MultiMod project in continue. This is a chance for the project to maximize its impact.

8.1.3. ANR FREDDA

Participants: Carole Gallet Delporte, Hugues Fauconnier, Pierre Fraigniaud.

Arnaud Sangnier (IRIF, Univ Paris Diderot) leads this project that grants 1 PhD. (This project began in October 2017).

Distributed algorithms are nowadays omnipresent in most systems and applications. It is of utmost importance to develop algorithmic solutions that are both robust and flexible, to be used in large scale applications. Currently, distributed algorithms are developed under precise assumptions on their execution context: synchronicity, bounds on the number of failures, etc. The robustness of distributed algorithms is a challenging problem that has not been much considered until now, and there is no systematic way to guarantee or verify the behavior of an algorithm beyond the context for which it has been designed. We propose to develop automated formal method techniques to verify the robustness of distributed algorithms and to support the development of robust applications. Our methods are of two kinds: statically through classical verification, and dynamically, by synthesizing distributed monitors, that check either correctness or the validity of the context hypotheses at runtime.

8.1.4. ANR Distancia

Participants: Pierre Charbit, Michel Habib, Laurent Viennot.

Victor Chepoi (Univ. Marseille) leads this project. P. Charbit coordinates locally. The project began in early-2018.

The theme of the project is Metric Graph Theory, and we are concerned both on theoretical foundations and applications. Such applications can be found in real world networks. For example, the hub labelling problem in road networks can be directly applied to car navigation applications. Understanding key structural properties of large-scale data networks is crucial for analyzing and optimizing their performance, as well as improving their reliability and security. In prior empirical and theoretical studies researchers have mainly focused on features such as small world phenomenon, power law degree distribution, navigability, and high clustering coefficients. Although those features are interesting and important, the impact of intrinsic geometric and topological features of large-scale data networks on performance, reliability and security is of much greater importance. Recently, there has been a surge of empirical works measuring and analyzing geometric characteristics of real-world networks, namely the Gromov hyperbolicity (called also the negative curvature) of the network. It has been shown that a number of data networks, including Internet application networks, web networks, collaboration networks, social networks, and others, have small hyperbolicity.

Metric graph theory was also indispensable in solving some open questions in concurrency and learning theory in computer science and geometric group theory in mathematics. Median graphs are exactly the 1–skeletons of CAT(0) cube complexes (which have been characterized by Gromov in a local-to-global combinatorial way). They play a vital role in geometric group theory (for example, in the recent solution of the famous Virtual Haken Conjecture). Median graphs are also the domains of event structures of Winskel, one of the basic abstract models of concurrency. This correspondence is very useful in dealing with questions on event structures.

Many classical algorithmic problems concern distances: shortest path, center and diameter, Voronoi diagrams, TSP, clustering, etc. Algorithmic and combinatorial problems related to distances also occur in data analysis. Low-distortion embeddings into l1-spaces (theorem of Bourgain and its algorithmical use by Linial et al.) were the founding tools in metric methods. Recently, several approximation algorithms for NP-hard problems were designed using metric methods. Other important algorithmic graph problems related to distances concern the construction of sparse subgraphs approximating inter-node distances and the converse, augmentation problems with distance constraints. Finally, in the distributed setting, an important problem is that of designing compact data structures allowing very fast computation of inter- node distances or routing along shortest or almost shortest paths. Besides computer science and mathematics, applications of structures involving distances can be found in archeology, computational biology, statistics, data analysis, etc. The problem of characterizing isometric subgraphs of hypercubes has its origin in communication theory and linguistics. . To take into account the recombination effect in genetic data, the mathematicians Bandelt and Dress developed in 1991 the theory of canonical decompositions of finite metric spaces. Together with geneticists, Bandelt successfully used it over the years to reconstruct phylogenies, in the evolutional analysis of mtDNA data in human genetics. One important step in their method is to build a reduced median network that spans the data but still contains all most parsimonious trees. As mentioned above, the median graphs occurring there constitute a central notion in metric graph theory.

With this project, we aim to participate at the elaboration of this new domain of Metric Graph Theory, which requires experts and knowledge in combinatorics (graphs, matroids), geometry, and algorithms. This expertise is distributed over the members of the consortium and a part of the success of our project it will be to share these knowledges among all the members of the consortium. This way we will create a strong group in France on graphs and metrics.

8.1.5. ANR HOSIGRA

Participants: Pierre Charbit, Michel Habib.

This project starting in early-2018, led by Reza Naserasr, explores the connection between minors and colorings, exploiting the notion of signed graphs. With the four colour theorem playing a central role in development of Graph Theory, the notions of minor and coloring have been branded as two of the most distinguished concepts in this field. The geometric notion of planarity has given birth to the theory of minors among others, and coloring have proven to have an algebraic nature through its extension to the theory of graph homomorphisms. Great many projects have been completed on both subjects, but what remains mostly a mystery is the correlation of the two subjects. The four color theorem itself, in slightly stronger form, claims that if a complete graph on five vertices cannot be formed by minor operation from a given graph, then the graph can be homomorphically mapped into the complete graph on four vertices (thus a 4-coloring). Commonly regarded as the most challenging conjecture on graph theory, the Hadwiger conjecture claims that five and four in this theorem can be replaced with n and $n-1$ respectively for any value of n. The correlation of these two concepts has been difficult to study, mainly for the following reason: While the coloring or homomorphism problems roots back into intersections of odd-cycles, the minor operation is irrelevant of the parity of cycles. To overcome this barrier, the notion of signed graphs has been used implicitly since 1970s when coloring results on graphs with no odd-K4 is proved, following which a stronger form of the Hadwiger conjecture, known as Odd Hadwiger conjecture, was proposed by P. Seymour and B. Gerards, independently. Being a natural subclass of Matroids and a superclass of graphs, the notion of minor of signed graphs is well studied and many results from graph minor are either already extended to signed graphs or it is considered by experts of the subject. Observing the importance, and guided by some earlier works, in particular that of B. Guenin, we then started the study of algebraic concepts (coloring and homomormphisms) for signed graphs. Several results have been obtained in the past decade, and this project aims at exploring more of this topic.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

Amos Korman has an ERC Consolidator Grant entitled “Distributed Biological Algorithms (DBA)”, started in May 2015. This project proposes a new application for computational reasoning. More specifically, the purpose of this interdisciplinary project is to demonstrate the usefulness of an algorithmic perspective in studies of complex biological systems. We focus on the domain of collective behavior, and demonstrate the benefits of using techniques from the field of theoretical distributed computing in order to establish algorithmic insights regarding the behavior of biological ensembles. The project includes three related tasks, for which we have already obtained promising preliminary results. Each task contains a purely theoretical algorithmic component as well as one which integrates theoretical algorithmic studies with experiments. Most experiments are strategically designed by the PI based on computational insights, and are physically conducted by experimental biologists that have been carefully chosen by the PI. In turn, experimental outcomes will be theoretically analyzed via an algorithmic perspective. By this integration, we aim at deciphering how a biological individual (such as an ant) “thinks”, without having direct access to the neurological process within its brain, and how such limited individuals assemble into ensembles that appear to be far greater than the sum of their parts. The ultimate vision behind this project is to enable the formation of a new scientific field, called algorithmic biology, that bases biological studies on theoretical algorithmic insights.

8.2.2. LIA Struco

Pierre Charbit is director of the LIA STRUCO, which is an Associated International Laboratory of CNRS between IÚUK, Prague, and IRIF, Paris. The director on the Czech side is Pr. Jaroslav Nešetřil. The primary theme of the laboratory is graph theory, more specifically: sparsity of graphs (nowhere dense classes of graphs, bounded expansion classes of graphs), extremal graph theory, graph coloring, Ramsey theory, universality and morphism duality, graph and matroid algorithms and model checking.

STRUCO focuses on high-level study of fundamental combinatorial objects, with a particular emphasis on comprehending and disseminating the state-of-the-art theories and techniques developed. The obtained insights shall be applied to obtain new results on existing problems as well as to identify directions and questions for future work.

One of the main goals of STRUCO is to provide a sustainable and reliable structure to help Czech and French researchers cooperate on long-term projects, disseminate the results to students of both countries and create links between these students more systematically. The chosen themes of the project indeed cover timely and difficult questions, for which a stable and significant cooperation structure is needed. By gathering an important number of excellent researchers and students, the LEA will create the required environment for making advances, which shall be achieved not only by short-term exchanges of researchers, but also by a strong involvement of Ph. D students in the learning of state-of-the-art techniques and in the international collaborations.

STRUCO is a natural place to federate and organize these many isolated collaborations between our two countries. Thus, the project would ensure long-term cooperations and allow young researchers (especially PhD students) to maintain the fruitful exchanges between the two countries in the future years, in a structured and federated way.

8.3. International Initiatives

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

Carole Delporte-Gallet and Hugues Fauconnier are members of the Inria-MEXICO Equipe Associée LiDiCo (At the Limits of Distributed Computability, https://sites.google.com/site/lidicoequipeassociee/).

8.3.2. Inria International Partners

8.3.2.1. Informal International Partners

Ofer Feinerman (Physics department of complex systems, Weizmann Institute of Science, Rehovot, Israel), is a team member in Amos Korman's ERC project DBA. This collaboration has been formally established by signing a contract between the CNRS and the Weizmann Institute of Science, as part of the ERC project.

Rachid Guerraoui (School of Computer and Communication Sciences, EPFL, Switzerland) maintains an active research collaboration with Gang team members (Carole Delporte, Hugues Fauconnier).

Sergio Rajsbaum (UNAM, Mexico) is a regular collaborator of the team, also involved formally in a joint French-Mexican research project (see next subsection).

Boaz Patt-Shamir (Tel Aviv University, Israel) is a regular collaborator of the team, also involved formally in a joint French-Israeli research project (see next subsection).

8.4. International Research Visitors

8.4.1. Visits to International Teams

- Laurent Viennot has visited Archontia Giannopoulou at National and Kapodistrian University of Athens from July 1st to July 7th.
- Michel Habib has visited Prof. M. Chen (Xiamen University of Technology) and Prof. Lin Cheng-Kuan (Fuzhou University) in China, 9-15 december.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. General Chair, Scientific Chair

Adrian Kosowski was one of the organizing chairs of the workshop "CELLS: Computing among Cells", co-located with DISC 2019, Budapest, October 2019.

9.1.2. Scientific Events: Selection

9.1.2.1. Member of the Conference Program Committees

Carole Delporte was PC member of DISC 2019

Carole Delporte was PC member of NETYS 2019

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Pierre Fraigniaud is a member of the Editorial Board of Distributed Computing (DC).
- Pierre Fraigniaud is a member of the Editorial Board of Theory of Computing Systems (TOCS).

9.1.4. Invited Talks

Carole Delporte, Collège de France, Protocoles de population, Paris mars 2019.

Michel Habib International Summer School on Networks and Evolution, ”Centralities in Networks”, Roscoff juin 2019.

Michel Habib Participation to a CIMPA Research School in Tabriz (IRAN) “Graphs, Algorithms and Randomness”, june 2019. Cours de 8h: ”Efficient graph algorithms and applications”.

Michel Habib "Diameter computations in graphs and shortest paths problems", Université Libre de Bruxelles, june 2019

Pierre Fraigniaud was invited speaker [18] at SIROCCO 2019.

9.1.5. Research Administration

- Hugues Fauconnier is director of the UFR d'informatique of Université Paris Diderot.
- Carole Delporte-Gallet is deputy director of the UFR d'informatique of Université Paris Diderot.
- Laurent Viennot is leader of the “Algorithms and discrete structures” department of the Institute de Recherche en Informatique Fondamentale (IRIF).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: Carole Delporte and Hugues Fauconnier, Algorithmique distribuée avec mémoire partagée, 6h, M2, Université Paris Diderot

Master: Hugues Fauconnier, Cours Algorithmes répartis, 33h, M2, Univ. Paris Diderot

Master: Carole Delporte, Cours et TP Protocoles des services internet, 44h, M2, Univ. Paris Diderot

Master: Carole Delporte, Cours Programmation répartie, 33h, M2, Univ. Paris Diderot

Master: Carole Delporte and Hugues Fauconnier, Théorie et pratique de la concurrence, 48h, M1, Université Paris Diderot

Licence: Carole Delporte and Hugues Fauconnier, Culture informatique, 16h, L2, Univ. Paris Diderot

Licence: Boufkhad Yacine, Algorithmique et Informatique, 132h, L1, IUT de l'Université Paris Diderot

Licence: Boufkhad Yacine, Programmation Orientée Objet, 60h, L2, IUT de l'Université Paris Diderot

Licence: Boufkhad Yacine, Traitement de données, 16h, L2, IUT de l'Université Paris Diderot

DU Santé connectée, 12h réseau, Université Paris Diderot, public divers.

Master: Pierre Fraigniaud, Algorithmique parallèle et distribuée, 24h, Ecole Centrale Supélec Paris, M2

Master: Adrian Kosowski, Competitive Programming - Modal Course, 60h, M1, École Polytechnique

Licence: Adrian Kosowski, Design and Analysis of Algorithms, 32h, L3, École Polytechnique

Master: Pierre Fraigniaud and Adrian Kosowski, Algorithmique distribuée pour les réseaux, 24h, M2, Master Parisien de Recherche en Informatique (MPRI)

Master: Fabien de Montgolfier, Grand Réseaux d'Interaction, 44h, M2, Univ Paris Diderot

Licence: Fabien de Montgolfier, Protocoles Réseau (TP/TD), 24h, M1, Univ Paris Diderot

Licence: Fabien de Montgolfier, Programmation avancée (cours/TD/projet, bio-informatique), 52h, L3, Univ. Paris Diderot

Master: Fabien de Montgolfier, Algorithmique avancée (bio-informatique), 26h, M1, Univ Paris Diderot

Licence: Fabien de Montgolfier, Algorithmique (TD), 26h, L3, Ecole d'Ingénieurs Denis Diderot

License : Laurent Viennot, Algorithms, 12h, L3-M1 BioInfo, Univ. Paris Diderot

Licence: Pierre Charbit, Elements d'Algorithmique, 24h, L2, Université Paris Diderot, France

Licence: Pierre Charbit, Automates finis, 36h, L2, Université Paris Diderot, France

Licence: Pierre Charbit, Internet et Outils, 52h, L1, Université Paris Diderot, France

Master: Pierre Charbit, Programmation Objet, 60h, M2Pro PISE, Université Paris Diderot, France

Master: Pierre Charbit, Algorithmique de Graphes, 12h, M2 MPRI, Université Paris Diderot, France

9.2.2. Supervision

PhD defended: Simon Collet (co-advised by Amos Korman and Pierre Fraigniaud) [1]. Title of thesis is: "Algorithmic Game Theory Applied to Biology". Started September 2015, defended on December 9th, 2019 at Paris University.

PhD defended: Mengchuan Zou (co-advised by Adrian Kosowski and Michel Habib) [2]. Title of thesis is: "Aspects of Efficiency in Selected Problems of Computation on Large Graphs". Started October 2016, defended on December 17th at Paris University.

PhD in progress: Brieuc Guinard (advised by Amos Korman). Title of thesis is: "Algorithmic Aspects of Random Biological Processes". Started October 2016.

PhD in progress: Zeinab Nehaï (advised by Hugues Fauconnier). Title of thesis is : Verification of block-chain. Started October 2018

9.2.3. Juries

Carole Delporte-Gallet was on the jury committee of the PhD thesis of Jules Chouquet "Une Géométrie du calcul Réseaux de preuve, Appel-Par-Pousse-Valeur et topologie du Consensus", december 2019, Université Paris Diderot.

Carole Delporte-Gallet was the referee and on jury committee of the HDR thesis of Janna Burman "Distributed Computing with Limited Resources", december 2019, Université Paris-Sud.

Carole Delporte-Gallet was on the jury committee of the PhD thesis of Marjorie Bournat "Graceful Degradation and Speculation for Robots in Highly Dynamic Environments, june 2019, Sorbonne Université.

Hugues Fauconnier was on jury committee of the HDR thesis of Janna Burman "Distributed Computing with Limited Resources", december 2019, Université Paris-Sud.

Michel Habib was on the jury committee of the PhD thesis of Keno Merckx "Optimization and Realizability Problems for Convex Geometries", june 2019, Université Libre de Bruxelles.

Michel Habib was on the jury committee of the PhD thesis of Mehdi Kosravian Ghadikolaei, "Extension of NP Optimization problems", Université de Dauphine (PSL), july 2019.

Michel Habib was on the jury committee of the PhD thesis of Antoine Roux, "Etude d'un code correcteur linéaire pour le canal à effacements de paquets et optimisation par comptage de forêts et calcul modulaire" at Sorbonne University, december 2019.

Laurent Viennot was referee and on the jury committee of the HDR thesis of David Ilcinkas on "Structural Information in Distributed Computing" at the University of Bordeaux, March 2019.

Laurent Viennot was referee and on the jury committee of the PhD thesis of Sébastien Ratel on "Densité, VC-dimension et étiquetages de graphes" at the University of Aix-Marseille, November 2019.

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] S. COLLET.*Algorithmic Game Theory Applied to Networks and Populations*, Université Paris 7 - Diderot, December 2019, https://hal.inria.fr/tel-02433566

[2] M. ZOU.*Aspects of Efficiency in Selected Problems of Computation on Large Graphs*, Université de Paris, December 2019, https://tel.archives-ouvertes.fr/tel-02436610

Articles in International Peer-Reviewed Journal

[3] J. ASPLUND, P. CHARBIT, C. FEGHALI.*Enclosings of decompositions of complete multigraphs in 2-edge-connected r -factorizations*, in "Discrete Mathematics", August 2019, vol. 342, n^o 8, p. 2195-2203 [*DOI :* 10.1016/J.DISC.2019.04.021], https://hal.inria.fr/hal-02429274

[4] A. BALLIU, P. FRAIGNIAUD.*Certification of Compact Low-Stretch Routing Schemes*, in "The Computer Journal", May 2019, vol. 62, n^o 5, p. 730-746 [*DOI :* 10.1093/COMJNL/BXY089], https://hal.inria.fr/hal-02433473

[5] P. BERENBRINK, R. KLASING, A. KOSOWSKI, F. MALLMANN-TRENN, P. UZNANSKI.*Improved Analysis of Deterministic Load-Balancing Schemes*, in "ACM Transactions on Algorithms", 2019, vol. 15, n^o 1, Art.10, https://arxiv.org/abs/1404.4344 - minor corrections; updated literature overview [*DOI :* 10.1145/3282435], https://hal.inria.fr/hal-00979691

[6] E. BIRMELE, F. DE MONTGOLFIER, L. PLANCHE.*On the Minimum Eccentricity Isometric Cycle Problem*, in "Electronic Notes in Theoretical Computer Science", August 2019, vol. 346, p. 159-169 [*DOI :* 10.1016/J.ENTCS.2019.08.015], https://hal.inria.fr/hal-02430756

[7] L. BOCZKOWSKI, A. KORMAN, E. NATALE.*Minimizing message size in stochastic communication patterns: fast self-stabilizing protocols with 3 bits*, in "Distributed Computing", June 2019, A preliminary version of this work appears as a 3-pages Brief Announcement in PODC 2016 (https://doi.org/10.1145/2933057.2933075) and as an extended abstract at SODA 2017 (https://doi.org/10.1137/1.9781611974782.168), forthcoming [*DOI :* 10.1007/S00446-018-0330-X], https://hal.archives-ouvertes.fr/hal-01965945

[8] A. CASTAÑEDA, C. DELPORTE-GALLET, H. FAUCONNIER, S. RAJSBAUM, M. RAYNAL.*Making Local Algorithms Wait-Free: the Case of Ring Coloring*, in "Theory of Computing Systems", May 2019, p. 344–365 [*DOI :* 10.1007/S00224-017-9772-Y], https://hal.archives-ouvertes.fr/hal-01672723

[9] V. CHEPOI, F. F. DRAGAN, M. HABIB, Y. VAXÈS, H. ALRASHEED.*Fast approximation of eccentricities and distances in hyperbolic graphs*, in "Journal of Graph Algorithms and Applications", 2019, vol. 23, n^o 2, p. 393-433 [*DOI :* 10.7155/JGAA.00496], https://hal.inria.fr/hal-02423453

[10] P. FRAIGNIAUD, A. KORMAN, Y. RODEH.*Parallel Bayesian Search with No Coordination*, in "Journal of the ACM (JACM)", June 2019, vol. 66, n^o 3, p. 1-28 [*DOI :* 10.1145/3304111], https://hal.archives-ouvertes.fr/hal-02351345

[11] P. FRAIGNIAUD, A. KORMAN, Y. RODEH.*Parallel Bayesian Search with No Coordination*, in "Journal of the ACM (JACM)", June 2019, vol. 66, n^o 3, p. 1-28 [*DOI :* 10.1145/3304111], https://hal.inria.fr/hal-02433483

[12] P. FRAIGNIAUD, E. NATALE.*Noisy rumor spreading and plurality consensus*, in "Distributed Computing", August 2019, vol. 32, n^o 4, p. 257-276 [*DOI :* 10.1007/S00446-018-0335-5], https://hal.inria.fr/hal-02433482

[13] P. FRAIGNIAUD, D. OLIVETTI.*Distributed Detection of Cycles*, in "ACM Transactions on Parallel Computing", December 2019, vol. 6, n^o 3, p. 1-20 [*DOI :* 10.1145/3322811], https://hal.inria.fr/hal-02433486

[14] P. FRAIGNIAUD, B. PATT-SHAMIR, M. PERRY.*Randomized proof-labeling schemes*, in "Distributed Computing", June 2019, vol. 32, n^o 3, p. 217-234 [*DOI :* 10.1007/S00446-018-0340-8], https://hal.inria.fr/hal-02433478

[15] N. GHAREGHANI, M. HABIB, M. R. HOOSHMANDASL, R. NASERASR, P. SHARIFANI.*When an optimal dominating set with given constraints exists*, in "Theoretical Computer Science", July 2019 [*DOI :* 10.1016/J.TCS.2019.02.012], https://hal.inria.fr/hal-02423707

[16] A. KOSOWSKI, D. PAJAK.*Does Adding More Agents Make a Difference? A Case Study of Cover Time for the Rotor-Router*, in "Journal of Computer and System Sciences", 2019, vol. 106, p. 80-93 [*DOI :* 10.1016/J.JCSS.2019.07.001], https://hal.inria.fr/hal-00950743

[17] A. WATSON, M. HABIB, E. BAPTESTE.*Phylosystemics: Merging Phylogenomics, Systems Biology, and Ecology to Study Evolution*, in "Trends in Microbiology", November 2019 [*DOI :* 10.1016/J.TIM.2019.10.011], https://hal.inria.fr/hal-02423510

International Conferences with Proceedings

[18] A. CASTAÑEDA, P. FRAIGNIAUD, A. PAZ, S. RAJSBAUM, M. ROY, C. TRAVERS.*A Topological Perspective on Distributed Network Algorithms*, in "26th International Colloquium on Structural Information and Communication Complexity (SIROCCO 2019)", L'Aquila, Italy, K. CENSOR-HILLEL, M. FLAMMINI (editors), LNCS - Lecture Notes in Computer Science, Springer, July 2019, vol. 11639, https://hal.inria.fr/hal-02433511

[19] *Best Paper*
A. CASTAÑEDA, P. FRAIGNIAUD, A. PAZ, S. RAJSBAUM, M. ROY, C. TRAVERS.*Synchronous t-Resilient Consensus in Arbitrary Graphs*, in "SSS 2019 - 21st International Symposium on Stabilization, Safety, and Security of Distributed Systems", Pisa, Italy, October 2019 [*DOI :* 10.1007/978-3-030-34992-9_5], https://hal.inria.fr/hal-02433524.

[20] P. CRESCENZI, P. FRAIGNIAUD, A. PAZ.*Trade-Offs in Distributed Interactive Proofs*, in "DISC 2019 - 33rd International Symposium on Distributed Computing", Budapest, Hungary, October 2019, https://hal.inria.fr/hal-02433529

[21] P. CRESCENZI, P. FRAIGNIAUD, A. PAZ.*Simple and Fast Distributed Computation of Betweenness Centrality*, in "IEEE INFOCOM 2020 - International Conference on Computer Communications", Pekin, China, April 2020, https://hal.inria.fr/hal-02433541

[22] G. DUCOFFE, M. HABIB, L. VIENNOT.*Fast Diameter Computation within Split Graphs*, in "COCOA 2019 - 13th Annual International Conference on Combinatorial Optimization and Applications", Xiamen, China, December 2019, https://arxiv.org/abs/1910.03438 , https://hal.inria.fr/hal-02307397

[23] G. DUCOFFE, M. HABIB, L. VIENNOT.*Diameter computation on H-minor free graphs and graphs of bounded (distance) VC-dimension*, in "ACM-SIAM Symposium on Discrete Algorithms (SODA20)", Salt Lake City, United States, January 2020, https://arxiv.org/abs/1907.04385 , https://hal.inria.fr/hal-02340382

[24] P. FRAIGNIAUD, P. MONTEALEGRE, R. OSHMAN, R. IVAN, I. TODINCA.*On Distributed Merlin-Arthur Decision Protocols*, in "SIROCCO 2019 - 26th International Colloquium on Structural Information and Communication Complexity", L'Aquila, Italy, K. CENSOR-HILLEL, M. FLAMMINI (editors), LNCS - Lecture Notes in Computer Science, 2019, vol. 11639, https://hal.inria.fr/hal-02433519

[25] S. GUPTA, A. KOSOWSKI, L. VIENNOT.*Exploiting Hopsets: Improved Distance Oracles for Graphs of Constant Highway Dimension and Beyond*, in "ICALP 2019 - 46th International Colloquium on Automata, Languages, and Programming", Patras, Greece, July 2019, https://arxiv.org/abs/1803.06977 , https://hal.inria.fr/hal-01737210

[26] M. HABIB, F. DE MONTGOLFIER, L. MOUATADID, M. ZOU.*A General Algorithmic Scheme for Modular Decompositions of Hypergraphs and Applications*, in "IWOCA 2019 - 30th International Workshop on Combinatorial Algorithms", Pise, Italy, July 2019, https://hal.inria.fr/hal-02423508

[27] A. KORMAN, Y. RODEH.*Multi-Round Cooperative Search Games with Multiple Players*, in "ICALP 2019 - 46th International Colloquium on Automata, Languages and Programming", Patras, Greece, July 2019, https://arxiv.org/abs/1811.01270 , https://hal.archives-ouvertes.fr/hal-02105524

[28] A. KOSOWSKI, P. UZNANSKI, L. VIENNOT.*Hardness of Exact Distance Queries in Sparse Graphs Through Hub Labeling*, in "PODC '19 - ACM Symposium on Principles of Distributed Computing", Toronto, Canada, 2019, p. 272-279, https://arxiv.org/abs/1902.07055 [*DOI :* 10.1145/3293611.3331625], https://hal.inria.fr/hal-02429930

[29] P. PENNA, L. VIENNOT.*Independent Lazy Better-Response Dynamics on Network Games*, in "CIAC 2019 2019 : 11th International Conference on Algorithms and Complexity", Rome, Italy, May 2019, https://arxiv.org/abs/1609.08953 , https://hal.inria.fr/hal-01373411

[30] D.-M. PHAN, L. VIENNOT.*Fast Public Transit Routing with Unrestricted Walking through Hub Labeling*, in "Special Event on Analysis of Experimental Algorithms (SEA2)", Kalamata, Greece, Special Event on Analysis of Experimental Algorithms, June 2019, https://arxiv.org/abs/1906.08971 , https://hal.inria.fr/hal-02161283

Conferences without Proceedings

[31] C. DELPORTE-GALLET, H. FAUCONNIER, P. FRAIGNIAUD, M. RABIE.*Brief Announcement: Distributed Computing in the Asynchronous LOCAL Model*, in "SSS 2019 - 21st International Symposium on Stabilization, Safety, and Security of Distributed Systems", Pisa, Italy, October 2019, https://hal.archives-ouvertes.fr/hal-02436289

[32] C. DELPORTE-GALLET, H. FAUCONNIER, Y. JURSKI, F. LAROUSSINIE, A. SANGNIER.*Towards Synthesis of Distributed Algorithms with SMT Solvers*, in "NETYS 2019 - International conference on networked systems", Marrakech, Morocco, June 2019, https://hal.archives-ouvertes.fr/hal-02434364

[33] C. DELPORTE-GALLET, H. FAUCONNIER, M. RAYNAL.*Participant-Restricted Consensus in Asynchronous Crash-Prone Read/Write Systems and Its Weakest Failure Detector*, in "PaCT 2019 - 15th International Conference on Parallel Computing Technologies", Almaty, Kazakhstan, August 2019, https://hal.archives-ouvertes.fr/hal-02434365

[34] C. DELPORTE-GALLET, H. FAUCONNIER, M. RAYNAL.*On the Weakest Failure Detector for Read/Write-Based Mutual Exclusion*, in "AINA 2019 - 33rd International Conference on Advanced Information Networking and Applications", Matsue, Japan, March 2020, https://hal.archives-ouvertes.fr/hal-02433606

Scientific Books (or Scientific Book chapters)

[35] C. DELPORTE-GALLET, P. A. ABDULLA.*Special Issue on NETYS 2016, Computing (journal)*, Springer, 2019, https://hal.inria.fr/hal-01955924

Other Publications

[36] H. BAKTASH, E. NATALE, L. VIENNOT.*A Comparative Study of Neural Network Compression*, October 2019, https://arxiv.org/abs/1910.11144 - working paper or preprint, https://hal.inria.fr/hal-02321581

Team KOPERNIC

Keeping wOrst case reasoning aPpropriatE foR differeNt critICalities

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
Paris

THEME
Embedded and Real-time Systems

Table of contents

Team KOPERNIC

Creation of the Team: 2018 July 03

Keywords:

Computer Science and Digital Science:

A1.1.1. - Multicore, Manycore
A1.5. - Complex systems
A1.5.1. - Systems of systems
A1.5.2. - Communicating systems
A2.3. - Embedded and cyber-physical systems
A2.3.1. - Embedded systems
A2.3.2. - Cyber-physical systems
A2.3.3. - Real-time systems
A2.4.1. - Analysis

Other Research Topics and Application Domains:

B5.2. - Design and manufacturing
B5.2.1. - Road vehicles
B5.2.2. - Railway
B5.2.3. - Aviation
B5.2.4. - Aerospace
B6.6. - Embedded systems

1. Team, Visitors, External Collaborators

Research Scientists

Liliana Cucu [Team leader, Inria, Researcher, HDR]
Yves Sorel [Inria, Senior Researcher]

External Collaborator

Adriana Gogonel [Statinf]

Technical Staff

Rihab Bennour [Inria, Engineer]
Mehdi Mezouak [Inria, Engineer, until Oct 2019]

PhD Students

Slim Ben Amor [Inria, PhD Student]
Evariste Ntaryamira [France Embassy at Burundi, PhD Student, from May 2016]
Walid Talaboulma [Inria, PhD Student]
Kevin Zagalo [Inria, PhD Student, from Oct 2019]

Post-Doctoral Fellow

Roberto Medina Bonilla [Inria, Post-Doctoral Fellow, from Feb 2019]

Visiting Scientist

George Lima [Inria, from Jul 2019 until Aug 2019]

Administrative Assistants

Christine Anocq [Inria, Administrative Assistant]
Nelly Maloisel [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

The Kopernic members are focusing their research on studying **time for embedded communicating systems**, also known as cyber-physical systems.

The term cyber-physical systems refers to a new generation of systems with integrated computational and physical capabilities that can interact with humans through many new modalities [15]. A defibrillator, a mobile phone, an autonomous car or an aircraft, they all are CPSs. Beside constraints like power consumption, security, size and weight, CPSs may have cyber components required to fulfill their functions within a limited time interval (a.k.a. dependability), often imposed by the environment, e.g., a physical process controlled by some cyber components. The appearance of communication channels between cyber-physical components, easing the CPS utilization within larger systems, forces cyber components with high criticality to interact with lower criticality cyber components. This interaction is completed by external events from the environnement that has a time impact on the CPS. Moreover, some programs of the cyber components may be executed on predictable processors and other programs on less predictable processors. For instance, a drone that supervises an airport area may be pictured continually interacting with the airport control tower and the pilotes of the airplanes. In this exemple, the drone, the tower and the airplanes belong to a large CPS.

Different research communities study separately the three design phases of these systems: the modeling, the design and the analysis of CPSs [23]. These phases are repeated iteratively until an appropriate solution is found. During the first phase, the behavior of a system is often described using model-based methods. Other methods exist, but model-driven approaches are widely used by both the research and the industry communities. A solution described by a model is proved (functionally) correct usually by a formal verification method used during the analysis phase (third phase described below).

During the second phase of the design, the physical components (e.g., sensors and actuators) and the cyber components (e.g., programs, messages and embedded processors) are chosen often among those available on the market. However, due to the ever increasing pressure of smartphone market, the microprocessor industry provides general purpose processors based on multicore and, in a near future, based on manycore processors. These processors have complex architectures that are not time predictable due to features like multiple levels of caches and pipelines, speculative branching, communicating through shared memory or/and through a network on chip, internet, etc. Therefore, nowadays the CPS industry is facing the great challenge of estimating the corresponding worst case execution times of programs executed on these processors. Indeed, the current complexity of both processors and programs does not allow to propose reasonable worst case bounds. Then, the phase of design ends with the implementation of the cyber components on such processors, where the models are transformed in programs (or messages for the communication channels) manually or by code generation techniques [17].

During the third phase of analysis, the correctness of the cyber components is verified at program level where the functions of the cyber component are implemented. The execution times of programs are estimated either by static analysis, by measurements or by a combination of both approaches [30].

The time properties of a cyber component are subject to variability factors. We understand by variability the distance between the smallest value and the largest value of a time property. With respect to the time properties of a CPS, the factors may be classified in three main classes:

- program structure: for instance, the execution time of a program that has two main branches is obtained, if appropriate composition principles apply, as the maximum between the largest execution time of each branch. In this case the branch is a variability factor on the execution time of the program;
- processor structure: for instance, the execution time of a program on a less predictable processor (e.g., one core, two levels of cache memory and one main memory) will have a larger variability than the execution time of the same program executed on a more predictable processor (e.g., one

core, one main memory). In this case the cache memory is a variability factor on the execution time of the program;

- execution environnement: for instance, the appearance of a pedestrian in front of a car triggers the execution of the program corresponding to the brakes in an autonomous car. In this case the pedestrian is a variability factor for triggering the execution of some programs. Moreover, the execution environnement may trigger branches of the programs, according to their structure.

Verifying that time properties of a CPS are met is often formalized as a scheduling problem [25], where the programs should be provided a start time within the schedule together with an assignment of resources (processor, memory, communication, etc.). The verification of a solution for a scheduling problem is known as schedulability analysis.

A cyber-physical system (CPS) has cyber (or computational) components and physical components that communicate. Our team deals with the problem of studying time properties (execution time of a program or a set of communicating programs, etc.) of the cyber components of a CPS. The cyber components may implement functions with different criticalities with respect to time and a solution should come with associated proofs of its *appropriateness* for each criticality. A solution is appropriate for a criticality level if all functions fulfill the expectations of that criticality level. Based on their mathematical foundations, the solutions are:

- either classic (or non-probabilistic) when all time properties are estimated and/or upper bounded by numerical values;
- or probabilistic when at least one time property is estimated and/or upper bounded by probability distributions.

The Kopernic members propose a **system-oriented solution** to the problem of studying time properties of the cyber components of a CPS. The solution is expected to be obtained by composing probabilistic and non-probabilistic approaches for CPSs.

We identify three main scientific objectives developed in Sections 3.1, 3.2 and 3.3. These objectives are presented from program level, where we use statistical approaches, to the level of all programs, where we use probabilistic and non-probabilistic approaches.

3. Research Program

3.1. Worst case execution time estimation of a program

Modern processors induce an increased variability of the execution time of programs, making difficult (or even impossible) a complete static analysis. Our objective is to propose a solution composing probabilistic and non-probabilistic approaches based both on static and on statistical analyses by answering the following **scientific challenges**:

1. **a classification of the variability of execution times** of a program with respect to the processor features. We will use as first measure our statistical estimator based on the Extreme Value Theory [18], [20]. An implementation of the estimator is available at http://inria-rscript.serveftp.com. The access to this later page requires a login (aoste) and a password (aoste). The difficulty of this challenge is related to the definition of an element belonging to the set of variability factors and its mapping to the execution time of the program.
2. **a compositional rule** of statistical models based on Bayesian approaches. The difficulty of this challenge comes from the fact that a global maximum cannot be obtained by upper bounding the corresponding local maxima. We will use as first rule of composition a Bayesian approach [22]. We consider as first statistical model those obtained by any static analysis of the program on a basic processor. Through the Bayesian approach we add iteratively the variability due to each processor feature as a new statistical model. The convergence of the global model is decided once no variability is detected at the level of the statistical estimator providing the bounds on the execution time of the program.

The problem of estimating the worst case execution time of a program is an excellent opportunity for the Extreme Values community to validate and to evolve as the context of obtaining measures is indefinitely reproducible.

3.2. Deciding the schedulability of all programs running within the same cyber component

In this context, the programs may have different time criticalities, but they share the same processor, possibly multicore [0]. Our objective is to propose a solution composing probabilistic and non-probabilistic approaches based on answers to the following **scientific challenges**:

1. **scheduling algorithms taking into account the interaction between different variability factors**. The proposed scheduling algorithms are the theoretical bases of a scheduler able to guarantee the time constraints of the cyber component. The existence of time parameters described by probability distributions imposes to answer to the challenge of revisiting scheduling algorithms that lose their optimality even in the case of an unicore processor [26]. Moreover, the multicore partionning problem is, also, recognized difficult for the non-probabilistic case [29];
2. **schedulability analyses** based on the algorithms proposed previously. In the case of predictable processors, the schedulability analyses accounting for operating systems costs increase the dependability of CPSs [28]. Morever, in presence of variability factors, the additivity property of non-probabilistic approaches is lost and new composition principles are required. We will propose new composition principles based on our preliminary results on the propagation of the probabilistic constraints [16]. The definition of these principles form the challenge related to this objective.

3.3. Deciding the schedulability of all programs communicating through predictable and non-predictable networks

In this case the programs of the same cyber component execute on the same processor and they may communicate with the programs of other cyber components through networks that may be predictable (network on chip) or non-predictable (internet, telecommunications). Our objective is to propose a solution to the challenge of analysing schedulability of programs, for which existing (worst case) probabilistic solutions exist [27], communicating through networks, for which probabilistic worst-case solutions [19] and average solutions exist [24]. Our solution is based on the results obtained for the two first objectives, making this third objective a longer-term one.

4. Application Domains

4.1. Avionics

This work is based on a direct collaboration between Airbus and Inria, complementary to collaborative projects like PIA LEOC Capacites and CIFRE thesis. The time critical solutions in this context are based on temporal and spatial isolation of the programs and the understanding of multicore interferences is crucial. Our contributions belong mainly to the solutions space for the objective identified in Section 3.1.

4.2. Railway

This work is based on a direct collaboration with Clearsy and SNCF, complementary to collaborative projects like PIA BGLE Departs and FUI 21 Waruna. The time critical solutions in this context concern both the proposition of an appropriate scheduler and associated schedulability analyses. Our contributions belong to the solutions space of problems dealt within the objectives identified in Section 3.1.

[0] This case is referred as a mixed criticality approach.

4.3. Autonomous cars

This work is based on a direct collaboration with RITS (Inria project team). The time critical solutions in this context concern the interaction between programs executed on multicore processors and messages transmitted through wireless communication channels. Our contributions belong to the solutions space of all three classes of problems dealt within the objectives identified in Section 3.2.

4.4. Drones

This work is based on the collaborative project FUI/FEDER 22 Ceos. As in the case of autonomous cars, there is an interaction between programs and messages, suggesting that our contributions in this context belong to the solutions space of all three classes of problems dealt within the objectives identified in Section 3.2.

5. Highlights of the Year

5.1. Highlights of the Year

The Kopernic research results on statistical estimation of execution time bounds has been transferred to a start-up, led by Adriana Gogonel, postdoctoral student in Kopernic team. The start-up, Statinf, has been a Carnot 2019 and Wilco 2019 laureate and it has integrated the Agoranov Deeptech incubator since September 2019.

The Kopernic leader, Liliana Cucu-Grosjean has been the IEEE RTSS2019 Track co-chair as well as the DATE2020 Real-time Systems Track co-chair.

6. New Software and Platforms

6.1. SynDEx

KEYWORDS: Distributed - Optimization - Real time - Embedded systems - Scheduling analyses

SCIENTIFIC DESCRIPTION: SynDEx is a system level CAD software implementing the AAA methodology for rapid prototyping and for optimizing distributed real-time embedded applications. It is developed in OCaML.

Architectures are represented as graphical block diagrams composed of programmable (processors) and non-programmable (ASIC, FPGA) computing components, interconnected by communication media (shared memories, links and busses for message passing). In order to deal with heterogeneous architectures it may feature several components of the same kind but with different characteristics.

Two types of non-functional properties can be specified for each task of the algorithm graph. First, a period that does not depend on the hardware architecture. Second, real-time features that depend on the different types of hardware components, ranging amongst execution and data transfer time, memory, etc.. Requirements are generally constraints on deadline equal to period, latency between any pair of tasks in the algorithm graph, dependence between tasks, etc.

Exploration of alternative allocations of the algorithm onto the architecture may be performed manually and/or automatically. The latter is achieved by performing real-time multiprocessor schedulability analyses and optimization heuristics based on the minimization of temporal or resource criteria. For example while satisfying deadline and latency constraints they can minimize the total execution time (makespan) of the application onto the given architecture, as well as the amount of memory. The results of each exploration is visualized as timing diagrams simulating the distributed real-time implementation.

Finally, real-time distributed embedded code can be automatically generated for dedicated distributed real-time executives, possibly calling services of resident real-time operating systems such as Linux/RTAI or Osek for instance. These executives are deadlock-free, based on off-line scheduling policies. Dedicated executives induce minimal overhead, and are built from processor-dependent executive kernels. To this date, executives kernels are provided for: TMS320C40, PIC18F2680, i80386, MC68332, MPC555, i80C196 and Unix/Linux workstations. Executive kernels for other processors can be achieved at reasonable cost following these examples as patterns.

FUNCTIONAL DESCRIPTION: Software for optimising the implementation of embedded distributed real-time applications and generating efficient and correct by construction code

NEWS OF THE YEAR: We improved the distribution and scheduling heuristics to take into account the needs of co-simulation.

- Participant: Yves Sorel
- Contact: Yves Sorel
- URL: http://www.syndex.org

6.2. EVT Kopernic

KEYWORDS: Embedded systems - Worst Case Execution Time - Real-time application - Statistics

SCIENTIFIC DESCRIPTION: The EVT-Kopernic tool is an implementation of the Extreme Value Theory (EVT) for the problem of the statistical estimation of worst-case bounds for the execution time of a program on a processor. Our implementation uses the two versions of EVT - GEV and GPD - to propose two independent methods of estimation. Their results are compared and only results that are sufficiently close allow to validate an estimation. Our tool is proved predictable by its unique choice of block (GEV) and threshold (GPD) while proposant reproducible estimations.

FUNCTIONAL DESCRIPTION: EVT-Kopernic is tool proposing a statistical estimation for bounds on worst-case execution time of a program on a processor. The estimator takes into account dependences between execution times by learning from the history of execution, while dealing also with cases of small variability of the execution times.

NEWS OF THE YEAR: Any statistical estimator should come with an representative measurement protocole based on the processus of composition, proved correct. We propose the first such principle of composition while using a Bayesien modeling taking into account iteratively different measurement models. The composition model has been described in a patent submitted this year with a scientific publication under preparation.

- Participants: Adriana Gogonel and Liliana Cucu
- Contact: Adriana Gogonel
- URL: http://inria-rscript.serveftp.com/

7. New Results

7.1. Uniprocessor Mixed-Criticality Real-Time Scheduling

In the context of the FUI CEOS project 8.1.1.1, last two years we transformed the free software program PX4, which performs the autopilot of the CEOS drone, in a graph of hard real-time tasks. This transformation was intended to achieve a schedulability analysis guaranteeing the autopilot is able to perform safety critical missions since its behaviour is deemed to be hard real-time, i.e., all deadlines of all tasks are satisfied. It is worth noting that the autopilot is one of the most important programs of the drone since it maintains its stability not only during hover phases but also during automatic flight missions from one GPS point to another. This transformation resulted in a "real-time autopilot" that we called PX4-RT.

For the first version of PX4-RT we chose, as periods, the periods used in the original version of PX4 which was not hard real-time as we shown last year. Then, since these periods was inherited from an automatic control analysis achieved by initial designers of PX4 in a non hard real-time context, we had to determine the right combination of periods of tasks, allowing on the one hand to correctly control the drone, and on the other hand, using a schedulabilty analysis, to satify all the deadlines. In order to achieve this goal, we used a hardware in the loop simulation (HitL) which simulates only the sensors and the actuators, whereas the PX4-RT program runs on the Pixhawk board based on an ARM Cortex-M4 uniprocessor. Eventually, we determined some period combinations that fit our needs, other combinations did not allow the drone to follow correctly the given mission or resulted in a crash. Moreover, we verified that all the right combinations led to a schedulable set of tasks, meaning that corresponding versions of PX4-RT were hard real-time. Finally, we used with success the best combination of periods to run PX4-RT on the real drone of CEOS during a simple flight. Of course, we plan to achieve numerous realistic flights planned in the three industrial use cases of the CEOS project.

In addition to this study intended to determine the right combination of periods, we addressed two other issues. In the first one we tried to decrease the worst case execution times (WCET) of tasks in order to increase the schedulability ratio. Such decrease allows to add on the same processor new tasks presently executed on other processors, e.g., mission planning, fault tolerance, etc. Since we found out that the Kalman filter had the largest measured execution time of all the tasks, we studied the Kalman filter algorithm implemented in PX4 to decrease its WCET. We suppressed the two states of the Kalman Filter corresponding to the wind speed estimation since our drone do not have a sensor measuring this speed. Then, we suppressed the three states of the Kalman filter corresponding to the accelerometer bias whose standard deviation was close to zero. Each of these modifications brought an improvement of 15 percent in term of largest measured execution time without decreasing the performances of the drone. In the second issue we started a theoretical study about relations between the stability of a set of automatic control laws and the schedulability of the corresponding set of real-time tasks. In the literature some results exist about one control law corresponding to one real-time task. To the best of our knowledge there is no result for a set of control laws that exchange data.

Finally, we deeply studied NuttX the real-time operating system used presently to support PX4 and PX4-RT autopilot programs. Indeed, we plan to modify the scheduler of this operating system in order to manage real-time tasks more safely. In order to do that we will draw inspiration from the technique proposed in our time triggered offline scheduler that accounts for the preemption and scheduler cost [14].

7.2. Multicore processor graph tasks scheduling

Due to widespread of multicore processors on embedded and real-time systems, we concentrate our work on the study of the schedulability of real-time tasks with precedence constraints on such processors. We consider preemptive fixed-priority scheduling policies. First, we have proposed a response time analysis for directed acyclic graphs task model with non-probabilistic execution time and preemptive fixed-priority scheduling policy [10]. Our response time analysis improves importantly the state of the art analyses, while allowing scalable extensions for response time analysis of tasks with worst case execution times described by probability distributions. We extend this response time analysis to similar task model with probabilistic worst case execution time with the advantage of providing efficient results also for task model with non-probabilistic worst case execution times. Our response time analysis is based on iterative equations which offer run-time enhancement compared to existing work [21] requesting the resolution of complex MILP optimization problem. In addition, we have defined priority on sub-task level enhancing the schedulability and reducing the worst-case response time. The proposed priority assignment algorithm is adapted for the studied task model and it outperforms several state-of-the art methods. We have also proposed a partitioning heuristic that assigns each sub-task to a given core. This heuristic takes into consideration communication delays between sub-tasks inside the same graph in order to minimize the communication while balancing different cores load and maximizing possible parallelism. The proposed heuristics and response time analysis (RTA) are validated on randomly generated task sets and on the PX4-RT drone autopilot programs developed by Kopernic team in FR FUI21 CEOS project.

7.3. Power consumption of probabilistic real-time systems

Energy consumption on real-time systems is a crucial problem nowadays as these systems are becoming complex and are expected to deliver more and more functionalities. At the same time, while the processing demand increases, the vast majority of these systems are powered by batteries and are deployed in hazardous environments making their maintenance difficult and impractical. Existing works on energy consumption and real-time systems are often based on a technique called Dynamic Voltage and Frequency Scaling (DVFS). The principle of this technique is to reduce the frequency of the processor in order to lower its input voltage, consequently reducing the energy required to power the processor. Nevertheless, by reducing the frequency of the processor, programs tend to take more time to complete their execution. In the context of real-time systems, programs need to finish their execution before a given deadline. Therefore, the goal of DVFS techniques is to derive proper frequencies that minimize energy consumption and still ensure that all deadlines of all the programs will be respected. Works carried during this postdoc are twofold. The first contribution consisted in observing how the Worst-Case Execution Time (WCET) of programs varies with regards to the frequency of the processor. Many existing works have considered that the WCET is completely scalable, i.e., a simple factor can be applied to derive a new WCET under a different frequency setup. Nevertheless, researchers have recognize that this hypothesis may be too optimistic since other components, that do not run at the same speed as the processor, e.g., the memory, are used by programs. We derived an experimental setup to observe how the execution of programs varied by setting different frequencies on the processor and the memory. We measured CPU cycles and execution times and it was clear from our experiments that the theoretical speedup bound that should be achieved when the processor is running at its maximum speed is never achieved. We also observed, that DVFS techniques could also be applied to the memory of the system, since some programs do not perform many memory request. Our experiments led to a short paper accepted for the Work-in-Progress session of the 40th Real-Time System Symposium. The paper also introduced the task model that will be used as a basis of the next contribution of the postdoc. This next contribution consists in developing RTA techniques for probabilistic real-time systems in order to derive hardware frequency setups. The inclusion of probabilistic real-time system is motivated by the ever-increasing demand of functionalities for this type of systems. To the best of our knowledge, DVFS techniques in conjunction with probabilistic real-time systems have never been studied. The solution to this optimization problem is ongoing work while preparing the submission of first results beginning of February 2020.

7.4. Data-oriented scheduling approaches

We consider the scheduling problem of tasks using an inter-task communication model based on a circular buffer, which eases the data consistency between tasks [13], [12]. The tasks are scheduled on one processor by a fixed priority preemptive scheduling algorithm and they have implicit deadlines. We provide a formal method calculating the optimal size for each of the buffers while ensuring data consistency, i.e., it is required that a buffer slot is accessed for reading the input data. This later slot will never be used by the producer task to write new data before the execution completion of the instances of all consumers that are currently reading from this slot. As a second contribution, we provide an analytical characterization of the temporal validity and reachability properties of the data flowing in between communicating tasks. These two properties are characterized by considering both tasks execution and data propagation orders. Moreover, we assume that a task instance reads all its inputs data at its activation time and writes back the output data at the completion time where this data becomes immediately available for consumption. Given that, they may be several data samples available in the buffer, we say that a data sample is fresh or temporal valid if, since the time instant it is produced, its producer has not completed another execution. Given that, we use buffers whose size may be larger than one, it is obvious that the consumer task will not implicitly know which data is temporally valid. In order to use the data that reflects the current status of the system environment (valid data), we introduce a novel parameter; the sub-sampling rate used within two scheduling algorithms. These scheduling algorithms ensure the data consistency and temporal validity, while deadlines are met.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. FUI

8.1.1.1. CEOS

This project was started on May 2017. Partners of the project are: ADCIS, ALERION, Aeroport de Caen, EDF, ENEDIS, RTaW, EDF, Thales Communications and Security, ESIEE engineering school and Lorraine University. The CEOS project delivers a reliable and secure system of inspections of pieces of works using professional mini-drone for Operators of Vital Importance coupled with their Geographical Information System. These inspections are carried out automatically at a lower cost than current solutions employing helicopters or off-road vehicles. Several software applications proposed by the industrial partners, are developed and integrated in the drone, within an innovative mixed-criticality approach using multi-core platforms.

8.2. European Initiatives

8.2.1. Collaborations with Major European Organizations

University of York: Real-Time System Group (UK)

Uncertainties in real-time systems: the utilization of extreme value theory has received increased efforts from our community and more rigorous principles are needed for its full understanding. Our two research teams have gathered these principles in a joint publication.

8.3. International Research Visitors

8.3.1. Visits of International Scientists

- Prof. Christopher Gill, Washington University in St. Louis (May 2019).
- Robert Davis, University of York (July 2019).

8.3.1.1. Internships

- Kartikeya Singh (India).

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. Local Chair

The Kopernic team has hosted the 10th edition of the RTSOPS in July 2019.

9.1.2. Scientific Events: Selection

9.1.2.1. Chair of Conference Program Committees

Liliana Cucu-Grosjean is the track chair of IEEE RTSS2019 and DATE2020 Track co-chair for real-time systems

9.1.2.2. Member of the Conference Program Committees

- Liliana Cucu: ACM RACS, RTNS, WFCS

9.1.3. Journal

9.1.3.1. Reviewer - Reviewing Activities

All members of the team are regularly serving as reviewers for the main journals of our domain: Information Processing Letter, Journal of Heuristics, Journal of Real-Time Systems, Journal of Systems Architecture, Journal of Signal Processing Systems, Leibniz Transactions on Embedded Systems, IEEE Transactions on Industrial Informatics, etc.

9.1.4. Scientific Expertise

- Yves Sorel: Steering Committee of System Design and Development Tools Group of Systematic Paris-Region Cluster.
- Yves Sorel: Steering Committee of Technologies and Tools Program of SystemX Institute for Technological Research (IRT).

9.1.5. Research Administration

- Liliana Cucu-Grosjean is co-chair of Inria Committee on gender equality and equal opportunities
- Yves Sorel is chair of the CUMI Paris center commission
- Liliana Cucu-Grosjean was member of Inria Evaluation Commission until September 2019
- Liliana Cucu-Grosjean is member of Inria Scientific Board and CAP CRCN since September 2019
- Liliana Cucu-Grosjean is member of Paris CLHSCT
- Liliana Cucu-Grosjean is member of CES Paris commission
- Yves Sorel is member of the CDT Paris center commission

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: Yves Sorel, Optimization of distributed real-time embedded systems, 38H, M2, University of Paris Sud, France

Master: Yves Sorel, Safe design of reactive systems, 18H, M2, ESIEE Engineering School, Noisy-Le-Grand, France

Engineering school: Liliana Cucu-Grosjean, Graph theory, 30H, ESIEE, Cergy, France

Master: Liliana Cucu-Grosjean, Software Engineering, 30H, ESIEE, Noisy-le-Grand, France

Master: Adriana Gogonel, Machine learning, 30H, ESIEE, Noisy-le-Grand, France

9.2.2. Supervision

PhD in progress: Kevin Zagalo, Statistical predictability of cyber-physical systems, UMPC, started on October 2019, supervised by Liliana Cucu and Prof. Avner Bar-Hen (CNAM)

PhD in progress: Evariste Ntaryamira, Analysis of embedded systems with time and security constraints, UPMC, started on May 2017, supervised by Liliana Cucu and Cristian Maxim (IRT SystemX)

PhD in progress: Slim Ben-Amor, Schedulability analysis of probabilistic real-time tasks under end to end constraints, UPMC, started on November 2016, supervised by Liliana Cucu.

PhD in progress: Walid Talaboulma, Probabilistic timing analysis in presence of dependences, UPMC, started November 2015, co-supervised by Liliana Cucu and Adriana Gogonel (Statinf)

9.2.3. Juries

- Yves Sorel was a HDR jury member for the thesis of Luca Santinelli (ONERA), defended in May 2019.
- Liliana Cucu-Grosjean was a Phd reviewer for the thesis of Roberto Medina (Telecom Paristech), defended in January 2019.
- Liliana Cucu-grosjean was a PhD jury member of Francesco Pozo (Malardelan, Sweden), defended in October 2019.

9.3. Popularization

9.3.1. Articles and contents

- Interview with the Statinf co-founders, published on Inria web site (https://www.inria.fr/centre/paris/actualites/statinf-sonde-les-caeurs-des-systemes-embarques)

9.3.2. Education

- Presentation of Adriana Gogonel at a day organized by a Creteil and Versailles Academy for outreaching Computer Science (April 2019, Paris)

9.3.3. Interventions

- Participation of Adriana Gogonel at a round table entitled "May we innovate within a research institute?" (March 2019, Rocquencourt).

10. Bibliography

Major publications by the team in recent years

[1] L. CUCU-GROSJEAN, A. GOGONEL.*Simulation Device*, March 2016, n° FR2016/050504, https://hal.archives-ouvertes.fr/hal-01666599

[2] L. CUCU-GROSJEAN, L. SANTINELLI, M. HOUSTON, C. LO, T. VARDANEGA, L. KOSMIDIS, J. ABELLA, E. MEZZETTI, E. QUIÑONES, F. J. CAZORLA.*Measurement-Based Probabilistic Timing Analysis for Multi-path Programs*, in "the 24th Euromicro Conference on Real-Time Systems, ECRTS", 2012, p. 91–101

[3] A. GOGONEL, L. CUCU-GROSJEAN.*Dispositif de caractérisation et/ou de modélisation de temps d'exécution pire-cas*, June 2017, n° 1000408053, https://hal.archives-ouvertes.fr/hal-01666535

[4] T. KLODA, A. BERTOUT, Y. SOREL.*Latency analysis for data chains of real-time periodic tasks*, in "the 23rd IEEE International Conference on Emerging Technologies and Factory Automation, ETFA'18", September 2018

[5] C. MAXIM, A. GOGONEL, I. M. ASAVOAE, M. ASAVOAE, L. CUCU-GROSJEAN.*Reproducibility and representativity: mandatory properties for the compositionality of measurement-based WCET estimation approaches*, in "SIGBED Review", 2017, vol. 14, n° 3, p. 24–31

[6] S. E. SAIDI, N. PERNET, Y. SOREL.*Scheduling Real-time HiL Co-simulation of Cyber-Physical Systems on Multi-core Architectures*, in "the 24th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications", August 2018

Publications of the year

Articles in International Peer-Reviewed Journal

[7] R. DAVIS, L. CUCU-GROSJEAN.*A Survey of Probabilistic Schedulability Analysis Techniques for Real-Time Systems*, in "Leibniz Transactions on Embedded Systems", 2019, vol. 6, n° 1, 53 [*DOI :* 10.4230/LITES-V006-I001-A004], https://hal.inria.fr/hal-02158985

[8] R. DAVIS, L. CUCU-GROSJEAN.*A Survey of Probabilistic Timing Analysis Techniques for Real-Time Systems*, in "Leibniz Transactions on Embedded Systems", 2019, vol. 6, n° 1, 60 [*DOI :* 10.4230/LITES-V006-I001-A003], https://hal.inria.fr/hal-02158973

[9] S. E. SAIDI, N. PERNET, Y. SOREL.*A method for parallel scheduling of multi-rate co-simulation on multi-core platforms*, in "Oil & Gas Science and Technology - Revue d'IFP Energies nouvelles", 2019, vol. 74, 49 [*DOI :* 10.2516/OGST/2019009], https://hal-ifp.archives-ouvertes.fr/hal-02141707

International Conferences with Proceedings

[10] S. BEN-AMOR, L. CUCU-GROSJEAN, D. MAXIM.*Worst-case response time analysis for partitioned fixed-priority DAG tasks on identical processors*, in "ETFA 2019 - 24th IEEE International Conference on Emerging Technologies and Factory Automation", Zaragoza, Spain, IEEE, 2019, p. 1423-1426 [*DOI :* 10.1109/ETFA.2019.8869147], https://hal.inria.fr/hal-02407007

[11] R. MEDINA, L. CUCU-GROSJEAN.*Work-in-Progress: System-wide DVFS for real-time systems with probabilistic parameters*, in "Real-Time System Symposium", York, United Kingdom, February 2020, https://hal.inria.fr/hal-02421157

[12] E. NTARYAMIRA, C. MAXIM, L. CUCU-GROSJEAN.*Data consistency and temporal validity under the circular buffer communication paradigm*, in "RACS '19 - Conference on Research in Adaptive and Convergent Systems", Chongqing, China, ACM Press, 2019, p. 51-56 [*DOI :* 10.1145/3338840.3355682], https://hal.inria.fr/hal-02409672

Conferences without Proceedings

[13] E. NTARYAMIRA, C. MAXIM, L. CUCU-GROSJEAN.*The temporal correlation of data in a multirate system*, in "RTNS'2019 - 27th International Conference on Real-Time Networks and Systems", Toulouse, France, November 2019, https://hal.archives-ouvertes.fr/hal-02362858

Research Reports

[14] Y. SOREL, F. NDOYE, W. TALABOULMA, M. MEZOUAK, L. CUCU-GROSJEAN.*Time Triggered Offline Scheduling of Data Dependent Real-Time Tasks Accounting for the Preemption and Scheduler Cost*, Inria Paris, December 2019, n^{o} RR-9318, https://hal.inria.fr/hal-02425501

References in notes

[15] R. BAHETI, H. GILL.*Cyber-physical systems*, IEEE, 2011

[16] S. BEN-AMOR, D. MAXIM, L. CUCU-GROSJEAN.*Schedulability analysis of dependent probabilistic real-time tasks*, in "the 24th International Conference on Real-Time Networks and Systems (RTNS)", 2016

[17] T. BOURKE, J. COLAÇO, B. PAGANO, C. PASTEUR, M. POUZET.*A Synchronous-Based Code Generator for Explicit Hybrid Systems Languages*, in "Compiler Construction - 24th International Conference, CC, Joint with ETAPS", 2015, p. 69–88

[18] S. COLES.*An introduction to statistical modeling of extreme values*, Springer, 2001

[19] L. CUCU.*Preliminary results for introducing dependent random variables in stochastic feasiblity analysis on CAN*, in "the WIP session of the 7th IEEE International Workshop on Factory Communication Systems (WFCS)", 2008

[20] L. CUCU-GROSJEAN, L. SANTINELLI, M. HOUSTON, C. LO, T. VARDANEGA, L. KOSMIDIS, J. ABELLA, E. MEZZETI, E. QUINONES, F. CAZORLA.*Measurement-Based Probabilistic Timing Analysis for Multi-path Programs*, in "the 24th Euromicro Conference on Real-time Systems (ECRTS)", 2012

[21] J. FONSECA, G. NELISSEN, V. NELIS, L. PINHO.*Response time analysis of sporadic DAG tasks under partitioned scheduling*, in "11th IEEE Symposium on Industrial Embedded Systems (SIES)", 05 2016, p. 1-10

[22] A. GOGONEL, L. CUCU-GROSJEAN.*Dispositif de caractérisation et/ou de modélisation de temps d'exécution pire-cas, brevet 1000408053*, INPI, 2017, https://hal.archives-ouvertes.fr/hal-01666535

[23] E. LEE, S. SESHIA.*Introduction to embedded systems - a cyber-physical systems approach*, MIT Press, 2017

[24] J. LEHOCZKY.*Real-Time Queueing Theory*, in "the 10th IEEE Real-Time Systems Symposium (RTSS)", 1996

[25] S. B. M. BERTOGNA, G. BUTTAZZO.*Multiprocessor Scheduling for Real-Time Systems*, Springer, 2015

[26] D. MAXIM, O. BUFFET, L. SANTINELLI, L. CUCU-GROSJEAN, R. I. DAVIS.*Optimal Priority Assignment Algorithms for Probabilistic Real-Time Systems*, in "the 19th International Conference on Real-Time and Network Systems (RTNS)", 2011

[27] D. MAXIM, L. CUCU-GROSJEAN.*Response Time Analysis for Fixed-Priority Tasks with Multiple Probabilistic Parameters*, in "the IEEE Real-Time Systems Symposium (RTSS)", 2013

[28] F. NDOYE, Y. SOREL.*Monoprocessor Real-Time Scheduling of Data Dependent Tasks with Exact Preemption Cost for Embedded Systems*, in "the 16th IEEE International Conference on Computational Science and Engieering (CSE)", 2013

[29] S. E. SAIDI, N. PERNET, Y. SOREL.*Automatic Parallelization of Multi-Rate FMI-based Co-Simulation On Multi-core*, in "the Symposium on Theory of Modeling & Simulation: DEVS Integrative M&S Symposium", 2017

[30] R. WILHELM, J. ENGBLOM, A. ERMEDAHL, N. HOLSTI, S. THESING, D. WHALLEY, G. BERNAT, C. FERDINAND, R. HECKMANN, T. MITRA, F. MUELLER, I. PUAUT, P. PUSCHNER, G. STASCHULAT, P. STENSTRÖEM.*The worst-case execution time problem: overview of methods and survey of tools*, in "Trans. on Embedded Computing Systems", 2008, vol. 7, n^{o} 3, p. 1-53

Project-Team MAMBA

Modelling and Analysis for Medical and Biological Applications

IN COLLABORATION WITH: Laboratoire Jacques-Louis Lions (LJLL)

IN PARTNERSHIP WITH:
CNRS

Sorbonne Université (UPMC)

RESEARCH CENTER
Paris

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team MAMBA

Creation of the Team: 2014 January 01, updated into Project-Team: 2015 April 01

Keywords:

Computer Science and Digital Science:

A3. - Data and knowledge
A3.1. - Data
A3.1.1. - Modeling, representation
A3.4. - Machine learning and statistics
A3.4.6. - Neural networks
A3.4.7. - Kernel methods
A6. - Modeling, simulation and control
A6.1. - Methods in mathematical modeling
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A6.1.4. - Multiscale modeling
A6.1.5. - Multiphysics 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
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A6.3. - Computation-data interaction
A6.3.1. - Inverse problems
A6.3.2. - Data assimilation
A6.4. - Automatic control
A6.4.1. - Deterministic control
A6.4.4. - Stability and Stabilization
A6.4.6. - Optimal control

Other Research Topics and Application Domains:

B1. - Life sciences
B1.1. - Biology
B1.1.2. - Molecular and cellular biology
B1.1.5. - Immunology
B1.1.6. - Evolutionnary biology
B1.1.7. - Bioinformatics
B1.1.8. - Mathematical biology
B1.2. - Neuroscience and cognitive science
B2. - Health
B2.2. - Physiology and diseases

B2.2.3. - Cancer
B2.2.4. - Infectious diseases, Virology
B2.2.6. - Neurodegenerative diseases
B2.3. - Epidemiology
B2.4. - Therapies
B2.4.1. - Pharmaco kinetics and dynamics
B2.4.2. - Drug resistance
B2.6.3. - Biological Imaging
B9.6.4. - Management science

1. Team, Visitors, External Collaborators

Research Scientists

Marie Doumic [Team leader, Inria, Senior Researcher, HDR]
Pierre-Alexandre Bliman [Inria, Senior Researcher, HDR]
Jean Clairambault [Inria, Emeritus, HDR]
Dirk Drasdo [Inria, Senior Researcher, HDR]
Luis Lopes Neves de Almeida [CNRS, Senior Researcher, HDR]
Diane Peurichard [Inria, Researcher]
Nastassia Pouradier Duteil [Inria, Researcher, from Sep 2018]
Philippe Robert [Inria, Senior Researcher, HDR]

Faculty Members

Ayman Moussa [Univ Pierre et Marie Curie, Associate Professor, until Aug 2019]
Benoît Perthame [Sorbonne Université, Professor, HDR]

External Collaborators

Jorge Hernandez [Univ Paris-Nord, from Feb 2019]
Jieling Zhao [IfADo-Leibniz-Institute Dortmund (Germany)]

Technical Staff

Florian Joly [Inria, Engineer, until Apr 2019]
Paul Van Liedekerke [IfADo Leibniz Institute, Dortmund (Germany), Engineer]

PhD Students

Emma Leschiera [Sorbonne Université, from October 2019]
Noemi David [Sorbonne Université, from October 2019]
Giorgia Ciavolella [Sorbonne Université, from October 2019]
Jesus Bellver Arnau [Inria, PhD Student, from Oct 2019]
Federica Bubba [Sorbonne Université, PhD Student]
Valeria Caliaro [Inria, PhD Student, from Oct 2019]
Noemi David [Inria, PhD Student, from Oct 2019]
Julia Delacour [Ecole Normale Supérieure Lyon, PhD Student]
Cecile Della Valle [Inria, PhD Student]
Adrien Ellis [Sorbonne Université, PhD Student, from Oct 2019]
Hugo Martin [Sorbonne Université, PhD Student, until Sep 2019]
Mathieu Mezache [Inria, PhD Student]
Anais Rat [Ecole centrale de Marseille, PhD Student, from Oct 2019]
Gaetan Vignoud [Ecole Nationale Supérieure des Mines de Paris , PhD Student]
Adrien Ellis [Inria, from Sept 2019]
Alexandre Poulain [Sorbonne Université, ERC ADORA]
Pastor Emmanuel Perez-Estigarribia [Polytechnic School, National University of Asuncion (Paraguay)]

Post-Doctoral Fellows

Gissell Estrada Rodriguez [Sorbonne Université, from October 2019]
Markus Schmidtchen [Sorbonne Université, from October 2019]
Cécile Carrère [Sorbonne Université, Post-Doctoral Fellow, until Oct 2019]
Jules Dichamp [Post-Doctoral Fellow, with IfADo Leibniz Institute, Dortmund (Germany)]
Sophie Hecht [Inria, Post-Doctoral Fellow, from Nov 2019]
Florian Joly [Inria, Post-Doctoral Fellow, from May 2019 until Oct 2019]
Xinran Ruan [Sorbonne Université, Post-Doctoral Fellow]

Administrative Assistants

Anna Bednarik [Inria, Administrative Assistant, from Jul 2019]
Andree Nicole Chamroux [Inria, Administrative Assistant, until Aug 2019]
Derya Gok [Inria, Administrative Assistant, until Nov 2019]

2. Overall Objectives

2.1. Overall Objectives

The MAMBA (Modelling and Analysis in Medical and Biological Applications) team is the continuation of the BANG (Biophysics, Numerical Analysis and Geophysics) team, which itself was a continuation of the former project-team M3N. Historically, the BANG team, headed by Benoît Perthame during 11 years (2003-2013), has developed models, simulations and numerical algorithms for problems involving dynamics of Partial Differential Equations (PDEs).

The dynamics of complex physical or biophysical phenomena involves many agents, e.g. proteins or cells, which can be seen as active agents. Mathematically, they can be represented either explicitly as individuals with their dynamics modelled e.g. through branching trees and piecewise deterministic Markov processes (PDMP), or stochastic differential equations, or under certain conditions be grouped or locally averaged, in which case their dynamics is mimicked by Ordinary or Partial Differential Equations (ODEs/PDEs).

Biology and medicine presently face the difficulty to make sense of the data newly available by means of recent signal acquisition methods and to take appropriate actions through possible treatment pathways. Modeling through agent-based or continuous models is a unique way to explain (model) the observations and then compute, control and predict the consequences of the mechanisms under study. These are the overall goals of Mamba.

3. Research Program

3.1. Introduction

Data and image analysis, statistical, ODEs, PDEs, and agent-based approaches are used either individually or in combination, with a strong focus on PDE analysis and agent-based approaches. Mamba was created in January 2014. It aims at developing models, simulations, numerical and control algorithms to solve questions from life sciences involving dynamics of phenomena encountered in biological systems such as protein intra-cellular spatio-temporal dynamics, cell motion, early embryonic development, multicellular growth, wound healing and liver regeneration, cancer evolution, healthy and tumor growth control by pharmaceuticals, protein polymerization occurring in neurodegenerative disorders, control of dengue epidemics, etc.

Another guideline of our project is to remain close to the most recent questions of experimental biology or medicine.In this context, we develop many close and fruitful collaborations with biologists and physicians, among which we can quote: the collaboration with St Antoine Hospital in Paris within the Institut Universitaire de Cancérologie of Sorbonne Université (IUC, Luis Almeida, Jean Clairambault, Dirk Drasdo, Benoît Perthame); Institut Jacques Monod (Luis Almeida); INRA Jouy-en-Josas (VIM team, headed by Human Rezaei and Vincent Béringue (Marie Doumic anc Philippe Robert); Wei-Feng Xue's team in the university of Canterbury (Marie Doumic and Philippe Robert); our collaborators within the HTE program (François Delhommeau at St Antoine, Thierry Jaffredo, and Delphine Salort at IBPS, Sorbonne Université, Paris; François Vallette at INSERM Nantes); Frédéric Thomas at CREEC, Montpellier; Hôpital Paul Brousse through ANR-IFlow and ANR-iLite; Institut de Biologie Physico-Chimique (IBPC, Paris, Teresa Teixeira's team; Marie Doumic); the close experimental collaborations that emerged through the former associated team QUANTISS (Dirk Drasdo), particularly at the Leibniz Institute for Working Environment and Human Factors in Dortmund, Germany; Yves Dumont at CIRAD, Montpellier.

We focus mainly on the creation, investigation and transfer of new mathematical models, methods of analysis and control, and numerical algorithms, but in selected cases software development as that of CellSys and TiQuant by D. Drasdo and S. Hoehme is performed. More frequently, the team develops "proof of concept" numerical codes in order to test the adequacy of our models to experimental biology.

We have organized the presentation of our research program in three methodological axes (Subsections 3.2, 3.3 and 3.4) and two application axes (Subsections 4.2 and 4.3). Evolving along their own logic in close interaction with the methodological axes, the application axes are considered as application-driven research axes in themselves. The methodological research axes are the following.

Axis 1 is devoted to work in physiologically-based design, analysis and control of population dynamics. It encompasses populations of bacteria, of yeasts, of cancer cells, of neurons, of aggregating proteins, etc. whose dynamics are represented by partial differential equations (PDEs), structured in evolving physiological traits, such as age, size, size-increment, time elapsed since last firing (neurons).

Axis 2 is devoted to reaction equations and motion equations of agents in living systems. It aims at describing biological phenomena such as tumor growth, chemotaxis and wound healing.

Axis 3 tackles the question of model and parameter identification, combining stochastic and deterministic approaches and inverse problem methods in nonlocal and multi-scale models.

3.2. Methodological axis 1: analysis and control for population dynamics

Personnel Pierre-Alexandre Bliman, Jean Clairambault, Marie Doumic, Benoît Perthame, Diane Peurichard, Nastassia Pouradier Duteil, Philippe Robert

Project-team positioning

Population dynamics is a field with varied and wide applications, many of them being in the core of MAMBA interests - cancer, bacterial growth, protein aggregation. Their theoretical study also brings a qualitative understanding on the interplay between individual growth, propagation and reproduction in such populations. In the past decades, many results where obtained in the BANG team on the asymptotic and qualitative behavior of such structured population equations, see e.g. [135], [73], [99], [84]. Other Inria teams interested by this domain are Mycenae, Numed and Dracula, with which we are in close contacts. Among the leaders of the domain abroad, we can cite among others our colleagues Tom Banks (USA), Graeme Wake (New Zealand), Glenn Webb (USA), Jacek Banasiak (South Africa), Odo Diekmann (Netherlands), with whom we are also in regular contact. Most remarkably and recently, connections have also been made with probabilists working on Piecewise Deterministic Markov Processes (F. Malrieu at the university of Rennes, Jean Bertoin at the ETH in Zurich, Vincent Bansaye at Ecole Polytechnique, Julien Berestycki at Cambridge, Amaury Lambert at College de France, M. Hoffmann at Paris Dauphine, Alex Watson in UCL, London and J. Bertoin in Zurich), leading to a better understanding of the links between both types of results – see also the Methodological axis 3.

Scientific achievements

We divide this research axis, which relies on the study of structured population equations, according to four different applications, bringing their own mathematical questions, e.g., stability, control, or blow-up.

Time asymptotics for nucleation, growth and division equations
Following the many results obtained in the BANG team on the asymptotic and qualitative behavior of structured population equation, we put our effort on the investigation of limit cases, where the trend to a steady state or to a steady exponential growth described by the first eigenvector fails to happen. In [78], the case of equal mitosis (division into two equally-sized offspring) with linear growth rate was studied, and strangely enough, it appeared that the general relative entropy method could also be adapted to such a non-dissipative case. Many discussions and common workshops with probabilists, especially through the ANR project PIECE coordinated by F. Malrieu, have led both communities to work closer.

In [96], the case of constant fragmentation rate and linear growth rate has been investigated in a deterministic approach, whereas similar questions were simultaneously raised but in a stochastic process approach in [75].

We also enriched the models by taking into account a nucleation term, modeling the spontaneous formation of large polymers out of monomers [147]. We investigated the interplay between four processes: nucleation, polymerization, depolymerization and fragmentation.

New perspectives are now to consider not only one species but several interacting ones, which may exhibit complex interplays which may lead to damped oscillations or to infinite growth; these are in collaboration with C. Schmeiser and within the Vienna associated team MaMoCeMa (J. Delacour's Ph.D) and with K. Fellner from Graz (M. Mezache's Ph.D).

Cell population dynamics and its control
One of the important incentives for such model design, source of many theoretical works, is the challenging question of drug-induced drug resistance in cancer cell populations, described in more detail below in the Applicative axis 1, Cancer. The adaptive dynamics setting used consists of phenotype-structured integro-differential [or reaction-diffusion, when phenotype instability is added under the form of a Laplacian] equations describing the dynamic behavior of different cell populations interacting in a Lotka-Volterra-like manner that represents common growth limitation due to scarcity of expansion space and nutrients. The phenotype structure allows us to analyse the evolution in phenotypic traits of the populations under study and its asymptotics for two populations [128], [125], [124], [126]. Space may be added as a complementary structure variable provided that something is known of the (Cartesian) geometry of the population [127], which is seldom the case.

Modelling Mendelian and non-Mendelian inheritances in density-dependent population dynamics
Classical strategies for controlling mosquitoes responsible of vector-borne disease are based on mechanical methods, such as elimination of oviposition sites; and chemical methods, such as insecticide spraying. Long term usage of the latter generates resistance [81], [110], transmitted to progeny according to Mendelian inheritance (in which each parent contributes randomly one of two possible alleles for a trait). New control strategies involve biological methods such as genetic control, which may either reduces mosquito population in a specific area or decreases the mosquito vector competence [61], [120], [156]. Among the latter, infection of wild populations by the bacterium *Wolbachia* appears promising (see also Applicative axis 2 below). Being maternally-transmitted, the latter obeys non-Mendelian inheritance law. Motivated by the effects of the (possibly unwanted) interaction of these two types of treatment, we initiated the study of modelling of Mendelian and non-Mendelian inheritances in density-dependent population dynamics. First results are shown in [59].

Control of collective dynamics

The term *self-organization* is used to describe the emergence of complex organizational patterns from simple interaction rules in collective dynamics systems. Such systems are valuable tools to model various biological systems or opinion dynamics, whether it be the collective movement of animal groups, the organization of cells in an organism or the evolution of opinions in a large crowd. A special case of self-organization is given by *consensus*, i.e. the situation in which all agents' state variables converge. Another phenomenon is that of *clustering*, when the group is split into clusters that each converge to a different state. We have designed optimal control strategies to drive collective dynamics to consensus. In the case where consensus and clustering are situations to be avoided (for example in crowd dynamics), we designed control strategies to keep the system away from clustering.

Models of neural network
Mean field limits have been proposed by biophysicists in order to describe neural networks based on physiological models. The various resulting equations are called integrate-and-fire, time elapsed models, voltage-conductance models. Their specific nonlinearities and the blow-up phenomena make their originality which has led to develop specific mathematical analysis [138], followed by [134], [119], [139], [83]. This field also yields a beautiful illustration for the capacity of the team to combine and compare stochastic and PDE modelling (see Methodological axis 3), in [89].

Models of interacting particle systems
The organisation of biological tissues during development is accompanied by the formation of sharp borders between distinct cell populations. The maintenance of this cell segregation is key in adult tissue homeostatis, and its disruption can lead tumor cells to spread and form metastasis. This segregation is challenged during tissue growth and morphogenesis due to the high mobility of many cells that can lead to intermingling. Therefore, understanding the mechanisms involved in the generation and maintain of cell segregation is of tremendous importance in tissue morphogenesis, homeostasis, and in the development of various invasive diseases such as tumors. In this research axis, we aim to provide a mathematical framework which enables to quantitatively link the segregation and border sharpening ability of the tissue to these cell-cell interaction phenomena of interest [72]. As agent-based models do not enable precise mathematical analysis of their solutions due to the lack of theoretical results, we turn towards continuous -macroscopic- models and aim to provide a rigorous link between the different models [71].

Models of population dynamics structured in phenotype
The collaboration of Jean Clairambault with Emmanuel Trélat and Camille Pouchol (from September this year assistant professor at MAP5 Paris-Descartes, University of Paris), together now with Nastassia Pouradier Duteil, has been continued and presently leads us to a possible quantitative biological identification of the structuring phenotypes of the model developed in [146], through a beginning collaboration with an Indian systems biologist (Mohit Kumar Jolly, IIS Bangalore). Our motivation in this collaboration is to couple a physiologically based system of 6 ODEs developed by our Indian collaborator with our phenotype-structured cell population dynamics model [13], [45].

In the framework of the HTE project EcoAML 2016-2020, Thanh Nam Nguyen, Jean Clairambault, Delphine Salort and Benoît Perthame, in collaboration with Thierry Jaffredo at IBPS-SU, have designed a phenotype-structured integrodifferential model of interactions between haematopoietic stem cells (healthy or leukaemic) and their supporting stromal cells [24]. In this model, without diffusion, to our relative astonishment, our postdoctoral fellow T.N. Nguyen predicts in particular that under special circumstances, a coexistence between healthy and leukaemic stem cell subpopulations is possible.The explanation of such possible theoretical coexistence still remains to be explained.

The idea of cooperation between cell subpopulations in a tumour is also studied using phenoytype-structured models of cell populations by Frank Ernesto Alvarez Borges, PhD student of Stéphane Mischler (Paris-Dauphine University), Mariano Rodríguez Ricard (University of Havana, Cuba) and Jean Clairambault, in collaboration with José Antonio Carrillo (Imperial College London). A feature of these models, in as much as conflicting continuous phenotypes (e.g., adhesivity vs. motility, or fecundity vs. viability, or fecundity vs. motility [0]) are supposed to structure a unique cell population, is that they can also represent the emergence of multicellularity in such a cell population, when two subpopulations of the same population, i.e., endowed with the same genome and represented w.r.t. relevant heterogeneity in the cell population by such conflicting phenotypes, are determined by two different choices of the 2-d phenotype. In a simplified representation when the two phenotypes are just extreme values of a 1-d continuous phenotype (e.g., 0 for total adhesivity and no motility, 1 for no adhesivity and complete motility) this situation may be related to the previously described case, developed in [24], in which two extreme values of a convex function linked to proliferation are occupied by the two extreme phenotype values (0 and 1), leading to the coexistence of two cell subpopulations.

Collaborations

- Nucleation, growth and fragmentation equations: **Klemens Fellner**, university of Graz, Austria, **Piotr Gwiązda**, Polish Academy of Sciences, Poland, **Christian Schmeiser**, university of Vienna.
- Cell population dynamics and its control: **Tommaso Lorenzi**, former Mamba postdoc, now at the University of St. Andrews, Scotland, maintains a vivid collaboration with the Mamba team. He is in particular an external member of the HTE program MoGlImaging (see also Applicative axis 1). **Emmanuel Trélat**, Sorbonne Université professor, member of LJLL and of the CAGE Inria team, is the closest Mamba collaborator for optimal control. **Benedetto Piccoli**, Professor at Rutgers University (Camden, New Jersey), is collaborating on the analysis and control of collective dynamics.
- Mendelian inheritance and resistance in density-dependent population dynamics: **Pastor Pérez-Estigarribia, Christian Schaerer**, Universidad Nacional de Asunción, Paraguay.
- Neural networks: **Delphine Salort,** Professor Sorbonne Université, Laboratory for computations and quantification in biology, and **Patricia Reynaud**, University of Nice, **Maria Cáceres**, University of Granada.
- Models of interacting particle systems: **Pierre Degond**, Imperial College London, Julien Barré, **MAPMO, Orléans**, **Ewelina Zatorska**, University College London

3.3. Methodological axis 2: reaction and motion equations for living systems

Personnel

Luis Almeida, Benoît Perthame, Diane Peurichard, Nastassia Pouradier Duteil, Dirk Drasdo

Project-team positioning

The Mamba team had initiated and is a leader on the works developed in this research axis. It is a part of a consortium of several mathematicians in France through the ANR Blanc project *Kibord*, which involves in particular members from others Inria team (DRACULA, COMMEDIA). Finally, we mention that from Sept. 2017 on, Mamba benefited from the ERC Advanced Grant ADORA (Asymptotic approach to spatial and dynamical organizations) of Benoît Perthame.

Scientific achievements

We divide this research axis, which relies on the study of partial differential equations for space and time organisation of biological populations, according to various applications using the same type of mathematical formalisms and methodologies: asymptotic analysis, weak solutions, numerical algorithms.

Aggregation equation

In the mathematical study of collective behavior, an important class of models is given by the aggregation equation. In the presence of a non-smooth interaction potential, solutions of such systems may blow up in finite time. To overcome this difficulty, we have defined weak measure-valued solutions in the sense of duality and its equivalence with gradient flows and entropy solutions in one dimension [117]. The extension to higher dimensions has been studied in [87]. An interesting consequence of this approach is the possibility to use the traditional finite volume approach to design numerical schemes able to capture the good behavior of such weak measure-valued solutions [109], [116].

[0] as proposed by John Maynard Keynes and Eös Száthmary in their book "The major transitions in evolution" (OUP 1995) as a condition of the emergence of multicellularity under environmental pressure

Identification of the mechanisms of single cell motion
In this research axis, we aim to study the mechanisms of single cell adhesion-based and adhesion free motion. This work is done in the frame of the recently created associated team MaMoCeMa (see Section 9) with the WPI, Vienna. In a first direction [150] with N. Sfakianakis (Heidelberg University), we extended the live-cell motility Filament Based Lamellipodium Model to incorporate the forces exerted on the lamellipodium of the cells due to cell-cell collision and cadherin induced cell-cell adhesion. We took into account the nature of these forces via physical and biological constraints and modelling assumptions. We investigated the effect these new components had in the migration and morphology of the cells through particular experiments. We exhibit moreover the similarities between our simulated cells and HeLa cancer cells.

In a second work done in collaboration with the group of biologist at IST (led by **Michael Sixt** Austria), we developed and analyzed a two-dimensional mathematical model for cells migrating without adhesion capabilities [118]. Cells are represented by their cortex, which is modelled as an elastic curve, subject to an internal pressure force. Net polymerization or depolymerization in the cortex is modelled via local addition or removal of material, driving a cortical flow. The model takes the form of a fully nonlinear degenerate parabolic system. An existence analysis is carried out by adapting ideas from the theory of gradient flows. Numerical simulations show that these simple rules can account for the behavior observed in experiments, suggesting a possible mechanical mechanism for adhesion-independent motility.

Free boundary problems for tumor growth
Fluid dynamic equations are now commonly used to describe tumor growth with two main classes of models: those which describe tumor growth through the dynamics of the density of tumoral cells subjected to a mechanical stress; those describing the tumor through the dynamics of its geometrical domain thanks to a Hele-Shaw-type free boundary model. The first link between these two classes of models has been rigorously obtained thanks to an incompressible limit in [137] for a simple model. This result has motivated the use of another strategy based on viscosity solutions, leading to similar results, in [121].

Since more realistic systems are used in the analysis of medical images, we have extended these studies to include active motion of cells in [136], viscosity in [141] and proved regularity results in [129]. The limiting Hele-Shaw free boundary model has been used to describe mathematically the invasion capacity of a tumour by looking for travelling wave solutions, in [140], see also Methodological axis 3. It is a fundamental but difficult issue to explain rigorously the emergence of instabilities in the direction transversal to the wave propagation. For a simplified model, a complete explanation is obtained in [122].

Two-way coupling of diffusion and growth
We are currently developing a mathematical framework for diffusion equations on time-evolving manifolds, where the evolution of the manifold is a function of the distribution of the diffusing quantity. The need for such a framework takes it roots in developmental biology. Indeed, the growth of an organism is triggered by signaling molecules called morphogens that diffuse in the organism during its development. Meanwhile, the diffusion of the morphogens is itself affected by the changes in shape and size of the organism. In other words, there is a complete coupling between the diffusion of the morphogens and the evolution of the shapes. In addition to the elaboration of this theoretical framework, we also collaborate with a team of developmental biologists from Rutgers University (Camden, New Jersey) to develop a model for the diffusion of Gurken during the oogenesis of *Drosophila*.

Migration of cells in extracellular matrix
A single cell based model has been developed that reproduces a large set of experimental observations of cells migrating in extracellular matrix based on physical mechanisms with mimimal internal cell dynamics. This includes individually migrating cells in micro-channels of different size, and their collective dynamics in case of many cells, as well as the impact of cell division and growth. The model explicitly mimics the extracellular matrix as the cells as deformable objects with explicit filopodia.

Collaborations

- Shanghai Jiao Tong University, joint publications with Min Tang on bacterial models for chemotaxis and free boundary problems for tumor growth.
- Imperial College London, joint works with José Antonio Carrillo on aggregation equation.
- University of Maryland at College Park, UCLA, Univ. of Chicago, Univ. Autónoma de Madrid, Univ. of St. Andrews (Scotland), joint works on mathematics of tumor growth models.
- Joint work with Francesco Rossi (Università di Padova, Italy) and Benedetto Piccoli (Rutgers University, Camden, New Jersey, USA) on Developmental PDEs.
- Cooperation with Shugo Yasuda (University of Hyogo, Kobe, Japan) and Vincent Calvez (EPI Dracula) on the subject of bacterial motion.
- Cooperation with Nathalie Ferrand (INSERM), Michèle Sabbah (INSERM) and Guillaume Vidal (Centre de Recherche Paul Pascal, Bordeaux) on cell aggregation by chemotaxis.
- Nicolas Vauchelet, Université Paris 13

3.4. Methodological axis 3: Model and parameter identification combining stochastic and deterministic approaches in nonlocal and multi-scale models

Personnel

Marie Doumic, Dirk Drasdo.

Project-team positioning

Mamba developed and addressed model and parameter identification methods and strategies in a number of mathematical and computational model applications including growth and fragmentation processes emerging in bacterial growth and protein misfolding, in liver regeneration [103], TRAIL treatment of HeLa cells [74], growth of multicellular spheroids [115], blood detoxification after drug-induced liver damage [149], [107].

This naturally led to increasingly combine methods from various fields: image analysis, statistics, probability, numerical analysis, PDEs, ODEs, agent-based modeling methods, involving inverse methods as well as direct model and model parameter identification in biological and biomedical applications. Model types comprise agent-based simulations for which Mamba is among the leading international groups, and Pharmocokinetic (PK) simulations that have recently combined in integrated models (PhD theses Géraldine Cellière, Noémie Boissier). The challenges related with the methodological variability has led to very fruitful collaborations with internationally renowned specialists of these fields, e.g. for bacterial growth and protein misfolding with Marc Hoffmann (Paris Dauphine) and Patricia Reynaud-Bouret (University of Nice) in statistics, with Tom Banks (Raleigh, USA) and Philippe Moireau (Inria M3DISIM) in inverse problems and data assimilation, and with numerous experimentalists.

Scientific achievements

Direct parameter identification is a great challenge particularly in living systems in which part of parameters at a certain level are under control of processes at smaller scales.

Estimation methods for growing and dividing populations

In this domain, all originated in two papers in collaboration with J.P. Zubelli in 2007 [143], [101], whose central idea was to used the asymptotic steady distribution of the individuals to estimate the division rate. A series of papers improved and extended these first results while keeping the deterministic viewpoint, lastly [78]. The last developments now tackle the still more involved problem of estimating not only the division rate but also the fragmentation kernel (i.e., how the sizes of the offspring are related to the size of the dividing individual) [97]. In parallel, in a long-run collaboration with statisticians, we studied the Piecewise Deterministic Markov Process (PDMP) underlying the equation, and estimated the division rate directly on sample observations of the process, thus making a bridge between the PDE and the PDMP approach in [100], a work which inspired also very recently other groups in statistics and probability [75], [112] and was the basis for Adélaïde Olivier's Ph.D thesis [132], [114] and of some of her more recent works [133][46] (see also axis 5).

Data assimilation and stochastic modeling for protein aggregation

Estimating reaction rates and size distributions of protein polymers is an important step for understanding the mechanisms of protein misfolding and aggregation (see also axis 5). In [63], we settled a framework problem when the experimental measurements consist in the time-dynamics of a moment of the population.

To model the intrinsic variability among experimental curves in aggregation kinetics - an important and poorly understood phenomenon - Sarah Eugène's Ph.D, co-supervised by P. Robert [105], was devoted to the stochastic modeling and analysis of protein aggregation, compared both with the deterministic approach traditionally developed in Mamba [147] and with experiments.

Collaborations

- **Marc Hoffmann,** Université Paris-Dauphine, for the statistical approach to growth and division processes [100], **M. Escobedo**, Bilbao and **M. Tournus**, Marseille, for the deterministic approach.
- **Philippe Moireau,** Inria M3DISIM, for the inverse problem and data assimilation aspects [69], [62]

4. Application Domains

4.1. Introduction

The team has two main application-driven research axes. Applicative axis 1 focuses on cancer, an application on which almost all team members work, with various approaches. A main focus of the team is to study cancer as a Darwinian evolutionary phenomenon in phenotype-structured cell populations. Optimal control methods take into account the two main pitfalls of clinical cancer therapeutics, namely unwanted toxic side effects in healthy cell populations and drug resistance in cancer cell populations. Other studies concern telomere shortening, and multi-scale models. Applicative axis 2 is devoted to growth, evolution and regeneration in populations and tissues. It involves protein aggregation and fragmentation models for neurodegenerative diseases (prion, Alzheimer), organ modeling, mainly of the liver, its damages induced by toxic molecules, and its regeneration after toxic insult. Newcomers in this applicative field are epidemiological modeling of propagation of insect vector-borne diseases by reaction-diffusion equations and of their optimal control, bacterial growth and wound healing.

4.2. Applicative axis 1: Focus on cancer

Personnel

Luis Almeida, Jean Clairambault, Marie Doumic, Dirk Drasdo, Benoît Perthame, Diane Peurichard, Nastassia Pouradier Duteil.

Project-team positioning

The MAMBA team designs and analyses mathematical models of tumor growth and therapy, at the cell population level, using agent-based or partial differential equations, with special interest in methodologies for therapeutic optimization using combined anticancer drug treatments. Rather than, or not only, modeling the effect of drugs on molecular targets, we represent these effects by their functional consequences on the fate of healthy and cancer cell populations: proliferation (velocity of the cell division cycle, decreasing it, e.g., by antagonizing growth factor receptors), apoptosis, cell death or senescence. Our goal in doing this is to circumvent the two main issues of anticancer therapy in the clinic, namely unwanted toxic side effects in populations of healthy cells and emergence of drug-induced drug resistance in cancer cell populations. This point of view leads us to take into account phenomena of transient and reversible resistance, observed in many cancer cell populations, by designing and analyzing models of cell populations structured in continuous phenotypes, relevant for the description of the behavior of cell populations exposed to drugs:

either degree of resistance to a given drug, or potential of resistance to drug-induced stress, proliferation potential, and plasticity. Such modeling options naturally lead us to to take into account in a continuous way (i.e., by continuous-valued phenotype or relevant gene expression) the wide phenotypic heterogeneity of cancer cell populations. They also lead us to adopt the point of view of adaptive dynamics according to which characteristic traits of cell populations evolve with tumor environmental pressure (drugs, cytokines or metabolic conditions, mechanical stress and spatial conditions), in particular from drug sensitivity to resistance. This position is original on the international scene of teams dealing with drug resistance in cancer.

Scientific achievements

Modeling Acute Myeloid Leukemia (AML) and its control by anticancer drugs by PDEs and Delay Differential equations
In collaboration with Catherine Bonnet (Inria DISCO, Saclay) and François Delhommeau (St Antoine hospital in Paris), together with DISCO PhD students José Luis Avila Alonso and Walid Djema, this theme has led to common published proceedings of conferences: IFAC, ACC, CDC, MTNS [66], [67], [68], [77], [95], [65]. These works study the stability of the haematopoietic system and its possible restabilization by combinations of anticancer drugs with functional targets on cell populations: proliferation, apoptosis, differentiation.

Adaptive dynamics setting to model and circumvent evolution towards drug resistance in cancer by optimal control
We tackle the problem to represent and inhibit - using optimal control algorithms, in collaboration with Emmanuel Trélat, proposed Inria team CAGE - drug-induced drug resistance in cancer cell populations. This theme, presently at the core of our works on cancer modeling with a evolutionary perspective on tumor heterogeneity, is documented in a series of articles [90], [92], [124], [125], [127]. Taking into account the two main pitfalls of cancer therapy, unwanted side effects on healthy cells and evolution towards resistance in cancer cells, it has attracted to our team the interest of several teams of biologists, with whom we have undertaken common collaborative works, funded by laureate answers to national calls (see ITMO Cancer HTE call).

This theme is also at the origin of methodological developments (see Research axis 1). In collaboration with Shensi Shen from Institut Gustave Roussy and Francois Vallette from Université de Nantes, we aim to develop simple non-spatial models to understand the mechanisms of drug resistance acquisition -and loss- in melanoma and glioblastoma. The models are systematically compared with in vitro and in vivo data generated by our collaborators and treated via image processing techniques developed in the team.

Senescence modeling by telomere shortening
In many animals, aging tissues accumulate senescent cells, a process which is beneficial to protect from cancer in the young organism. In collaboration with Teresa Teixeira and Zhou Xu from IBCP, we proposed a mathematical model based on the molecular mechanisms of telomere replication and shortening and fitted it on individual lineages of senescent Saccharomyces cerevisiae cells, in order to decipher the causes of heterogeneity in replicative senescence [79].

Biomechanically mediated growth control of cancer cells
Model simulations indicate that the response of growing cell populations on mechanical stress follows a simple universal functional relationship and is predictable over different cell lines and growth conditions despite the response curves look largely different. We developed a hybrid model strategy in which cells were represented by coarse-grained individual units calibrated in a high resolution cell model and parameterized each model cell by measurable biophysical and cell-biological parameters. Cell cycle progression in our model is controlled by volumetric strain, the latter being derived from a bio-mechanical relation between applied pressure and cell compressibility. After parameter calibration from experiments with mouse colon carcinoma cells growing against the resistance of an elastic alginate capsule, the model adequately predicts the growth curve in i) soft and rigid capsules, ii) in different experimental conditions where the mechanical stress is generated by osmosis via a high molecular weight dextran solution, and iii) for other cell types with different growth kinetics. Our model simulation results suggest that the growth response of cell population upon externally applied mechanical stress is the same, as it can be quantitatively predicted using the same growth progression function [123].

Bio-mechanical models of tissue growth
The degenerate Cahn-Hilliard equation is a standard model to describe living tissues. It takes into account cell populations undergoing short-range attraction and long-range repulsion effects. In this framework, we consider the usual Cahn-Hilliard equation with a singular single-well potential and degenerate mobility. These degeneracy and singularity induce numerous difficulties, in particular for its numerical simulation. To overcome these issues, we propose in [hal-02274417] a relaxation system formed of two second order equations which can be solved with standard packages. This system is endowed with an energy and an entropy structure compatible with the limiting equation. Here, we study the theoretical properties of this system; global existence and convergence of the relaxed system to the degenerate Cahn-Hilliard equation. We also study the long-time asymptotics which interest relies on the numerous possible steady states with given mass.

Free boundary multiphase models of tumor growth
Multiphase mechanical models are now commonly used to describe living tissues including tumour growth. The specific model we study here consists of two equations of mixed parabolic and hyperbolic type which extend the standard compressible porous media equation, including cross-reaction terms. We study the incompressible limit, when the pressure becomes stiff, which generates a free boundary problem. We establish the complementarity relation and also a segregation result. Several major mathematical difficulties arise in the two species case which are addressed in [43]. Firstly, the system structure makes comparison principles fail. Secondly, segregation and internal layers limit the regularity available on some quantities to BV. Thirdly, the Aronson-Bénilan estimates cannot be established in our context. We are lead, as it is classical, to add correction terms. This procedure requires technical manipulations based on BV estimates only valid in one space dimension. Another novelty is to establish an L^1 version in place of the standard upper bound.

Philosophy of cancer
The quite natural idea that cancer is a disease of the control of coherent multicellularity, expressed when cohesion of tissues and coherence of (unknown, except maybe for the case of a centralised circadian clock) synchronising signals fail to ensure it, by a regression towards unicellularity, stopping in this "reverse evolution path" at a coarse, incoherent multicellularity state [0] continues to be developed and popularised by Jean Clairambault in seminars and workshops, and published in review articles [13], [45]. This view, and the investigation of the immune system in the design of such coherence of all multicellular organisms [0] is naturally inscribed in a *philosophy of cancer* perspective, and from a mathematical viewpoint, to multicellularity genes - and links between them and unicellularity genes - seen as a *hyperstructure* [0] above structures consisting of the genes of unicellularity, i.e., those that make a single cell a coherent living system, such hyperstructure being failed in cancer; this view is presently under development with colleagues from universities of the Paris region, together with Nils Baas at NTNU, Trondheim, Norway). This perspective, that makes use of category theory as a structuring point of view to apprehend multicellularity and cancer, is also meant to endow us with an innovative methodology to apply *topological data analysis (TDA)* to investigate cancer genome data.

Modelling of TMZ induced drug resistance
Temozolomide (TMZ) is a standard chemotherapy treatment in patients with glioblastoma. Resistance to this drug is correlated to the presence of a specific enzyme, which activity in cancer cells creates a drug-induced cell death resistant phenotype. Understanding the transition of cancer cells to a resistant phenotype is still a topic of research where multiple hypothesis have been studied: From an adaptive process to an inherent resistance to treatment. Moreover it has been recently shown that TMZ treatment has an influence on the spatial structuration of cancer cell aggregates. In the frame of the HTE project MoGlImaging and through the recent hiring of a post-doctoral candidate (Gissell Estrada Rodriguez), we are currently developping a mathematical framework to study and analyse the evolution of a population of glioblastoma cells that are exposed to TMZ. Based on the experimental data generated by our partner team led by F. Valette (Inserm Nantes), we propose a Keller-Segel type model for the formation of spheroid as a result of a chemoattractant produced by cancer

[0] Metazoa 1.0, as theorised by PCW Davies and CH Lineweaver in their article "Cancer tumors as Metazoa 1.0: tapping genes of ancient ancestors", Physical Biology 2011, that popularised the so-called atavistic hypothesis of cancer

[0] this latter point partly, however nicely, developed in Thomas Pradeu's book "The limits of the self", OUP 2012

[0] See on this point, e.g., Nils Baas: "On the philosophy of higher structures", Int. J. General Systems 2019

cells, and study the influence of a chemotherapeutic agent on the structuration of the cancer cell aggregates. By confronting the model results to experimental data, different modelling choices are currently explored to identify which key mechanisms could be responsible for the apparition of drug resistance in glioblastoma.

Collaborations

- AML modelling: **Catherine Bonnet**, DISCO Inria team, Saclay, and **François Delhommeau**, INSERM St Antoine (also collaborator in the INSERM HTE laureate project EcoAML, see below).
- INSERM HTE laureate project MoGlImaging, headed by E. Moyal (Toulouse): **François Vallette**, CRCNA and INSERM Nantes
- INSERM HTE laureate project EcoAML, headed by **François Delhommeau**, INSERM St Antoine: François Delhommeau, Thierry Jaffredo (IBPS), Delphine Salort (LCQB-IBPS)
- Adaptive dynamics to model drug resistance and optimal control to circumvent it:

  Alexandre Escargueil, Michèle Sabbah (1 PhD thesis in common), St Antoine Hospital, Paris

  Emmanuel Trélat (1 PhD thesis in common) at Inria team CAGE and Laboratoire Jacques-Louis Lions at Sorbonne Université.

  Frédéric Thomas at CREEC, Montpellier.

  Tommaso Lorenzi (Univ. of St Andrews).
- Telomere shortening: **Teresa Teixeira** and **Zhou Xu** (IBCP, Paris), **Philippe Robert** (Inria RAP).
- TRAIL treatment: **Gregory Batt,** Inria Saclay and Inst. Pasteur (France)
- Biomechanical control of cancer cells: **Pierre Nassoy**, Bioimaging and Optofluidics Group, LP2N – UMR 5298. IOGS, CNRS & University of Bordeaux

4.3. Applicative axis 2: Growth, evolution and regeneration in populations and tissues

Luis Almeida, Pierre-Alexandre Bliman, Marie Doumic, Dirk Drasdo, Benoît Perthame, Diane Peurichard, Nastassia Pouradier Duteil, Philippe Robert

Project-team positioning

The applications in this category span very different subjects from amyloid diseases, dengue fever, wound healing, liver regeneration and toxicity, up to bacterial growth and development of organisms. As the applications, the methods span a wide range. Those concerning identification of models and parameters with regard to data have partially been outlined in axis 3. Focus in this axis is on the model contribution to the biologically and/or medically relevant insights and aspects.
Liver-related modelling is partially performed within the Inria team MIMESIS (Strasbourg) with the focus on real-time, patient-specific biomechanical liver models to guide surgery and surgeons. Internationally, spatial temporal liver related models are developed in Fraunhofer MEVIS (Bremen), by T. Ricken (TU Dortmund), and P. Segers group (Leuven).

Different from these, Mamba has a strong focus on spatial-temporal modeling on the histological scale, integration of molecular processes in each individual cell, and single-cell (agent) based models [102]. Works by Schliess [149], [107] have been highlighted in editorials.

Mathematical modeling of protein aggregation is a relatively recent domain, only a few other groups have emerged yet; among them we can cite the Inria team Dracula, with whom we are in close contact, and e.g., the work by Jean-Michel Coron (Sorbonne Université) and Monique Chyba (Hawaii, USA) in control, and Suzanne Sindi (USA) for the modeling of the yeast prion. We have interactions with all these groups and organized a workshop in June 2017, gathering both the biophysics and applied mathematics communities.

Scientific achievements

Amyloid disease

Application to protein aggregation in amyloid diseases is a long-standing interest of Mamba, dating back to 2010 [85], and developed through the collaboration with Human Rezaei's team at Inra. More recently, with Wei-Feng Xue in Canterbury, we investigated the intrinsic variability among identical experiments of nucleation [98], [106], Sarah Eugène's Ph.D subject (co-supervised by Philippe Robert) [105].

In collaboration with Tom Banks first [70], [69] and then Philippe Moireau, we developed quantitative comparisons between model and data. Through data assimilation and statistical methods [63], we proposed new models and mechanisms.

Biological control of arboviroses
Sterile Insect Technique (SIT) [104] is a biological control method relying on massive releases of sterile male insects into the wild. The latter compete with wild males to mate with the females, and induce no offspring to the latter, thus reducing the next generation's population. This can result in a progressive reduction, or even disparition, of the target population.

A related technique is based on the infection by *Wolbachia* [111]. This symbiotic bacterium is maternally transmitted from infected females to their offspring, but induces *cytoplasmic incompatibility* [151], [80]: mating between infected males and uninfected females gives no offspring. Releases of *Wolbachia* infected males alone is thus comparable to classical SIT.

On the other hand, releasing both infected males and females in sufficient quantity may result in infection of the wild population. This gives rise to an interesting new control principle, as *Wolbachia* has been shown to severely reduce the insect vectorial ability to transmit dengue, zika or chikungunya, indirectly by lifespan and fertility reduction, and directly by reducing the ability of the viruses to proliferate within the organism [130].

We proposed new insights on the practical and theoretical issues raised by the implementation of the previous methods. Concerning the SIT, we obtained control synthesis results through impulsive periodic release of controlled amplitude [10], and through optimal control approach [42]. Concerning Wolbachia technique, we investigated general control principles [39] capable of spreading the infection.

Wound healing 1: epithelial tissues
We studied cell motion in epithelial gap closure, a form of collective cell migration that is a very widespread phenomenon both during development and adult life - it is essential for both the formation and for the maintenance of epithelial layers. Due to their importance, *in vivo* wound healing and morphogenetic movements involving closure of holes in epithelia have been the object of many studies. In our works [0] we considered wound healing and epithelial gap closure in both in vivo (in particular drosophila pupa) and in vitro (MDCK cell and human keratinocytes). We found some similarities in the geometry dependence of the wound closure strategies between these two situations, indicating the existence of conserved mechanisms that should be widespread across living beings. We are concentrating on the study of actin cable formation.

Wound healing 2: adipose tissues
After injury, if regeneration can be observed in hydra, planaria and some vertebrates, regeneration is rare in mammals and particularly in humans. In this research axis, we investigated the mechanisms by which biological tissues recover after injury. We explored this question on adipose tissue, using the mathematical framework recently developed in [144]. Our assumption is that simple mechanical cues between the Extra-Cellular Matrix (ECM) and differentiated cells can explain adipose tissue morphogenesis and that regeneration requires after injury the same mechanisms. We validated this hypothesis by means of a two-dimensional Individual Based Model (IBM) of interacting adipocytes and ECM fiber elements [145]. The model successfully generated regeneration or scar formation as functions of few key parameters, and seemed to indicate that the fate of injury outcome could be mainly due to ECM rigidity.

[0] ravasio:hal-01245750, vedula:hal-01298859

Following these encouraging results, the team is currently taking a step further in the model validation and confrontation to experimental data. The first direction concerns the development of a 3D framework to validate the mechanisms observed in 2D. In collaboration with S. Merino-Aceituno from the University of Vienna, efforts are being made in the development of a complete synthetic tissue model coupling the ECM and cell modelling with a vascularization model.A PhD project has been started to implement the coupled models and reduce the CPU time with the long-term goal to develop a usable software which would serve to investigate the role of different mechanisms in tissue development (not restricted to adipose tissues). Finally, further developments in collaboration with Imperial College London aim at pursuing the derivation of macroscopic PDE models from the agent-based formalisms.

Mathematical modelling of axolotl regeneration

Tissue response after injury/amputation induces one or two alternatives: scar formation versus regeneration (complete recovery of tissue shape and functions). In most mammals, regeneration is considered largely impaired for the benefit of a fibrotic scar after injury automatically associated with dysfunctions, but complete regeneration has been largely described and investigated in animal models such as zebra fish, salamander, or axolotl. Despite several processes regulating regeneration have been identified at different scales -from diffusing molecules and cellular gene expression patterns up to tissue mechanics-, how these mechanisms individually or collectively play a role in the regulation of regenerative processes remains poorly understood. In order to give insights into the mechanisms of tissue regeneration, Valeria Caliaro started an Inria PhD project in october 2019, in collaboration with Osvaldo Chara, internationally recognized group leader of SysBio in Argentina. This project focuses on the role of cell proliferation in space and time along the two first phases of regeneration after injury: (i) initiation of a regeneration response, (ii) tissue patterning during regenerate growth. The first part of the project aims at building an agent-based model featuring few key mechanisms regulating cell proliferation after injury. The model construction is based on recent works where the authors developed a mathematical model given by ordinary differential equations (ODEs)[2] and a mathematical framework in 1D [3] showing that acceleration of the cell cycle is the major driver of regenerative spinal cord outgrowth in axolotls. Building on both mathematical models and introducing heuristic rules which rely on Prof O. Chara expertise, we propose a 2D-ABM using methodologies borrowed from socio-dynamics and collective behavior studies (based on many interacting agent systems). While the focus is made on proliferation-based mechanisms, other mechanisms responsible for collective behavior such as volume exclusion, diffusion or aggregation will be tested and compared with experimental data. The resulting model will provide a synthetic tissue model which will serve to investigate regeneration in cellular systems, focusing on cell proliferation properties. The second part of the PhD will be devoted to the derivation of continuous models from the agent-based formalism. This will provide a large scale ‘synthetic tissue’ model to explore the role of large scale effects in general tissue models. Model validation and calibration will be ensured by quantitative comparison with biological data already present in the literature and generated by the SysBio group of O. Chara, particularly the representative images of regenerative spinal cords after tail amputation. By varying the model parameters and observing the resulting alteration of the spinal cord size and architecture as a consequence of these variations, it will be possible to provide ‘in silico’ setting experiments to guide and plan future in vivo or ex vivo experiments. Altogether, the project is expected to provide a mechanistic understanding of the cellular mechanisms driving spinal cord regeneration, and to identify how spatial structuration can influence cell differentiation and growth.

[1] Rodrigo Albors A, Tazaki A, Rost F, Nowoshilow S, Chara O & Tanaka EM. 2015. Planar cell polarity-mediated induction of neural stem cell expansion during axolotl spinal cord regeneration. Elife. 4:e10230. [2] Rost F, Rodrigo Albors A, Mazurov V, Brusch L, Deutsch A, Tanaka EM & Chara O. 2016. Accelerated cell divisions drive the outgrowth of the regenerating spinal cord in axolotls. Elife. 5. pii: e20357. [3] Cura Costa E, Rodrigo Albors A, Tanaka EM & Chara O. Spatiotemporal distribution of cell proliferation rules spinal cord regeneration in the axolotl. MS in preparation.

Quantitative cell-based model predicts mechanical stress response of growing tumor spheroids

Model simulations indicate that the response of growing cell populations on mechanical stress follows the same functional relationship and is predictable over different cell lines and growth conditions despite experimental response curves look largely different. We developed a hybrid model strategy in which cells are represented by coarse-grained individual units calibrated with a high resolution cell model and parameterized by measurable biophysical and cell-biological parameters. Cell cycle progression in our model is controlled by volumetric strain, the latter being derived from a bio-mechanical relation between applied pressure and cell compressibility. After parameter calibration from experiments with mouse colon carcinoma cells growing against the resistance of an elastic alginate capsule, the model adequately predicts the growth curve in i) soft and rigid capsules, ii) in different experimental conditions where the mechanical stress is generated by osmosis via a high molecular weight dextran solution, and iii) for other cell types with different growth kinetics from the growth kinetics in absence of external stress. Our model simulation results suggest a generic, even quantitatively same, growth response of cell populations upon externally applied mechanical stress, as it can be quantitatively predicted using the same growth progression function ([52])

Modeling of morphogen diffusion in *Drosophila* oogenesis
In collaboration with a team of developmental biologists of Rutgers University (Camden, New Jersey), we have built a model for the diffusion of the Gurken morphogen during *Drosophila* oogenesis, taking into account a wide variety of biological mechanisms such as diffusion of the morphogen, reactions of components of the EGFR signaling pathway, movement of the source of morphogen, shift of the overlying follicle cells and growth of the egg chamber. This model, together with a complete numerical code developed in Matlab, provides a tool to understand how each mechanism influences the signal distribution. The overall aim of the project is to use this tool to guide future experiments, and to understand what mechanisms contribute to the different distributions of signal among species.

Bacterial population growth
We exploited all the methods developed to estimate the division rate of a population (see axis 3) to address a seminal question of biology: is it a size-sensing or a timing mechanism which triggers bacterial growth? In [148], we showed that a sizer model is robust and fits the data well. Several studies from other groups came at the same time, showing a renewed interest on a question dated back to Jacques Monod's PhD thesis (1941). Of special interest is the "adder" model, for which we are currently developing new estimation methods [46].

A quantitative high resolution computational mechanics cell model for growing and regenerating tissues
Mathematical models are increasingly designed to guide experiments in biology, biotechnology, as well as to assist in medical decision making. They are in particular important to understand emergent collective cell behavior. For this purpose, the models, despite still abstractions of reality, need to be quantitative in all aspects relevant for the question of interest. Considered was as showcase example the regeneration of liver after drug-induced depletion of hepatocytes, in which the surviving and dividing hepatocytes must squeeze in between the blood vessels of a network to refill the emerged lesions. Here, the cells' response to mechanical stress might significantly impact the regeneration process. We present a 3D high-resolution cell-based model integrating information from measurements in order to obtain a refined and quantitative understanding of the impact of cell-biomechanical effects on the closure of drug-induced lesions in liver. Our model represents each cell individually and is constructed by a discrete, physically scalable network of viscoelastic elements, capable of mimicking realistic cell deformation and supplying information at subcellular scales. The cells have the capability to migrate, grow, and divide, and the nature and parameters of their mechanical elements can be inferred from comparisons with optical stretcher experiments. Due to triangulation of the cell surface, interactions of cells with arbitrarily shaped (triangulated) structures such as blood vessels can be captured naturally. Comparing our simulations with those of so-called center-based models, in which cells have a largely rigid shape and forces are exerted between cell centers, we find that the migration forces a cell needs to exert on its environment to close a tissue lesion, is much smaller than predicted by center-based models. To stress generality of the approach, the liver simulations were complemented by monolayer and multicellular spheroid growth simulations. In summary, our model can give quantitative insight in many tissue organization processes, permits hypothesis testing in silico, and guide experiments in situations in which cell mechanics is considered important [123].

Liver regeneration and disease: towards a full virtual liver model at histological scale

In our work towards a full virtual liver model at histological level, a number of steps were performed. The models under points (1)-(4) focus on either a single or a few liver lobules. A liver lobule is the smallest repetitive functional and anatomical building block of liver, while (5) addresses a much larger organisational building block of the liver, a liver lobe that consists of thousands to hundreds of thousands of lobules depending on the species. A second strand (6), (7) addresses image analysis, which in most cases forms the entrance to modeling as it provides the data necessary to generate model hypotheses and to parameterize a model.

(1) Cell types: In a former work by Hoehme et. al. ([113]) a model of liver regeneration after drug-induced damage was established considering hepatocytes and blood vessels. This model has now been expanded to include all relevant cell types, including hepatocytes, blood vessels, hepatic stellate cells, Kupffer cells, invading macrophages and other immune cells. Thereby it is now possible to study perturbations in the temporal scenario of damage and regeneration after signaling events or cells types are knocked down individually or collectively. This model is currently compared to respective perturbation experiments.

(2) Liver disease: Degenerative liver diseases such as liver fibrosis and cirrhosis develop out of a disturbed balance of degenerative and regenerative processes. The model under (1) has thereby been extended by the formation of extracellular matrix, mimicked as fiber networks, to capture the disease process leading to liver fibrosis. In that process characteristic streets form that modify the mechanics, perfusion behavior and detoxification capacity of the liver.

(3) Consequence of liver fibrosis: Whole-slide scans from fibrotic liver in a mouse model has been analysed at different time points after emergence of the disease with regard to the degree of excess matrix to mimic the possible consequences of fibrotic inclusions on perfusion and function of liver within a multiscale model that considers ammonia detoxification in each individual hepatocyte as well as blood flow and transport processes in the liver lobule.

(4) Bile flux: Bile flux has been for decades believed to be controlled by convection at the level of liver lobules as well as at the level of the entire organ. By a methodology based on correlative imaging for quantitative intravital flux analysis no directed advection was detectable in bile canaliculi at the resolution limit. Instead, after active transport across hepatocyte membranes bile salts within the liver lobules are transported in the canaliculi by a diffusion-dominated process. Only in the interlobular ducts i.e., at super-lobular level, diffusion is augmented by advection. In silico simulations of bile transport in real 3D bile network microarchitectures can quantitatively explain the data assuming diffusive transport as sole mechanism.

(5) Liver regeneration after partial hepatectomy (partial organ removal): Partial hepatectomy is an adequate therapy in case of diseases or events that destructed only part of the liver. A typical case is a primary tumor or a metastasis affecting only a single liver lobe. Within an biophysical agent-based model capturing many aspects of the cell mechanics we studied regrowth of liver after partial organ removal in mouse calibrated with multivariate experimental data. Our model predicts characteristic proliferation pattern that change from small animals (as mouse) to large animals (as pig).

(6) Bile duct ligation: Bile duct ligation (BDL) is an experimental procedure that mimics obstructive cholestatic disease. One of the early consequences of BDL in rodents is the appearance of so-called bile infarcts that correspond to Charcot-Gombault necrosis in human cholestasis. The mechanisms causing bile infarcts and their pathophysiological relevance are unclear. Therefore, intravital two photon–based imaging of BDL mice was performed with fluorescent bile salts (BS) and non-BS organic anion analogues. Key findings were followed up by matrix-assisted laser desorption ionization imaging, clinical chemistry, immunostaining, and gene expression analyses. Our group performed analysis of intravital imaging. The key finding is that bile microinfarcts occur in the acute phase after BDL in a limited number of dispersed hepatocytes followed by larger infarcts involving neighboring hepatocytes, and they allow leakage of bile from the BS-overloaded biliary tract into blood, thereby protecting the liver from BS toxicity; in the chronic phase after BDL, reduced sinusoidal BS uptake is a dominant protective factor, and the kidney contributes to the elimination of BS until cholemic nephropathy sets in [108].

(7) Periportalisation during liver fibrosis formation: Within a liver lobule, the function of hepatocytes is zonated i.e., certain functions are only executed by either hepatocytes close to the center (pericentral region) or hepatocytes in the periphery of the lobule (periportal region). Little is known about how liver fibrosis influences lobular zonation. To address this question, three mouse models of liver fibrosis were used, CCl4 administration repeated for 2, 6 and 12 months to induce pericentral damage, as well as bile duct ligation (21 days) and a particular mdr2-mouse model to study periportal fibrosis. Analyses were performed by RNA-sequencing, immunostaining of zonated proteins and image analysis. Image analysis was performed by our group. The key result was that liver fibrosis leads to strong alterations of lobular zonation, where the pericentral region adopts periportal features. Beside adverse consequences, periportalization supports adaptation to repeated doses of hepatotoxic compounds [17].

Toxicity extrapolation from in vitro to in vivo

In vivo toxicity prediction from in vitro data is a major objective in toxicology as it permits bypassing animal experiments, and as the predictive power of animal experiments for human is limited. Objective was the prediction of paracetamol (acetaminophen)-induced hepatotoxicity from in vitro experiments. For this purpose, numerous iterations between in vitro experiments, in vivo experiments and simulations were performed for mouse. Using a recent thesis (Géraldine Cellière'ns PhD thesis [88]) as a start point, two candidate mechanisms could be identified both explaining the in vivo data after calibration of the in silico model with in vitro toxicity data.

Relating imaging on microscopic scales with imaging on macroscopic scales: From Diffusion-Weighted MRI Calibrated With Histological Data: an Example From Lung Cancer

Diffusion-weighted magnetic resonance imaging (DWI) is a key non-invasive imaging technique for cancer diagnosis and tumor treatment assessment, reflecting Brownian movement of water molecules in tissues. Since densely packed cells restrict molecule mobility, tumor tissues produce usually higher signal (less attenuated signal) on isotropic maps compared with normal tissues. However, no general quantitative relation between DWI data and the cell density has been established. In order to link low-resolution clinical cross-sectional data with high resolution histological information, we developed an image processing and analysis chain, which was used to study the correlation between the diffusion coefficient (D value) estimated from DWI and tumor cellularity from serial histological slides of a resected non-small cell lung cancer tumor. Color deconvolution followed by cell nuclei segmentation was performed on digitized histological images to determine local and cell-type specific 2d (two-dimensional) densities. From these, the 3d cell density was inferred by a model-based sampling technique, which is necessary for the calculation of local and global 3d tumor cell count. Next, DWI sequence information was overlaid with high resolution CT data and the resected histology using prominent anatomical hallmarks for co-registration of histology tissue blocks and non-invasive imaging modalities' data. The integration of cell numbers information and DWI data derived from different tumor areas revealed a clear negative correlation between cell density and D value. Importantly, spatial tumor cell density can be calculated based on DWI data. In summary, our results demonstrate that tumor cell count and heterogeneity can be predicted from DWI data, which may open new opportunities for personalized diagnosis and therapy optimization [157]. The work of that paper has been further advanced to adapt the procedures for clinical use (in preparation).

Collaborations

- Biological control of arboviroses: **Nicolas Vauchelet** (Université Paris 13); **Grégoire Nadin** (LJLL, Sorbonne Université); **Yannick Privat** (Université de Strasbourg); **D. Villela, C. Struchiner** (Fiocruz, Brazil); **Jorge Zubelli** (IMPA, Brazil); **Alain Rapaport** (INRA-Montpellier), **Y. Dumont** (CIRAD-Montpellier); **Ch. Schaerer, P. Pérez-Estigarribia** (UNA, Paraguay), **O. Vasilieva** (Universidad del Valle, Cali, Colombia), **D. Cardona-Salgado** (Universidad Autónoma de Occidente, Cali, Colombia).
- Protein aggregation in amyloid diseases: **Human Rezaei**'s team at Inra Jouy-en-Josas (France) and **W-F Xue**'s team in at university of Kent (Great Britain); **Tom Banks** at the North Carolina State University (USA) and **Philippe Moireau** (M3DISIM)
- Bacterial growth and division: **Lydia Robert**, Sorbonne Université (France)

- Liver research & toxicology: **JG. Hengstler** group (IfADo, Dortmund, Germany); **R. Gebhardt** (Univ. Leipzig); **U. Klingmueller** (DKFZ, Heidelberg); **Irène Vignon-Clementel** (Inria, COMMEDIA)
- Growth in capsules and biomechanics: **Pierre Nassoy**, Institut dOptique Graduate School, Talence, France; **Josef Kaes**, Peter Debye Institute for Soft Matter Physics, Physics, Univ. Leipzig, Germany.
- Wound healing: **Patrizia Bagnerini** (Genova, Numerical methods), **Benoît Ladoux** (Institut Jacques Monod et Mechanobiology Institute Singapore, Biophysics) and **Antonio Jacinto** (CEDOC, Lisbon, Biology and Medicine). (Adipose tissue regeneration) team of **L. Casteilla** (StromaLab, Toulouse). (Axolotl regeneration) team of O. Chara, SysBio group, Argentina.
- Diffusion of morphogen: Center for Computational and Integrative Biology, Rutgers University (Camden, New Jersey), joint work with Professor Nir Yakoby's Drosophila Laboratory
- Linking micro and macro-image information: Oliver Sedlaczek, Univ. and DKFZ Heidelberg, Kai Breuhahn, Univ. Heidelberg.

5. Highlights of the Year

5.1. Highlights of the Year

- Marie Doumic gave a plenary talk at the AIP Conference (Applied Inverse Problems) in Grenoble, July 8-12th (around 600 participants).
- Diane Peurichard and Nastassia Pouradier Duteil won a Mittag-Leffler and EWS-EMS Call to organize a summer school at the Mittag-Leffler institute in July 2020.
- the STIC AmSud cooperative project NEMBICA, between France, Chile, Paraguay and Colombia, headed by Pierre-Alexandre Bliman, has been accepted (2020-2021).

6. New Software and Platforms

6.1. TiQuant

Tissue Quantifier

KEYWORDS: Systems Biology - Bioinformatics - Biology - Physiology

FUNCTIONAL DESCRIPTION: Systems biology and medicine on histological scales require quantification of images from histological image modalities such as confocal laser scanning or bright field microscopy. The latter can be used to calibrate the initial state of a mathematical model, and to evaluate its explanatory value, which hitherto has been little recognized. We generated a software for image analysis of histological material and demonstrated its use in analysing liver confocal micrografts, called TiQuant (Tissue Quantifier). The software is part of an analysis chain detailing protocols of imaging, image processing and analysis in liver tissue, permitting 3D reconstructions of liver lobules down to a resolution of less than a micrometer.

- Author: Dirk Drasdo
- Contact: Dirk Drasdo

6.2. TiSim

Tissue Simulator

KEYWORDS: Systems Biology - Bioinformatics - Biology - Physiology

SCIENTIFIC DESCRIPTION: TiSim (Tissue Simulator) is a versatile and efficient simulation environment for tissue models. TiSim is a software for agent-based models of multicellular systems. It permits model development with center-based models and deformable cell models, it contains modules for monolayer and multicellular spheroid simulations as well as for simulations of liver lobules. Besides agent-based simulations, the flow of blood and the transport of molecules can be modelled in the extracellular space, intracellular processes such as signal transduction and metabolism can be simulated, for example over an interface permitting integration of SBML-formulated ODE models. TiSim is written in modern C++ , keeping central model constituents in modules to be able to reuse them as building blocks for new models. For user interaction, the GUI Framework Qt is used in combination with OpenGL for visualisation. The simulation code is in the process of being published. The modeling strategy and approaches slowly reach systems medicine and toxicology. The diffusion of software is a fundamental component as it provides the models that are complex and difficult to implement (implementing a liver lobule model from scratch takes about 2-2.5yrs) in form of a software to the developer and users who like to build upon them. This increases significantly the speed of implementing new models. Moreover, standardization is indispensible as it permits coupling different software tools that may have implemented models at different scales / levels.

FUNCTIONAL DESCRIPTION: TiSim is a software that permits agent-based simulations of multicellular systems. - center-based lattice-free agent-based model - modular - C++, Qt, OpenGL, GUI, batch mode - permits multiscale simulations by integration of molecular pathways (for signaling, metabolisms, drug) into each individual cell - applications so far: monolayer growth, multicellular spheroids - Boolean networks (development time = coding time (60 MMs) + model development time (264 MMs)) - in follow-up version 1: - liver lobule regeneration - SBML interface - in follow-up version 2: - deformable cell model (by triangulation of cell surface) - deformable rod models - extracellular matrix - vascular flow and transport TiSim can be directly fed by processed image data from TiQuant.

- Participants: Andreas Buttenschoen, Dirk Drasdo, Eugenio Lella, Géraldine Cellière, Johannes Neitsch, Margaretha Palm, Nick Jagiella, Noémie Boissier, Paul van Liedekerke, Stefan Hoehme and Tim Johann
- Partner: IZBI, Université de Leipzig
- Contact: Dirk Drasdo

6.3. Platforms

6.3.1. TiSim

The deformable cell model [123], [52] has been integrated in addition to the center-based model in the software TiSim (Tissue Simulator), a follow-up of former CellSys [113]. Center-based models of cells represent forces between cells as forces between cell centers but lacks an explicit representation of cell shape. The deformable cell model represents cell shape explicitly. Applications are monolayers, multicellular spheroids and simulations of liver regeneration, whereby intracellular pathways can be integrated. The model shall be distributed as binary and will permit to use the deformable cell model to calibrate intercellular forces at high cell densities, where the two-body force models so far applied in center-based models fail.

6.3.2. TiQuant

This image processing and analysis software ([103]) now integrates a machine learning component. This is fundamental as it is more general and permits quicker adaptation to new images.

7. New Results

7.1. Direct and inverse Problems in Structured-population equations

7.1.1. Modelling Polymerization Processes

In 2017, we evidenced the presence of several polymeric species by using data assimilation methods to fit experimental data from H. Rezaei's lab [64]; new experimental evidence reinforced these findings [19], [35]. The challenges are now to propose mathematical models capable of tracking such diversity while keeping sufficient simplicity to be tractable to analysis.

In collaboration with Klemens Fellner from the university of Graz, we propose a new model, variant of the Becker-Döring system but containing two monomeric species, capable of displaying sustained though damped oscillations as is experimentally observed. We also proposed a statistical test to validate or invalidate the presence of oscillations in experimental highly nonstationary signals [55].

7.1.2. Asymptotic behaviour of structured-population equations

Pierre Gabriel and Hugo Martin studied the mathematical properties of a model of cell division structured by two variables – the size and the size increment – in the case of a linear growth rate and a self-similar fragmentation kernel [16]. They first show that one can construct a solution to the related two dimensional eigenproblem associated to the eigenvalue 1 from a solution of a certain one dimensional fixed point problem. Then they prove the existence and uniqueness of this fixed point in the appropriate L^1 weighted space under general hypotheses on the division rate. Knowing such an eigenfunction proves useful as a first step in studying the long time asymptotic behaviour of the Cauchy problem.
Etienne Bernard, Marie Doumic and Pierre Gabriel proved in [9] that for the growth-fragmentation equation with fission into two equal parts and linear growth rate, under fairly general assumptions on the division rate, the solution converges towards an oscillatory function, explicitely given by the projection of the initial state on the space generated by the countable set of the dominant eigenvectors of the operator. Despite the lack of hypo-coercivity of the operator, the proof relies on a general relative entropy argument in a convenient weighted L^2 space, where well-posedness is obtained via semigroup analysis. They also propose a non-dissipative numerical scheme, able to capture the oscillations.

Pierre Gabriel and Hugo Martin then extended this asymptotic result in the framework of measure solutions [50]. To do so they adopt a duality approach, which is also well suited for proving the well-posedness when the division rate is unbounded. The main difficulty for characterizing the asymptotic behavior is to define the projection onto the subspace of periodic (rescaled) solutions. They achieve this by using the generalized relative entropy structure of the dual problem.

7.1.3. Estimating the division rate from indirect measurements of single cells

7.1.3.1. Marie Doumic and Adélaïde Olivier

Is it possible to estimate the dependence of a growing and dividing population on a given trait in the case where this trait is not directly accessible by experimental measurements, but making use of measurements of another variable? The article [46] adresses this general question for a very recent and popular model describing bacterial growth, the so-called incremental or adder model - the model studied by Hugo Martin and Pierre Gabriel in [16]. In this model, the division rate depends on the increment of size between birth and division, whereas the most accessible trait is the size itself. We prove that estimating the division rate from size measurements is possible, we state a reconstruction formula in a deterministic and then in a statistical setting, and solve numerically the problem on simulated and experimental data. Though this represents a severely ill-posed inverse problem, our numerical results prove to be satisfactory.

7.2. Stochastic Models of Biological Systems

7.2.1. Stochastic models for spike-timing dependent plasticity

7.2.1.1. Ph. Robert and G. Vignoud

Synaptic plasticity is a common mechanism used to model learning in stochastic neural networks, STDP is a great example of such mechanisms. We develop a simple framework composed by two neurons and one synaptic weight, seen as stochastic processes and study the existence and stability of such distributions, for a wide range of classical synaptic plasticity models. Using two simple examples of STDP, the calcium-based rule and the all-to-all pair-based rule, we apply stochastic averaging principles and obtain differential equations for the limit processes, based on the invariant distributions of the fast system when the slow variables are considered fixed. We study a general stochastic queue to approximate the calcium-based rule and are able to have an analytical solution for the invariant distribution of the fast synaptic processes. We also detail some simpler systems, either through some approximations or simulations to put into light the influences of different biologically-linked parameters on the dynamics of the synaptic weight.

7.2.2. Online Sequence Learning In The Striatum With Anti-Hebbian Spike-Timing-Dependent Plasticity

7.2.2.1. G. Vignoud. Collaboration with J. Touboul (Brandeis University)

Spike-Timing Dependent Plasticity (STDP) in the striatum is viewed as a substrate for procedural learning. Striatal projecting neurons (SPNs) express anti-Hebbian plasticity at corticostriatal synapses, (a presynaptic cortical spike followed by a postsynaptic striatal spike leads to the weakening of the connection, whereas the reverse pairing leads to potentiation). SPNs need to integrate many inputs to spike, and as such, their main role is to integrate context elements to choose between different sensorimotor associations. In this work, we develop a simple numerical model of the striatum, integrating cortical spiking inputs to study the role of anti-Hebbian STDP in pattern recognition and sequence learning. Cortical neurons are seen as binary input neurons and one striatal SPN is modeled as a leaky-integrate-and-fire neuron. Combined informations from the output, reward and timing between the different spikes modify the intensity of each connection, through two mechanisms: anti-Hebbian STDP and dopaminergic signaling, using three-factor learning rules. We have added a second output neuron with collateral inhibition which leads to an improvement of the global accuracy. In another project, we studied the dynamics of learning, by shutting off/on the dopaminergic plasticity, and compare it to DMS/DLS experimental and behavioral experiments. We show that anti-Hebbian STDP favors the learning of complete sequence of spikes, such as is needed in the striatum, whereas, even if Hebbian STDP helps to correlate the spiking of two connected neurons, it is not sufficient to integrate of long sequences of correlated inputs spikes.

7.2.3. D1/D2 detection from action-potential properties using machine learning approach in the dorsal striatum

7.2.3.1. G. Vignoud. Collaboration, with Team Venance (CIRB/Collège de France)

Striatal medium spiny neurons (MSNs) are segregated into two subpopulations, the D1 receptor-expressing MSNs (the direct striatonigral pathway) and the D2 receptorexpressing MSNs (the indirect striatopallidal pathway). The fundamental role of MSNs as output neurons of the striatum, and the necessary distinction between D1- and D2-expressing neurons accentuate the need to clearly distinguish both subpopulations in electrophysiological recordings in vitro and in vivo. Currently, fluorescent labelling of the dopaminergic receptors in mice enables a clear differentiation. However, multiplying in vivo the number of genetic markers (optogenetics, fluorescence) hinders possibilities for other genetic manipulations. Moreover, electrophysiological properties of fluorescents neurons can slightly differ from “native” cells and false-positive can be observed. The lack of a proper way to separate D1- and D2-MSNs based on electrophysiological properties led us to devise a detection algorithm based on action potential profile. We used more than 450 D1/D2 labelled MSNs from in vitro patch-clamp recordings (different experimentalists, different setups and protocols), to characterize and identify properties that facilitate the MSN discrimination. After analyzing passive and active MSN membrane properties, we built an extensive dataset and fed it into classical machine learning classification methods. The training of the different algorithms (k-nearest neighbors, random forest, deep neural networks, ...) was performed with the scikit-learn Python library, and the optimized classifier was able to correctly discriminate neurons in the dorsolateral striatum at 76% (and up to 83% if we allow the classifier to reject some MSNs). This study developed an efficient classification algorithm for D1/D2-MSNs, facilitating cell discrimination without specific genetic fluorescent labelling, leaving some room for other genetic markers and optogenetic labeling.

7.2.4. The Stability of Non-Linear Hawkes Processes

7.2.4.1. Ph. Robert and G. Vignoud

We have investigated the asymptotic properties of self-interacting point processes introduced by Kerstan (1964) and Hawkes and Oakes (1974). These point processes have the property that the intensity at some point $t \in (-\infty, +\infty)$ is a functional of all points of the point process before t. Such a process is said to be stable if it has a version whose distribution is invariant by translation. By using techniques of coupling and Markovian methods, we have been able to obtain some existence and uniqueness results with weaker conditions than in the current literature.

7.2.5. Mathematical Models of Gene Expression

7.2.5.1. Ph. Robert

In Robert [30] we analyze the equilibrium properties of a large class of stochastic processes describing the fundamental biological process within bacterial cells, *the production process of proteins*. Stochastic models classically used in this context to describe the time evolution of the numbers of mRNAs and proteins are presented and discussed. An extension of these models, which includes elongation phases of mRNAs and proteins, is introduced. A convergence result to equilibrium for the process associated to the number of proteins and mRNAs is proved and a representation of this equilibrium as a functional of a Poisson process in an extended state space is obtained. Explicit expressions for the first two moments of the number of mRNAs and proteins at equilibrium are derived, generalizing some classical formulas. Approximations used in the biological literature for the equilibrium distribution of the number of proteins are discussed and investigated in the light of these results. Several convergence results for the distribution of the number of proteins at equilibrium are in particular obtained under different scaling assumptions.

7.2.6. Stochastic modelling of molecular motors

7.2.6.1. Marie Doumic, Dietmar Oelz, Alex Mogilner

It is often assumed in biophysical studies that when multiple identical molecular motors interact with two parallel microtubules, the microtubules will be crosslinked and locked together. The aim of the article [4] is to examine this assumption mathematically. We model the forces and movements generated by motors with a time-continuous Markov process and find that, counter-intuitively, a tug-of-war results from opposing actions of identical motors bound to different microtubules. The model shows that many motors bound to the same microtubule generate a great force applied to a smaller number of motors bound to another microtubule, which increases detachment rate for the motors in minority, stabilizing the directional sliding. However, stochastic effects cause occasional changes of the sliding direction, which has a profound effect on the character of the long-term microtubule motility, making it effectively diffusion-like. Here, we estimate the time between the rare events of switching direction and use them to estimate the effective diffusion coefficient for the microtubule pair. Our main result is that parallel microtubules interacting with multiple identical motors are not locked together, but rather slide bidirectionally. We find explicit formulae for the time between directional switching for various motor numbers.

7.3. Analysis and control of mosquitoe populations

7.3.1. Control Strategies for Sterile Insect Techniques

We proposed different models to serve as a basis for the design of control strategies relying on releases of sterile male mosquitoes (*Aedes spp*) and aiming at elimination of wild vector population. Different types of releases were considered (constant, periodic or impulsive) and sufficient conditions to reach elimination were provided in each case [152]We also estimated sufficient and minimal treatment times. A feedback approach was introduced, in which the impulse amplitude is chosen as a function of the actual wild population [152].

7.3.2. Optimal replacement strategies, application to Wolbachia

We modelled and designed optimal release control strategy with the help of a least square problem. In a nutshell, one wants to minimize the number of uninfected mosquitoes at a given time horizon, under relevant biological constraints. We derived properties of optimal controls and studied a limit problem providing useful asymptotic properties of optimal controls [8], [42].

7.3.3. Oscillatory regimes in population models

Understanding mosquitoes life cycle is of great interest presently because of the increasing impact of vector borne diseases. Observations yields evidence of oscillations in these populations independent of seasonality, still unexplained. We proposed [33] a simple mathematical model of egg hatching enhancement by larvae which produces such oscillations that conveys a possible explanation.

On the other hand, population oscillations may be induced by seasonal changes. We considered a biological population whose environment varies periodically in time, exhibiting two very different "seasons", favorable and unfavorable. We addressed the following question: the system's period being fixed, under what conditions does there exist a critical duration above which the population cannot sustain and extincts, and below which the system converges to a unique periodic and positive solution? We obtained [153], [154] sufficient conditions for such a property to occur for monotone differential models with concave nonlinearities, and applied the obtained criterion to a two-dimensional model featuring juvenile and adult insect populations.

7.3.4. Feedback control principles for population replacement by Wolbachia

The issue of effective scheduling of the releases of *Wolbachia*-infected mosquitoes is an interesting problem for Control theory. Having in mind the important uncertainties present in the dynamics of the two populations in interaction, we attempted to identify general ideas for building release strategies, which should apply to several models and situations [39]. These principles were exemplified by two interval observer-based feedback control laws whose stabilizing properties were demonstrated when applied to a model retrieved from [76].

7.4. Bacterial motion by Rerun and tumble

Collective motion of chemotactic bacteria such as Escherichia coli relies, at the individual level, on a continuous reorientation by runs and tumbles. It has been established that the length of run is decided by a stiff response to a temporal sensing of chemical cues along the pathway. We describe a novel mechanism for pattern formation stemming from the stiffness of chemotactic response relying on a kinetic chemotaxis model which includes a recently discovered formalism for the bacterial chemotaxis [142]. We prove instability both for a microscopic description in the space-velocity space and for the macroscopic equation, a flux-limited Keller-Segel equation, which has attracted much attention recently. A remarkable property is that the unstable frequencies remain bounded, as it is the case in Turing instability. Numerical illustrations based on a powerful Monte Carlo method show that the stationary homogeneous state of population density is destabilized and periodic patterns are generated in realistic ranges of parameters. These theoretical developments are in accordance with several biological observations.

This motivates also our study of traveling wave and aggregation in population dynamics of chemotactic cells based on the FLKS model with a population growth term [86]. Our study includes both numerical and theoretical contributions. In the numerical part, we uncover a variety of solution types in the one-dimensional FLKS model additionally to standard Fisher/KPP type traveling wave. The remarkable result is a counter-intuitive backward traveling wave, where the population density initially saturated in a stable state transits toward an un- stable state in the local population dynamics. Unexpectedly, we also find that the backward traveling wave solution transits to a localized spiky solution as increasing the stiffness of chemotactic response. In the theoretical part, we obtain a novel analytic formula for the minimum traveling speed which includes the counter-balancing effect of chemotactic drift vs. reproduction/diffusion in the propagating front. The front propagation speeds of numerical results only slightly deviate from the minimum traveling speeds, except for the localized spiky solutions, even for the backward traveling waves. We also discover an analytic solution of unimodal traveling wave in the large-stiffness limit, which is certainly unstable but exists in a certain range of parameters.

Another activity concerns the relation between the tumbling rate and the internal state of bacteria. The study [58] aims at deriving models at the macroscopic scale from assumptions on the microscopic scales. In particular we are interested in comparisons between the stiffness of the response and the adaptation time. Depending on the asymptotics chosen both the standard Keller-Segel equation and the flux-limited Keller-Segel (FLKS) equation can appear. An interesting mathematical issue arises with a new type of equilibrium equation leading to solution with singularities.

7.5. Numerical methods for cell aggregation by chemotaxis

Three-dimensional cultures of cells are gaining popularity as an in vitro improvement over 2D Petri dishes. In many such experiments, cells have been found to organize in aggregates. We present new results of three-dimensional in vitro cultures of breast cancer cells exhibiting patterns. Understanding their formation is of particular interest in the context of cancer since metastases have been shown to be created by cells moving in clusters. In the paper [82], we propose that the main mechanism which leads to the emergence of patterns is chemotaxis, i.e., oriented movement of cells towards high concentration zones of a signal emitted by the cells themselves. Studying a Keller-Segel PDE system to model chemotactical auto-organization of cells, we prove that it is subject to Turing instability if a time-dependent condition holds. This result is illustrated by two-dimensional simulations of the model showing spheroidal patterns. They are qualitatively compared to the biological results and their variability is discussed both theoretically and numerically.

This motivates to study parabolic-elliptic Keller-Segel equation with sensitivity saturation, because of its pattern formation ability, is a challenge for numerical simulations. We provide two finite-volume schemes that are shown to preserve, at the discrete level, the fundamental properties of the solutions, namely energy dissipation, steady states, positivity and conservation of total mass [131]. These requirements happen to be critical when it comes to distinguishing between discrete steady states, Turing unstable transient states, numerical artifacts or approximate steady states as obtained by a simple upwind approach. These schemes are obtained either by following closely the gradient flow structure or by a proper exponential rewriting inspired by the Scharfetter-Gummel discretization. An interesting fact is that upwind is also necessary for all the expected properties to be preserved at the semi-discrete level. These schemes are extended to the fully discrete level and this leads us to tune precisely the terms according to explicit or implicit discretizations. Using some appropriate monotonicity properties (reminiscent of the maximum principle), we prove well-posedness for the scheme as well as all the other requirements. Numerical implementations and simulations illustrate the respective advantages of the three methods we compare.

7.6. Focus on cancer

Modelling Acute Myeloid Leukaemia (AML) and its control by anticancer drugs by PDEs and Delay Differential equations

This theme has continued to be developed in collaboration with Catherine Bonnet, Inria DISCO (Saclay) [93], [94]. Without control by drugs, but with representation of mutualistic interactions between tumor cells and their surrounding support stroll cells, it has also, in collaboration with Delphine Salort and Thierry Jaffredo (LCQB-IBPS) given rise to a recent work by Thanh Nam Nguyen, hired as HTE and ERC postdoctoral fellow at LCQB, submitted as full article [24].

Adaptive dynamics setting to model and circumvent evolution towards drug resistance in cancer by optimal control

The research topic “Evolution and cancer”, designed in the framework of adaptive dynamics to represent and overcome acquired drug resistance in cancer, initiated in [128], [127] and later continued in [91], [92], [126], has been recently summarised in [60] and has been the object of the PhD thesis work of Camille Pouchol, see above “Cell population dynamics and its control” . It is now oriented, thanks to work underway by Cécile Carrère, Jean Clairambault, Tommaso Lorenzi and Grégoire Nadin, in particular towards the mathematical representation of *bet hedging* in cancer, namely a supposed optimal strategy consisting for cancer cell populations under life-threatening cell stress in diversifying their phenotypes according to several resistance mechanisms, such as overexpression of ABC transporters (P-glycoprotein and many others), of

DNA repair enzymes or of intracellular detoxication processes. According to different deadly insults the cancer cell population is exposed to, some phenotypes may be selected, any such successful subpopulation being able to store the cell population genome (or subclones of it if the cell population is already genetically heterogeneous) and make it amenable to survival and renewed replication.

Philosophy of cancer biology
This new research topic in Mamba, dedicated to explore possibly underinvestigated, from the mathematical modelling point of view, parts of the field of cancer growth, evolution and therapy, has been the object of a presentation by Jean Clairambault at the recent workshop “Philosophy of cancer biology’

https://www.philinbiomed.org/event/philosophy-of-cancer-biology-workshop/.

This workshop gathered most members worldwide of this small, but very active in publishing, community of philosophers of science whose field of research is “philosophy of cancer”, as they call it themselves. This topic offers a clear point of convergence between mathematics, biology and social and human sciences.

7.7. Deformable Cell Modeling: biomechanics and Liver regeneration

- Biomechanically mediated growth control of cancer cells The key intriguing novelty was that the same agent-based model after a single parameter has been calibrated with growth data for multicellular spheroids without application of external mechanical stress by adapting a single parameter, permitted to correctly predict the growth speed of multicellular spheroids of 5 different cell lines subject of external mechanical stress. Hereby the same mechanical growth control stress function was used without any modification [123]. The prediction turned out to be correct independent of the experimental method used to exert the stress, whereby once a mechanical capsule has been used, once dextran has been used in the experiments.
- Regeneration of liver with the Deformable Cell Model. The key novelty was the implementation of the model itself, but an interesting novel result is that the DCM permits closure of a pericentral liver lobule lesion generated by drug-induced damage with about 5 times smaller active migration force due to the ability of the cell to strongly deform and squeeze into narrow spaces between the capillaries. This finding stresses that a precise mechanical description is important in view of quantitatively correct modeling results [155]. The deformable cell model however could be used to calibrate the interaction forces of the computationally much cheaper center-based model to arrive at almost the same results.

8. Partnerships and Cooperations

8.1. National Initiatives

Mamba (Marie Doumic and Philippe Robert) participates to the GDR "MeDyna" (mechanisms and dynamics of assemblies of peptides and proteins), coordinated by Stéphane Bressanelli from IBPC.

8.1.1. ANR

8.1.1.1. ANR Blanc 2014-2018 “Kibord”

This project gathers several members of the MAMBA team together with the ENS Cachan and Université Paris-Dauphine on the mathematical study of PDE models with application to biology.

8.1.1.2. ANR iLITE 2016 - 2020

Jean-Charles Duclos-Vallée, Paul Brousse Hospital, Villejuif. Partners are several departments in Paul Brousse Hospital, ENS Cachan, University of Compiègne and several companies all over France, and COMMEDIA team, Inria Paris. The pursued objective is the bioengineering design of an artificial liver intended for liver replacement.

8.1.1.3. ANR InTelo 2017-2020

Telomere dynamics, headed by Teresa Teixeira (IBPC, Paris).

8.1.1.4. INCa/DGOS; PRT-K 2018-2021

Khê HOANG-XUAN, Hôpital Universitaire La Pitié Salpêtrière, Paris. Mathematical modeling at micro and macroscopic level of primary central nervous system lymphomas (PCNSL).

8.1.2. ITMO Cancer 2016 - 2020, HTE call (heterogeneity of tumours in their ecosystems)

8.1.2.1. ITMO Cancer EcoAML

Early leukaemogenesis in Acute Myelogenous Leukaemia (AML), 8 teams headed by François Delhommeau (CDR St Antoine, Paris).

8.1.2.2. ITMO Cancer MoGlImaging

Treatment-induced treatment resistance and heterogeneity in glioblastoma, 8 teams headed by Elizabeth Moyal (INSERM, Toulouse).

8.2. International Initiatives

- **STIC AmSud 20-STIC-05**
  - Title: New Methods for Biological Control of the Arboviruses
  - International Partner (Institution - Laboratory - Researcher):

    CIRAD (Montpellier), UMR MISTEA (Montpellier), Université Paris 13, Université de Bordeaux, Université de Strasbourg, Université Paris-Dauphine - PSL; Universidad de Buenos Aires and Universidad Nacional de Salta (Argentina); Universidad de Chile (Chile); Universidad del Quindio, Universidad Autónoma de Occidente and Universidad del Valle (Colombia); National University of Asuncion (Paraguay).
  - Duration: 2020 - 2021
  - Start year: 2020
  - The main focus of this project is modeling and analysis, using mathematical methods, of new strategies aimed at controling the spread of the dengue fever and other vector-borne diseases similar to Dengue and transmitted by Aedes mosquitoes, like Chikungunya and Zika virus.
  - The key topics are the following.
    - Spatial aspects of biological control techniques
    - Estimation issues for vector-borne epidemics
    - Optimal and non-optimal control approaches for biological control techniques
    - Modelling the effects of conventional control methods on the success of biological control
    - Modelling the competition effects in larval phase during biological control
    - Modelling and efficacy measures for self-propagating genetic interventions
    - Genome-scale models for Wolbachia
- **ERC Advanced grant No 740623 ADORA**

  ADORA is the acronym for *Asymptotic approach to spatial and dynamical organizations*.

  Adora ERC project aims at understanding of spatial, social and dynamical organization of large numbers of agents, presently a fundamental issue in science. ADORA focuses on problems motivated by biology because, more than anywhere else, access to precise and numerous data has opened the route to novel and complex mathematical models. The address ed problems are written in terms of

nonlinear partial differential equations. The flux-limited Keller-Segel system, the integrate-and-fire Fokker-Planck equation, kinetic equations with internal state, nonlocal parabolic equations and constrained Hamilton-Jacobi equations are among examples of the equations under investigation.

The role of mathematics is not only to understand the analytical structure of these new problems, but it is also to **explain the qualitative behavior of solutions and to quantify their properties**. The challenge arises here because these goals should be achieved through a hierarchy of scales. Indeed, the problems under consideration share the common feature that the large scale behavior cannot be understood precisely without access to a hierarchy of finer scales, down to the individual behavior and sometimes its molecular determinants.

Major difficulties arise because the numerous scales present in these equations have to be discovered and singularities appear in the asymptotic process which yields deep compactness obstructions. Our vision is that **the complexity inherent to models of biology can be enlightened** by mathematical analysis and a classification of the possible asymptotic regimes.

However an enormous effort is needed to uncover the equations intimate mathematical structures, and bring them at the level of conceptual understanding they deserve being given the applications motivating these questions which range from medical science or neuroscience to cell biology.

8.2.1. MaMoCeMa

Title: Mathematical modeling of cell motility and of autophagy

International Partner (Institution - Laboratory - Researcher):

University of Vienna (Austria) - Wolfgang Pauli Institute - Christian Schmeiser

Start year: 2018

Numerous fruitful collaborations have been developed these last years between the WPI and the Inria team MAMBA. Diane Peurichard – newly recruited permanent member of the team MAMBA- worked two years (2016-2017) with Christian Schmeiser -member of the present project- through a post-doctoral contract at the university of Vienna. In collaboration with the biologists of IST, they developed mathematical tools to understand how cells move through adhesion-based and adhesion-free motion with applications in cancer development, prevalent theme of the team MAMBA. Collaborations WPI-MAMBA have been maintained and ensured by the sabbatical of Marie Doumic (2016-2018) -, working at the university of Vienna with Christian Schmeiser and the PhD student Julia Delacour. They have initiated a collaboration on the mathematical modeling of autophagy, which requires both C. Schmeiser's expertise in biomechanics and M. Doumic's knowledge on aggregation processes. This team will also benefit of the strong links that C. Schmeiser has developed with the two biologists teams of S. Martens (on autophagy) and M. Sixt (on cell movement).

8.2.2. Participation in Other International Programs

- **BMBF (Germany) / LiSym; 2016-2020** LiSym addresses liver diseases and regeneration, namely, steatosis, fibrosis and cirrhosis, and acute on chronic liver failure. (Dirk Drasdo)
- **BMBF (Germany) / MSDILI; 2016-2019** MS-DILI addresses multiscale modeling of drug-induced liver disease focusing on the role of APAP. Dirk Drasdo participates in this project. (Dirk Drasdo)

8.3. International Research Visitors

Visitors in Paris (LJLL) invited by J. Clairambault: Zineb Kaïd, PhD student, University Abou Bekr Belkaïd, Tlemcen, Nov. 25 - Dec. 13; Jean-François Mascari, researcher, IAC-CNR, Rome, Dec. 9-13.

Visitors in Paris (Inria) invited by D. Drasdo: Jieling Zhao, Postdoc from IfADo, Jules Dichamp, Postdoc from ifADo, Paul van Liedekerke, Research engineer from IfADo.

Visitors in Paris (LJLL) invited by D. Peurichard and M. Doumic: P. Degond (Imperial College London) Nov 20 - 22, C. Schmeiser (University of Vienna), Dec. 1 - 7, Claudia Wytrzen (University of Vienna) feb 4-8 2019, M. Tournus and M. Escobedo (February 18-22), .

Visitors in Paris (LJLL) invited by P. A. Bliman: Prof. Héctor Jairo Martínez Romero (Universidade del Valle, Cali, Colombia) for two weeks, together with Oscar Eduardo Escobar Lasso, PhD student who was present one month.

Visitors in Paris (LJLL) invited by B. Perthame: Shugo Yasuda (University of Hyogo, Kobe, Japan), Min Tang (SJTU, China), Maria Caceres (Granada, Spain), Zhenan Zhou (Peking University), Weizhu Bao (Singapore university).

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. General Chair, Scientific Chair

- J. Clairambault has organised with Stylianos Xanthopoulos (University of the Aegean), Nikolaos Sfakianakis (University of St. Andrews), Nikos Kavallaris (University of Chester) and Pierre-Alexandre Bliman the Summer school and workshop *MBMC-Samos 2019* “Mathematical Biology on the Mediterranean Conference”, Samos island, Greece, September 1-14, see http://actuarweb.aegean.gr/mbmc2019. This conference, envisioned to be the first of a series of such MBMCs, consisted of two twin events, the Summer school and the workshop (1st and 2nd week) that were funded by Inria DGDS and DPEI, together with EMS and ESMTB.
- Dirk Drasdo is member of the Program committee of the SBMC 2020 (Heidelberg, Germany), May, 25-27.
- L. Almeida, B. Perthame and D. Peurichard coordinate a math-bio work group every month at LJLL
- D. Peurichard and D. Drasdo coordinate the internal 'Open MAMBA seminar' - seminar for PhD and post-doctoral students of the MAMBA team
- Pierre-Alexandre Bliman, together with Christian Schaerer (Polytechnic School, Universidad Nacional de Asunción, Paraguay), organized the Workshop on Advanced Control Techniques for Mosquito-Borne Diseases 2019 (WTAEM 2019) 11–17 november 2019, https://sites.google.com/view/wtaem/wtaem
- Benoit Perthame organized a France-China summer school 1-5 july 2019 (partnership LIA-SFMA, Institute of Natural Science and Department of mathematics, Shanghai Jiao Tong University)

9.1.1.2. Member of the Organizing Committees

- On October 7th and 8th, Marie Doumic co-organized the symposium *Deciphering the functional mechanisms of biological macromolecules* with Raphaël Guérois (CEA) and Marc Baaden (IBPC).
- Marie Doumic was a member of the organisation committee for the 50th anniversary of J-L. Lions laboratory (November 27-29).
- Marie Doumic co-organised the Workshop "Mathematical Models in Cancer", August 1-2, Vienna, Austria.

9.1.2. Scientific Events: Selection

9.1.2.1. Member of the Conference Program Committees

- Pierre-Alexandre Bliman was in the Editorial Board of the conference “European Control Conference”, Naples (Italy), 25-28 June 2019.

9.1.2.2. Reviewer

- Pierre-Alexandre Bliman, reviewer for the conferences IEEE Conference on Decision Control, European Control Conference.

- Dirk Drasdo, CELLMECH 2019 (June 3-6, 2019): Quantitative single-cell-based modelling reveals predictable response of growing tumour spheroids on external mechanical stress, and how this informs liver regeneration
- Dirk Drasdo, 3rd Workshop on BioFabrication & Cancer, 24-25.09.2019: Quantitative single-cell-based modelling reveals predictable response of growing tumour spheroids on external mechanical stress, and how this informs liver regeneration

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Dirk Drasdo is associate editorial member of J. Theor. Biol. Royal Society Open Science, and The Scientific World Journal.
- Marie Doumic is associate editor of Kinetic and Related Models and of the Journal of Mathematical Biology.
- Diane Peurichard and Nastassia Pouradier Duteil are guest editors of a special issue for the journal Mathematical Biosciences and Engineering, planned in september 2020.
- Benoit Perthame is associate editor of Bol. SeMA, Milan Journal of Mathematics, M3AS, Mathematical Medicine and Biology (IMA journal), Comm. in PDE,

9.1.3.2. Reviewer - Reviewing Activities

We reviewed for:

BBA Reviews Cancer, Mathematical Modelling of Natural Phenomena, Mathematical Biosciences, Mathematical Biosciences and Engineering, Bioinformatics, BBA Molecular Cell Research, Computers in Biology and Medicine, Biosystems, J Math Biol, Cancer Research, Bull Math Biol, J Theor Biol, Science Advances ;

Symmetry, Mathematical Biosciences and Engineering ;

IEEE Transactions on Automatic Control, Mathematical Biosciences;

Transactions of the London Mathematical Society, Journal of Mathematical Biology;

Annales de l'Institut Henri Lebesgue, SIAM Applied Math, J. Math. Biol, Asymptotic Analysis.

9.1.4. Invited Talks

- P.-A. Bliman's presentations at seminars and workshops:
  - Laboratoire Analyse, Géométrie et Applications UMR 7539, Université Paris 13, February 2019;
  - Institut de Mathématiques de Bordeaux UMR 5251, March 2019;
  - Universidad Nacional de Asunción, Asunción, Paraguay, March 2019 Université de Guyane, Cayenne, April 2019;
  - Universidade Federal de Amapá, Macapá, Brazil, April 2019;
  - Laboratoire Jacques-Louis Lions, Sorbonne Université, Paris, June 2019;
  - COPPE, Universidade Federal de Rio do Janeiro, Brazil, July 2019;
  - Fifth International Conference on Computational and Mathematical Population Dynamics, Fort Lauderdale (FL), USA May 2019;
  - Mathematical Biology on the Mediterranean Conference, Samos, September 2019;
  - Workshop on Advanced Control Techniques for Mosquito-Borne Diseases 2019, Asunción, Paraguay, November 2019.
- J. Clairambault's presentations at seminars and workshops:
  - Maths club Paris VII, January 2019;
  - LJLL 50-year anniversary, Roscoff March 2019;

  - Seminar Math-Bio, LAGA, Université Paris XIII, Villetaneuse, April 2019;
  - Mathematical Methods and Models in Biosciences Conference, IMPAN, Bedlewo, Poland, June 2019;
  - Mathematical Biology on the Mediterranean Conference, Samos, Greece, September 2019;
  - Symposium on Modelling approaches for cancer therapy, Cité scientifique, Lille, September 2019;
  - Q-Bio symposium Pasteur Institute, December 2019.
- D. Peurichard's presentations at seminars and workshops:
  - Colloque d'ouverture 50 ans du LJLL, Roscoff, March 2019;
  - ICIAM 9th International Congress on Industrial and Applied Mathematics, Valencia, july 15th-19th;
  - Math and cancer workshop in Vienna, 1-3th august, 2019;
  - MAMOVI days (MAthématiques de la MOdélisation du VIvant), Tours, 4th-6th sep, 2019;
  - Mathematical Biology on the Mediterranean Conference, Samos, Greece, September 2019;
  - PDE-biology workshop, Orsay, 25-26 nov 2019;
  - Math-bio seminar, Bordeaux, December, 2019.
- N. Pouradier Duteil's presentations at seminars and workshops:
  - Colloque d'ouverture 50 ans du LJLL, Roscoff, March 2019;
  - Journées Picard, Roscoff, March 2019;
  - Séminaire du laboratoire Jacques-Louis Lions, Paris, March 2019;
  - Workshop "PDEs and Applications in Life Sciences", Penn State, USA, October 2019;
  - INdAM workshop "Recent advances in kinetic equations and applications", Rome, Italy, November 2019.
- M. Doumic's presentations at seminars, workshops and conferences:
  - Inria-KAIST (Korean Advanced Institute of Science and Technology) meeting, Paris, March 28th.
  - Amiens PDE seminar, May 27th.
  - PDE Day, June 7th, Le Havre.
  - Four-hour course at the 2019 Summer School on Mathematical Biology, July 1-4, Shanghai.
  - Plenary talk at the AIP 2019 Conference (Applied Inverse Problems), July 7-11, Grenoble.
  - First meeting of MeDyNa GDR, October 7-10, Sainte Montaine.
  - DEA conference in Krakow, September 18-20.
  - CENTURI Conference in Marseille, November 18th, with W-F. Xue.

9.1.5. Leadership within the Scientific Community

Dirk Drasdo is member of the scientific leadership board of the German flagship project LiSyM (Liver Systems Medicine) financed by BMBF (Germany).

Marie Doumic is a member of the board of the GDR Medyna (self-assembly of peptides and proteins).

9.1.6. Scientific Expertise

M. Doumic is a member of ITMO BMSV (interdisciplinay institute for structural and biomolecular basis of living matter) experts' committee, representing Inria.

M. Doumic is a member of the scientific committee of the mathematical institute INSMI of CNRS, and a member of the interdisciplinary committee CID 51 of CNRS.

9.1.7. Research Administration

- Dirk Drasdo is associated with IfADo Leibniz-Institute. He directs a research group bi-localized at Inria de Paris and IfADo, Dortmund.
- Dirk Drasdo is member of the scientific leadership board of the German flagship project LiSyM (Liver Systems Medicine) financed by BMBF (Germany).
- D. Peurichard is a member of the CSD (Comité de Suivi Doctoral) at Inria
- P.-A. Bliman is the coordinator of the ECOS-Nord project C17M01 “News methods for control of dengue and arboviroses”, and international coordinator of the STIC AmSud project 20-STIC-05 “New Methods for Biological Control of the Arboviruses" epidemics”.
- M. Doumic is a member of the organisation committee for mentorship at Inria Paris.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Most non permanent researchers have teaching activities in Sorbonne Université. We detail below only some of the teaching activities of permanent researchers.

- N. Pouradier Duteil, “Travaux de Recherche Encadrés”, 10h, M1, Jussieu
- D. Drasdo, “Agent-based models of tissue organization”, Paris 24 h / yr, M2 course, UPMC, Paris, France
- D. Drasdo: "Integrated and spatial-temporal multiscale modeling of liver guide in vivo experiments in healthy & chronic disease states: a blue print for systems medicine?", M2 course, 1 h, Ecole Polytechnique, France
- D. Peurichard, TD M1 in Jussieu, "Fondements des méthodes numériques", 40h
- D. Peurichard, M1 TD in Jussieu, "Approximation des EDP", 24h
- D. Peurichard, M2 4 hours course in Strasbourg in the interdisciplinary program entitled ’Physique Cellulaire’
- B. Perthame, M2 course “Introduction to mathematical biology”
- M. Doumic, M2 Course "Inverse Problems in Population Dynamics"

(24 hours)

9.2.2. Supervision

PhD in progress:

- P.-A. Bliman is co-advisor of the PhD student Pastor E. Pére-Estigarribia (with Ch. Schaerer, at UNA, Paraguay);
- L. Almeida is Emma Leschiera’s co-supervisor (with Chloé Audebert, Sorbonne Université, and Tommaso Lorenzi, St Andrew’s university) and Jesus Bellver Arnau’s supervisor;
- M. Doumic is Mathieu Mezache’s co-supervisor (with H. Rezaei, Inra - defended in December 2019), Cécile della Valle’s co-supervisor (with Philippe Moireau, Inria Saclay), Julia Delacour’s co-supervisor (with Christian Schmeiser, Vienna), Hugo Martin’s co-supervisor (with Pierre Gabriel, Université de Versailles-St Quentin - defended in July 2019), Adrien Ellis’ co-supervisor (with Marc Hoffmann, Université Paris Dauphine - begun in October 2019) and Anaïs Rat co-supervisor (with Magali Tournus, Centrale Marseille - begun in October 2019);

- B. Perthame is Alexandre Poulain's supervisor, Giorgia Ciavovella's co-supervisor (with Roberto Natalini, Roma), Federica Bubba's co-supervisor (with Pasquale Ciarletta, Politecnico di Milano) and Noemi David's co-supervisor (with, University of Bologna);
- Diane Peurichard is Valeria Caliaro's supervisor;
- Philippe Robert is Gaëtan Vignoud's advisor.

M2 internship: P.-A. Bliman was Mahamadou Sylla's advisor; M. Doumic was Adrien Ellis' advisor; D. Peurichard was Valeria Caliaro's advisor; Luis Almeida is Jesus Bellver Arnau's supervisor and Jorge Hernandez co-supervisor.

9.2.3. Committees

- J. Clairambault's international project evaluation activity in 2019: ICREA, EPSRC
- J. Clairambault's PhD defence committees: Ghassen Haddad (JC supervisor), Paris March 1st; Mohammed Ladjimi, Lille, September 25.
- J. Clairambault is member of the ANR evaluation panel "Mathematics and their interactions for biology and health"
- M. Doumic is a member of the interdisciplinary committee 51 of CNRS (CID51): selection committee for junior and senior research scientists of CNRS.
- M. Doumic was a member of the Inria Paris research center selection committee for junior scientists.
- M. Doumic's Ph.D defence committees: Hugo Martin (MD supervisor), July Paris, Frédérique Robin (MD chair), September 26, Paris; Emma Horton (MD reviewer), November 14th, Bath; Mathieu Mezache (MD supervisor), December 17th, Paris; Samuel Nordmann, October 11th, Paris.
- D. Peurichard is a member of the Inria "Comission des emplois scientifiques" for selecting PhD, delegation and post-doctoral candidates at Inria.
- D. Peurichard is a member of the ARP/SRP selection comittee for Inria.

9.3. Popularization

Diane Peurichard presented a talk in the framework of the "Remise des prix des Olympiades de mathématiques de Paris" (prize for national mathematics competition for young french students), Wednesday 29th, June 2019.

Diane Peurichard presented her work at the Inria scientific days in Lyon from June 5th- 7th, 2019.

At the occasion of scientific seminaries, Pierre-Alexandre Bliman gave interviews to TV channels in Paraguay: Paraguai TV (March 28, 2019); and in Brazil: Globo Amapá (April 16, 2019), Balanço Geral AP (April 17th, 2019).

Marie Doumic gave a presentation of her research for middle-school pupils in internship (December 2019) and for highschool mathematics teachers (March 2019, in Lille), and did "speed-dating interviews" with high school students (February 12th, NGO "Sephora Berrebi").

Benoit Perthame gave public talks at Tamkang university (Taipei) and Nancy.

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] H. MARTIN.*Data analysis and integro-differential models in cellular biology*, Sorbonne Université - Laboratoire Jacques-Louis Lions, July 2019, https://hal.archives-ouvertes.fr/tel-02386939

[2] M. MEZACHE.*Oscillatory processes during the aggregation and the fragmentation of the amyloid fibrils*, Sorbonne Université - Laboratoire Jacques-Louis Lions ; Inria Paris, December 2019, https://hal.archives-ouvertes.fr/tel-02435325

[3] P. VAN LIEDEKERKE.*Quantitative modeling of cell and tissue mechanics with agent-based models*, Inria Paris, Sobonne Université, March 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02064216

Articles in International Peer-Reviewed Journal

[4] J. ALLARD, M. DOUMIC, A. MOGILNER, D. OELZ.*Bidirectional sliding of two parallel microtubules generated by multiple identical motors*, in "Journal of Mathematical Biology", April 2019 [*DOI :* 10.1007/S00285-019-01369-W], https://hal.archives-ouvertes.fr/hal-02181135

[5] L. ALMEIDA, P. BAGNERINI, G. FABRINI, B. D. HUGHES, T. LORENZI.*Evolution of cancer cell populations under cytotoxic therapy and treatment optimisation: insight from a phenotype-structured model*, in "ESAIM: Mathematical Modelling and Numerical Analysis", July 2019, vol. 53, n^o^ 4, p. 1157-1190 [*DOI :* 10.1051/M2AN/2019010], https://hal.archives-ouvertes.fr/hal-02433962

[6] L. ALMEIDA, F. BUBBA, B. PERTHAME, C. POUCHOL.*Energy and implicit discretization of the Fokker-Planck and Keller-Segel type equations*, in "Networks and Heterogeneous Media", March 2019, vol. 14, n^o^ 1, https://arxiv.org/abs/1803.10629 [*DOI :* 10.3934/NHM.2019002], https://hal.archives-ouvertes.fr/hal-01745769

[7] L. ALMEIDA, M. DUPREZ, Y. PRIVAT, N. VAUCHELET.*Mosquito population control strategies for fighting against arboviruses*, in "Mathematical Biosciences and Engineering", 2019, vol. 16, n^o^ 6, p. 6274-6297, https://arxiv.org/abs/1901.05688 [*DOI :* 10.3934/MBE.2019313], https://hal.archives-ouvertes.fr/hal-01984426

[8] L. ALMEIDA, Y. PRIVAT, M. STRUGAREK, N. VAUCHELET.*Optimal releases for population replacement strategies, application to Wolbachia*, in "SIAM Journal on Mathematical Analysis", 2019, vol. 51, n^o^ 4, p. 3170–3194, https://arxiv.org/abs/1909.02727 [*DOI :* 10.1137/18M1189841], https://hal.archives-ouvertes.fr/hal-01807624

[9] E. BERNARD, M. DOUMIC, P. GABRIEL.*Cyclic asymptotic behaviour of a population reproducing by fission into two equal parts*, in "Kinetic and Related Models ", June 2019, vol. 12, n^o^ 3, p. 551-571, https://arxiv.org/abs/1609.03846 [*DOI :* 10.3934/KRM.2019022], https://hal.archives-ouvertes.fr/hal-01363549

[10] P.-A. BLIMAN, D. CARDONA-SALGADO, Y. DUMONT, O. VASILIEVA.*Implementation of Control Strategies for Sterile Insect Techniques*, in "Mathematical Biosciences", 2019, vol. 314, p. 43-60 [*DOI :* 10.1016/J.MBS.2019.06.002], https://hal.inria.fr/hal-01943683

[11] V. BOEUF, P. ROBERT.*A Stochastic Analysis of a Network with Two Levels of Service*, in "Queueing Systems", August 2019, vol. 92, n^o^ 3-4, 30, https://arxiv.org/abs/1708.09590 , forthcoming [*DOI :* 10.1007/S11134-019-09617-Y], https://hal.archives-ouvertes.fr/hal-01583704

[12] F. BUBBA, C. POUCHOL, N. FERRAND, G. VIDAL, L. ALMEIDA, B. PERTHAME, M. SABBAH.*A chemotaxis-based explanation of spheroid formation in 3D cultures of breast cancer cells*, in "Journal of Theoretical Biology", 2019, vol. 479, p. 73-80, https://arxiv.org/abs/1810.13162 [*DOI :* 10.1016/J.JTBI.2019.07.002], https://hal.archives-ouvertes.fr/hal-01895999

[13] J. CLAIRAMBAULT.*An evolutionary perspective on cancer, with applications to anticancer drug resistance modelling and perspectives in therapeutic control*, in "Journal of Mathematical Study", 2019, p. 1 - 21, forthcoming, https://hal.inria.fr/hal-02126730

[14] J. CLAIRAMBAULT, C. POUCHOL.*A survey of adaptive cell population dynamics models of emergence of drug resistance in cancer, and open questions about evolution and cancer*, in "BIOMATH", May 2019, vol. 8, n° 1, 23, Copyright : 2019 Clairambault et al. This article is distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited [*DOI :* 10.11145/J.BIOMATH.2019.05.147], https://hal.inria.fr/hal-02132713

[15] R. DESSALLES, V. FROMION, P. ROBERT.*Models of protein production with cell cycle*, in "PLoS ONE", January 2020, vol. 15, n° 1, 25, https://arxiv.org/abs/1711.06378 [*DOI :* 10.1371/JOURNAL.PONE.0226016], https://hal.inria.fr/hal-01643441

[16] P. GABRIEL, H. MARTIN.*Steady distribution of the incremental model for bacteria proliferation*, in "Networks and Heterogeneous Media", March 2019, vol. 14, n° 1, p. 149-171, https://arxiv.org/abs/1803.04950 [*DOI :* 10.3934/NHM.2019008], https://hal.archives-ouvertes.fr/hal-01742140

[17] A. GHALLAB, M. MYLLYS, C. H. HOLLAND, A. ZAZA, W. MURAD, R. HASSAN, Y. A. AHMED, T. ABBAS, E. A. ABDELRAHIM, K. M. SCHNEIDER, M. MATZ-SOJA, J. REINDERS, R. GEBHARDT, M.-L. BERRES, M. HATTING, D. DRASDO, J. SAEZ-RODRIGUEZ, C. TRAUTWEIN, J. G. HENGSTLER.*Influence of Liver Fibrosis on Lobular Zonation*, in "Cells", December 2019, vol. 8, n° 12, 1556 [*DOI :* 10.3390/CELLS8121556], https://hal.inria.fr/hal-02425184

[18] T. HECK, D. A. VARGAS, B. SMEETS, H. RAMON, P. VAN LIEDEKERKE, H. VAN OOSTERWYCK.*The role of actin protrusion dynamics in cell migration through a degradable viscoelastic extracellular matrix: Insights from a computational model*, in "PLoS Computational Biology", 2020, vol. 16, e1007250 [*DOI :* 10.1371/JOURNAL.PCBI.1007250], https://hal.sorbonne-universite.fr/hal-02486182

[19] A. IGEL-EGALON, F. LAFERRIÈRE, M. MOUDJOU, J. BOHL, M. MEZACHE, T. KNÄPPLE, L. HERZOG, F. REINE, C. JAS-DUVAL, M. DOUMIC, H. REZAEI, V. BERINGUE.*Early stage prion assembly involves two subpopulations with different quaternary structures and a secondary templating pathway*, in "Communications Biology", December 2019, vol. 2, n° 1 [*DOI :* 10.1038/S42003-019-0608-Y], https://hal.umontpellier.fr/hal-02317910

[20] G. JANKOWIAK, D. PEURICHARD, A. REVERSAT, C. SCHMEISER, M. SIXT.*Modelling adhesion-independent cell migration*, in "Mathematical Models and Methods in Applied Sciences", 2020, forthcoming [*DOI :* 10.1142/S021820252050013X], https://hal.archives-ouvertes.fr/hal-02076901

[21] F. JOLY, G. SOULEZ, S. LESSARD, C. KAUFFMANN, I. VIGNON-CLEMENTEL.*A cohort longitudinal study identifies morphology and hemodynamics predictors of abdominal aortic aneurysm growth*, in "Annals of Biomedical Engineering", October 2019, forthcoming [*DOI :* 10.1007/S10439-019-02375-1], https://hal.archives-ouvertes.fr/hal-02302464

[22] P. V. LIEDEKERKE, J. NEITSCH, T. JOHANN, E. WARMT, I. GONZÀLEZ-VALVERDE, S. HOEHME, S. GROSSER, J. KAES, D. DRASDO.*A quantitative high-resolution computational mechanics cell model for growing and regenerating tissues*, in "Biomechanics and Modeling in Mechanobiology", November 2019 [*DOI :* 10.1007/S10237-019-01204-7], https://hal.inria.fr/hal-01956023

[23] D. MARTIROSYAN, P. ROBERT.*The Equilibrium States of Large Networks of Erlang Queues*, in "Advances in Applied Probability", 2020, https://arxiv.org/abs/1811.04763 , forthcoming, https://hal.inria.fr/hal-01920171

[24] T. N. NGUYEN, J. CLAIRAMBAULT, T. JAFFREDO, B. PERTHAME, D. SALORT.*Adaptive dynamics of hematopoietic stem cells and their supporting stroma: A model and mathematical analysis*, in "Mathematical Biosciences and Engineering", May 2019, vol. 16, n^o 05, p. 4818–4845 [*DOI :* 10.3934/MBE.2019243], https://hal.archives-ouvertes.fr/hal-01963820

[25] B. PERTHAME, N. VAUCHELET, Z. WANG.*The Flux Limited Keller-Segel System; Properties and Derivation from Kinetic Equations*, in "Revista Matemática Iberoamericana", 2019, https://arxiv.org/abs/1801.07062 , forthcoming, https://hal.archives-ouvertes.fr/hal-01689571

[26] D. PEURICHARD, M. OUSSET, J. PAUPERT, B. AYMARD, A. LORSIGNOL, L. CASTEILLA, P. DEGOND.*Extra-cellular matrix rigidity may dictate the fate of injury outcome*, in "Journal of Theoretical Biology", May 2019, vol. 469, p. 127-136 [*DOI :* 10.1016/J.JTBI.2019.02.017], https://hal.archives-ouvertes.fr/hal-02345773

[27] B. PICCOLI, N. POURADIER DUTEIL, E. TRÉLAT.*Sparse control of Hegselmann-Krause models: Black hole and declustering*, in "SIAM Journal on Control and Optimization", 2019, vol. 57, n^o 4, p. 2628–2659 [*DOI :* 10.1137/18M1168911], https://hal.archives-ouvertes.fr/hal-01699261

[28] C. POUCHOL, E. TRÉLAT, E. ZUAZUA.*Phase portrait control for 1D monostable and bistable reaction-diffusion equations*, in "Nonlinearity", 2019, vol. 32, n^o 3, p. 884–909, https://arxiv.org/abs/1805.10786 [*DOI :* 10.1088/1361-6544/AAF07E], https://hal.archives-ouvertes.fr/hal-01800382

[29] M. RABÉ, S. DUMONT, A. ÁLVAREZ-ARENAS, H. JANATI, J. BELMONTE-BEITIA, G. F. CALVO, C. THIBAULT-CARPENTIER, Q. SÉRY, C. CHAUVIN, N. JOALLAND, F. BRIAND, S. BLANDIN, E. SCOTET, C. PECQUEUR, J. CLAIRAMBAULT, L. OLIVER, V. PÉREZ-GARCÍA, A. M. NADARADJANE, P.-F. CARTRON, C. GRATAS, F. VALLETTE.*Identification of a transient state during the acquisition of temozolomide resistance in glioblastoma*, in "Cell Death and Disease", January 2020, vol. 11, n^o 1 [*DOI :* 10.1038/S41419-019-2200-2], https://hal.archives-ouvertes.fr/hal-02429194

[30] P. ROBERT.*Mathematical Models of Gene Expression*, in "Probability Surveys", October 2019, vol. 16, 56, https://arxiv.org/abs/1905.02578 , https://hal.inria.fr/hal-02316326

[31] P. ROBERT, W. SUN.*On the Asymptotic Distribution of Nucleation Times of Polymerization Processes*, in "SIAM Journal on Applied Mathematics", October 2019, vol. 79, n^o 5, 27, https://arxiv.org/abs/1712.08347 [*DOI :* 10.1137/19M1237508], https://hal.inria.fr/hal-01672800

[32] P. ROBERT, A. VÉBER.*A Scaling Analysis of a Star Network with Logarithmic Weights*, in "Stochastic Processes and their Applications", 2019, https://arxiv.org/abs/1609.04180 [*DOI :* 10.1016/J.SPA.2018.06.002], https://hal.archives-ouvertes.fr/hal-01377703

[33] M. STRUGAREK, L. DUFOUR, N. VAUCHELET, L. ALMEIDA, B. PERTHAME, D. A. M. VILLELA.*Oscillatory regimes in a mosquito population model with larval feedback on egg hatching*, in "Journal of Biological Dynamics", 2019, vol. 13, n^o 1, p. 269-300, https://arxiv.org/abs/1801.03701 [*DOI :* 10.1080/17513758.2019.1593524], https://hal.archives-ouvertes.fr/hal-01674280

[34] W. SUN, P. ROBERT.*Analysis of Large Urn Models with Local Mean-Field Interactions*, in "Electronic Journal of Probability", April 2019, vol. 24, n^{o} 45, 33, https://arxiv.org/abs/1802.05064 , forthcoming, https://hal.inria.fr/hal-01710964

[35] J. TORRENT, D. MARTIN, S. NOINVILLE, Y. YIN, M. DOUMIC, M. MOUDJOU, V. V. BERINGUE, H. REZAEI.*Pressure Reveals Unique Conformational Features in Prion Protein Fibril Diversity*, in "Scientific Reports", 2019, vol. 9, 2802 [*DOI :* 10.1038/S41598-019-39261-8], https://hal.sorbonne-universite.fr/hal-02057700

[36] B. UJVARI, C. JACQUELINE, D. MISSÉ, V. AMAR, J. FITZPATRICK, G. JENNINGS, C. BECKMANN, S. ROME, P. BIRO, R. GATENBY, J. BROWN, F. THOMAS, L. ALMEIDA.*Obesity paradox in cancer: Is bigger really better?*, in "Evolutionary Applications", April 2019, vol. 12, n^{o} 6, p. 1092-1095 [*DOI :* 10.1111/EVA.12790], https://hal.archives-ouvertes.fr/hal-02433971

[37] P. VAN LIEDEKERKE, J. NEITSCH, T. JOHANN, K. ALESSANDRI, P. NASSOY, D. DRASDO.*Quantitative cell-based model predicts mechanical stress response of growing tumor spheroids over various growth conditions and cell lines*, in "PLoS Computational Biology", March 2019, vol. 15, n^{o} 3, e1006273 [*DOI :* 10.1371/JOURNAL.PCBI.1006273], https://hal.archives-ouvertes.fr/hal-02371813

[38] P. VAN LIEDEKERKE, J. NEITSCH, T. JOHANN, K. ALESSANDRI, P. NASSOY, D. DRASDO.*Quantitative cell-based model predicts mechanical stress response of growing tumor spheroids over various growth conditions and cell lines*, in "PLoS Computational Biology", March 2019, vol. 15, n^{o} 3, e1006273 [*DOI :* 10.1371/JOURNAL.PCBI.1006273], https://hal.inria.fr/hal-02425194

International Conferences with Proceedings

[39] P.-A. BLIMAN.*Feedback Control Principles for Biological Control of Dengue Vectors*, in "European Control Conference ECC19", Naples, Italy, June 2019, 6, The last version of this manuscript has been accepted for publication in the Proceedings of European Control Conference ECC19. A complete version of the report appeared in arXiv, and is available at: http://arxiv.org/abs/1903.00730, https://hal.inria.fr/hal-01944958

[40] R. USHIROBIRA, D. EFIMOV, P.-A. BLIMAN.*Estimating the infection rate of a SIR epidemic model via differential elimination*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, June 2019, https://hal.inria.fr/hal-02418553

Other Publications

[41] L. ALMEIDA, B. PERTHAME, X. RUAN.*An Asymptotic Preserving Scheme for Capturing Concentrations in Age-structured Models Arising in Adaptive Dynamics*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02438316

[42] P.-A. BLIMAN, D. CARDONA-SALGADO, Y. DUMONT, O. VASILIEVA.*Optimal Control Approach for Implementation of Sterile Insect Techniques*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02418508

[43] F. BUBBA, B. PERTHAME, C. POUCHOL, M. SCHMIDTCHEN.*Hele-Shaw limit for a system of two reaction-(cross-)diffusion equations for living tissues*, January 2019, https://arxiv.org/abs/1901.01692 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01970313

[44] F. BUBBA, A. POULAIN.*A Positivity-Preserving Finite Element Scheme for the Relaxed Cahn-Hilliard Equation with Single-Well Potential and Degenerate Mobility*, October 2019, https://arxiv.org/abs/1910.13211 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02336530

[45] J. CLAIRAMBAULT, C. POUCHOL.*A survey of adaptive cell population dynamics models of emergence of drug resistance in cancer, and open questions about evolution and cancer*, May 2019, working paper or preprint, https://hal.inria.fr/hal-02126727

[46] M. DOUMIC, A. OLIVIER, L. ROBERT.*Estimating the division rate from indirect measurements of single cells*, July 2019, https://arxiv.org/abs/1907.05108 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02175633

[47] T. DĘBIEC, B. PERTHAME, M. SCHMIDTCHEN, N. VAUCHELET.*Incompressible limit for a two-species model with coupling through Brinkman's law in any dimension*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02461406

[48] D. FANG, S. JIN, P. MARKOWICH, B. PERTHAME.*Implicit and Semi-implicit Numerical Schemes for the Gradient Flow of the Formation of Biological Transport Networks*, January 2019, working paper or preprint, https://hal.sorbonne-universite.fr/hal-01984371

[49] N. FOURNIER, B. PERTHAME.*Monge-Kantorovich distance for PDEs: the coupling method*, March 2019, https://arxiv.org/abs/1903.11349 - working paper or preprint, https://hal.sorbonne-universite.fr/hal-02080155

[50] P. GABRIEL, H. MARTIN.*Periodic asymptotic dynamics of the measure solutions to an equal mitosis equation*, September 2019, https://arxiv.org/abs/1909.08276 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02290410

[51] E.-H. LAAMRI, B. PERTHAME.*Reaction-diffusion systems with initial data of low regularity*, August 2019, https://arxiv.org/abs/1908.09693 - working paper or preprint, https://hal.sorbonne-universite.fr/hal-02271103

[52] P. V. LIEDEKERKE, J. NEITSCH, T. JOHANN, K. ALESSANDRI, P. NASSOY, D. DRASDO.*Quantitative agent-based modeling reveals mechanical stress response of growing tumor spheroids is predictable over various growth conditions and cell lines*, January 2019, working paper or preprint, https://hal.inria.fr/hal-01956017

[53] D. MARTIROSYAN, V. NERSESYAN.*Multiplicative ergodic theorem for a non-irreducible random dynamical system*, January 2020, https://arxiv.org/abs/1801.09440 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01695046

[54] S. T. MCQUADE, B. PICCOLI, N. POURADIER DUTEIL.*Social Dynamics Models with Time-Varying Influence*, April 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02090560

[55] M. MEZACHE, M. HOFFMANN, H. REZAEI, M. DOUMIC.*Testing for high frequency features in a noisy signal*, November 2019, https://arxiv.org/abs/1911.12719 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02263522

[56] S. NORDMANN, B. PERTHAME.*Dynamics of concentration in a population structured by age and a phenotypic trait with mutations. Convergence of the corrector*, January 2020, https://arxiv.org/abs/2001.04323 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02437168

[57] B. PERTHAME, A. POULAIN.*Relaxation of the Cahn-Hilliard equation with singular single-well potential and degenerate mobility*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02274417

[58] B. PERTHAME, W. SUN, M. TANG, S. YASUDA.*Multiple asymptotics of kinetic equations with internal states*, July 2019, working paper or preprint, https://hal.sorbonne-universite.fr/hal-02194421

[59] P. E. PÉREZ-ESTIGARRIBIA, P.-A. BLIMAN, C. E. SCHAERER.*A fast-slow model for adaptive resistance evolution*, December 2019, https://arxiv.org/abs/1912.10733 - working paper or preprint, https://hal.inria.fr/hal-02418511

References in notes

[60] L. ALMEIDA, R. H. CHISHOLM, J. CLAIRAMBAULT, T. LORENZI, A. LORZ, C. POUCHOL, E. TRÉLAT.*Why Is Evolution Important in Cancer and What Mathematics Should Be Used to Treat Cancer? Focus on Drug Resistance*, in "Trends in Biomathematics: Modeling, Optimization and Computational Problems: Selected works from the BIOMAT Consortium Lectures, Moscow 2017", Springer International Publishing, August 2018, p. 107-120, https://hal.inria.fr/hal-01945593

[61] L. ALPHEY.*Genetic control of mosquitoes*, in "Annual review of entomology", 2014, vol. 59

[62] A. ARMIENTO.*Inverse problems and data assimilation methods applied to protein polymerisation*, Université Paris 7 - Diderot, January 2017, https://hal.inria.fr/tel-01447286

[63] A. ARMIENTO, M. DOUMIC, P. MOIREAU, H. REZAEI.*Estimation from Moments Measurements for Amyloid Depolymerisation*, in "Journal of Theoretical Biology", March 2016 [*DOI :* 10.1016/J.JTBI.2016.02.037], https://hal.archives-ouvertes.fr/hal-01248255

[64] A. ARMIENTO, P. MOIREAU, D. MARTIN, N. LEPEJOVA, M. DOUMIC, H. REZAEI.*The mechanism of monomer transfer between two structurally distinct PrP oligomers*, in "PLoS ONE", July 2017, vol. 12, n^{o} 7 [*DOI :* 10.1371/JOURNAL.PONE.0180538], https://hal.archives-ouvertes.fr/hal-01574346

[65] J. L. AVILA ALONSO, C. BONNET, J. CLAIRAMBAULT, H. OZBAY, S.-I. NICULESCU, F. MERHI, A. BALLESTA, R. TANG, J.-P. MARIE.*Analysis of a New Model of Cell Population Dynamics in Acute Myeloid Leukemia*, in "Delay Systems : From Theory to Numerics and Applications", T. VYHLÍDAL, J.-F. LAFAY, R. SIPAHI (editors), Advances in Delays and Dynamics, Springer, January 2014, vol. 1, p. 315-328 [*DOI :* 10.1007/978-3-319-01695-5_23], https://hal.inria.fr/hal-00932779

[66] J. L. AVILA ALONSO, C. BONNET, E. FRIDMAN, F. MAZENC, J. CLAIRAMBAULT.*Stability analysis of PDE's modelling cell dynamics in Acute Myeloid Leukemia*, in "53rd IEEE Conference on Decision and Control", Los Angeles, United States, December 2014, https://hal.inria.fr/hal-01110304

[67] J. L. AVILA ALONSO, C. BONNET, H. OZBAY, J. CLAIRAMBAULT, S.-I. NICULESCU.*A coupled model for healthy and cancer cells dynamics in Acute Myeloid Leukemia*, in "The 19th World Congress of the International Federation of Automatic Control", Cape Town, Souh Africa, August 2014, https://hal.inria.fr/hal-00940245

[68] J. L. AVILA ALONSO, C. BONNET, H. OZBAY, J. CLAIRAMBAULT, S.-I. NICULESCU.*A discrete-maturity Interconnected Model of Healthy and Cancer Cell Dynamics in Acute Myeloid Leukemia*, in "Mathematical Theory of Networks and Systems", Groningen, Netherlands, July 2014, https://hal.inria.fr/hal-00940305

[69] H. T. BANKS, M. DOUMIC, C. KRUSE, S. PRIGENT, H. REZAEI.*Information Content in Data Sets for a Nucleated-Polymerization Model*, in "Journal of Biological Dynamics", June 2015, vol. 9, n^{o} 1, 26 [*DOI :* 10.1080/17513758.2015.1050465], https://hal.inria.fr/hal-01123847

[70] H. T. BANKS, M. DOUMIC-JAUFFRET, C. KRUSE.*A numerical scheme for the early steps of nucleation-aggregation Models*, in "Journal of Mathematical Biology", January 2017, vol. 74, n^{o} 1-2, p. 259-287 [*DOI :* 10.1007/S00285-016-1026-0], https://hal.inria.fr/hal-00954437

[71] J. BARRÉ, J. A. CARILLO, P. DEGOND, E. ZATORSKA, D. PEURICHARD. *Particle interactions mediated by dynamical networks: assessment of macroscopic descriptions*, in " Journal of Nonlinear Science", 2017

[72] J. BARRÉ, P. DEGOND, E. ZATORSKA, D. PEURICHARD. *Modelling pattern formation through differential repulsion*, in "Networks and Heterogeneous media", 2019, to appear

[73] F. BEKKAL BRIKCI, J. CLAIRAMBAULT, B. PERTHAME.*Analysis of a molecular structured population model with possible polynomial growth for the cell division cycle*, in "Math. Comput. Modelling", 2008, vol. 47, n^{o} 7-8, p. 699–713

[74] F. BERTAUX, S. STOMA, D. DRASDO, G. BATT.*Modeling Dynamics of Cell-to-Cell Variability in TRAIL-Induced Apoptosis Explains Fractional Killing and Predicts Reversible Resistance*, in "PLoS Computational Biology", 2014, vol. 10, n^{o} 10, 14 [*DOI :* 10.1371/JOURNAL.PCBI.1003893.S016], https://hal.inria.fr/hal-00942885

[75] J. BERTOIN, A. R. WATSON.*Probabilistic aspects of critical growth-fragmentation equations*, in "Advances in Applied Probability", 9 2015

[76] P.-A. BLIMAN, M. S. ARONNA, F. C. COELHO, M. A. H. B. DA SILVA.*Ensuring successful introduction of Wolbachia in natural populations of Aedes aegypti by means of feedback control*, in "Journal of Mathematical Biology", August 2017 [*DOI :* 10.1007/S00285-017-1174-X], https://hal.inria.fr/hal-01579477

[77] C. BONNET, J. L. AVILA ALONSO, H. OZBAY, J. CLAIRAMBAULT, S.-I. NICULESCU, P. HIRSCH.*A Discrete-Maturity Interconnected Model of Healthy and Cancer Cell Dynamics in Acute Myeloid Leukemia*, in "The 10th AIMS Conference on Dynamical Systems,Differential Equations and Applications", Madrid, Spain, July 2014, https://hal.inria.fr/hal-01110309

[78] T. BOURGERON, M. DOUMIC, M. ESCOBEDO.*Estimating the division rate of the growth-fragmentation equation with a self-similar kernel*, in "Inverse Problems", Jan 2014, vol. 30, n^{o} 2, 025007, http://dx.doi.org/10.1088/0266-5611/30/2/025007

[79] T. BOURGERON, Z. XU, M. DOUMIC, M. T. TEIXEIRA.*The asymmetry of telomere replication contributes to replicative senescence heterogeneity*, in "Scientific Reports", October 2015, vol. 5, 15326 [*DOI :* 10.1038/SREP15326], http://hal.upmc.fr/hal-01272075

[80] K. BOURTZIS.*Wolbachia-Based Technologies for Insect Pest Population Control*, in "Advances in Experimental Medicine and Biology", Springer, New York, NY, 02 2008, vol. 627

[81] A. BROWN.*Insecticide resistance in mosquitoes: a pragmatic review.*, in "Journal of the American Mosquito Control Association", 1986, vol. 2, n^{o} 2, p. 123–140

[82] F. BUBBA, C. POUCHOL, N. FERRAND, G. VIDAL, L. ALMEIDA, B. PERTHAME, M. SABBAH.*A chemotaxis-based explanation of spheroid formation in 3D cultures of breast cancer cells*, in "Journal of Theoretical Biology", 2019, vol. 479, p. 73-80 [*DOI :* 10.1016/J.JTBI.2019.07.002], https://hal.archives-ouvertes.fr/hal-01895999

[83] M. J. CACERES, B. PERTHAME.*Beyond blow-up in excitatory integrate and fire neuronal networks: refractory period and spontaneous activity*, in "Journal of Theoretical Biology", 2014, vol. 350, p. 81-89 [*DOI :* 10.1016/J.JTBI.2014.02.005], http://hal.upmc.fr/hal-00874746

[84] V. CALVEZ, M. DOUMIC, P. GABRIEL.*Self-similarity in a general aggregation–fragmentation problem. Application to fitness analysis*, in "Journal de Mathématiques Pures et Appliquées", 2012, vol. 98, n^{o} 1, p. 1 - 27 [*DOI :* 10.1016/J.MATPUR.2012.01.004], http://www.sciencedirect.com/science/article/pii/S002178241200013X

[85] V. CALVEZ, N. LENUZZA, M. DOUMIC, J.-P. DESLYS, F. MOUTHON, B. PERTHAME.*Prion dynamic with size dependency - strain phenomena*, in "J. of Biol. Dyn.", 2010, vol. 4, n^{o} 1, p. 28–42

[86] V. CALVEZ, B. PERTHAME, S. YASUDA.*Traveling Wave and Aggregation in a Flux-Limited Keller-Segel Model*, in "Kinetic and Related Models ", 2018, vol. 11, n^{o} 4, p. 891–909 [*DOI :* 10.3934/KRM.2018035], https://hal.sorbonne-universite.fr/hal-01591490

[87] J. A. CARRILLO, F. JAMES, F. LAGOUTIÈRE, N. VAUCHELET.*The Filippov characteristic flow for the aggregation equation with mildly singular potentials*, in "Journal of Differential Equations", 2016, vol. 260, n^{o} 1, p. 304-338, 33 pages, https://hal.archives-ouvertes.fr/hal-01061991

[88] G. CELLIÈRE.*Multi-scale modeling of hepatic drug toxicity and its consequences on ammonia detoxification*, Université Paris 6 - Pierre et Marie Curie, July 2017

[89] J. CHEVALLIER, M. J. CACERES, M. DOUMIC, P. REYNAUD-BOURET.*Microscopic approach of a time elapsed neural model*, in "Mathematical Models and Methods in Applied Sciences", December 2015, 2669 [*DOI :* 10.1142/S021820251550058X], http://hal.upmc.fr/hal-01159215

[90] R. H. CHISHOLM, T. LORENZI, J. CLAIRAMBAULT.*Cell population heterogeneity and evolution towards drug resistance in cancer: Biological and mathematical assessment, theoretical treatment optimisation*, in "BBA - General Subjects", June 2016, vol. 1860, p. 2627 - 2645 [*DOI :* 10.1016/J.BBAGEN.2016.06.009], https://hal.inria.fr/hal-01321535

[91] R. H. CHISHOLM, T. LORENZI, A. LORZ, A. K. LARSEN, L. N. D. ALMEIDA, A. ESCARGUEIL, J. CLAIRAMBAULT.*Emergence of Drug Tolerance in Cancer Cell Populations: An Evolutionary Outcome of Selection, Nongenetic Instability, and Stress-Induced Adaptation*, in "Cancer Research", Jan 2015, vol. 75, n^{o} 6, p. 930–939 [*DOI :* 10.1158/0008-5472.CAN-14-2103], http://dx.doi.org/10.1158/0008-5472.CAN-14-2103

[92] R. H. CHISHOLM, T. LORENZI, A. LORZ, A. K. LARSEN, L. N. DE ALMEIDA, A. ESCARGUEIL, J. CLAIRAMBAULT.*Emergence of Drug Tolerance in Cancer Cell Populations: An Evolutionary Outcome of Selection, Nongenetic Instability, and Stress-Induced Adaptation*, in "Cancer Research", March 2015, vol. 75, n^{o} 6, p. 930-939 [*DOI :* 10.1158/0008-5472.CAN-14-2103], https://hal.inria.fr/hal-01237893

[93] 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

[94] 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

[95] W. DJEMA, F. MAZENC, C. BONNET, J. CLAIRAMBAULT, P. HIRSCH, F. DELHOMMEAU.*Stability of a Delay System Coupled to a Differential-Difference System Describing the Coexistence of Ordinary and Mutated Hematopoietic Stem Cells*, in "Conference on Decision and Control ", Las Vegas, United States, December 2016, https://hal.inria.fr/hal-01389870

[96] M. DOUMIC, M. ESCOBEDO.*Time Asymptotics for a Critical Case in Fragmentation and Growth-Fragmentation Equations*, in "Kinetic and Related Models ", June 2016, vol. 9, n^{o} 2, 47 [*DOI :* 10.3934/KRM.2016.9.251], https://hal.inria.fr/hal-01080361

[97] M. DOUMIC, M. ESCOBEDO, M. TOURNUS.*Estimating the division rate and kernel in the fragmentation equation*, in "Annales de l'Institut Henri Poincaré (C) Non Linear Analysis", 2018, https://arxiv.org/abs/1804.08945 , forthcoming [*DOI :* 10.1016/J.ANIHPC.2018.03.004], https://hal.archives-ouvertes.fr/hal-01501811

[98] M. DOUMIC, S. EUGENE, P. ROBERT.*Asymptotics of Stochastic Protein Assembly Models*, in "SIAM Journal on Applied Mathematics", November 2016, vol. 76, n^{o} 6, 20 [*DOI :* 10.1137/16M1066920], https://hal.inria.fr/hal-01301266

[99] M. DOUMIC, P. GABRIEL.*Eigenelements of a General Aggregation-Fragmentation Model*, in "Mathematical Models and Methods in Applied Sciences", 2009, vol. 20, n^{o} 05, 757, http://arxiv.org/abs/0907.5467

[100] M. DOUMIC, M. HOFFMANN, N. KRELL, L. ROBERT.*Statistical estimation of a growth-fragmentation model observed on a genealogical tree*, October 2012, 46 pages, 4 figures, https://hal.archives-ouvertes.fr/hal-00763601

[101] M. DOUMIC, B. PERTHAME, J. ZUBELLI.*Numerical Solution of an Inverse Problem in Size-Structured Population Dynamics*, in "Inverse Problems", 2009, vol. 25, n^{o} 4, 045008

[102] D. DRASDO, A. BUTTENSCHÖN, P. VAN LIEDEKERKE.*Agent-Based Lattice Models of Multicellular Systems*, in "Numerical Methods and Advanced Simulation in Biomechanics and Biological Processes", Elsevier, 2018, p. 223-238 [*DOI :* 10.1016/B978-0-12-811718-7.00012-5], https://hal.inria.fr/hal-01968847

[103] D. DRASDO, S. HOEHME, J. G. HENGSTLER.*How predictive quantitative modeling of tissue organization can inform liver disease pathogenesis*, in "Journal of Hepatology", October 2014, vol. 61, n^{o} 4, p. 951–956 [*DOI :* 10.1016/J.JHEP.2014.06.013], https://hal.inria.fr/hal-01110644

[104] V. A. DYCK, J. HENDRICHS, A. S. ROBINSON.*The Sterile Insect Technique, Principles and Practice in Area-Wide Integrated Pest Management*, Springer, Dordrecht, 2006, 787

[105] S. EUGENE.*Stochastic modelling in molecular biology: a probabilistic analysis of protein polymerisation and telomere shortening*, UPMC LJLL, September 2016, https://hal.inria.fr/tel-01377561

[106] S. EUGENE, W.-F. XUE, P. ROBERT, M. DOUMIC-JAUFFRET.*Insights into the variability of nucleated amyloid polymerization by a minimalistic model of stochastic protein assembly*, in "Journal of Chemical Physics", May 2016, vol. 144, n^{o} 17, 12 [*DOI :* 10.1063/1.4947472], https://hal.inria.fr/hal-01205549

[107] A. GHALLAB, G. CELLIÈRE, S. HENKEL, D. DRIESCH, S. HOEHME, U. HOFMANN, S. ZELLMER, P. GODOY, A. SACHINIDIS, M. BLASZKEWICZ, R. REIF, R. MARCHAN, L. KUEPFER, D. HÄUSSINGER, D. DRASDO, G. GEBHARDT, J. G. HENGSTLER.*Model-guided identification of a therapeutic strategy to reduce hyperammonemia in liver diseases*, in "Journal of Hepatology", November 2015, vol. 64, n^{o} 4, p. 860–871 [*DOI :* 10.1016/J.JHEP.2015.11.018], https://hal.archives-ouvertes.fr/hal-01257127

[108] A. GHALLAB, U. HOFMANN, S. SEZGIN, N. VARTAK, R. HASSAN, A. ZAZA, P. GODOY, K. M. SCHNEIDER, G. GUENTHER, Y. A. AHMED, A. ABBAS, V. KEITEL, L. KUEPFER, S. DOOLEY, F. LAMMERT, C. TRAUTWEIN, M. SPITELLER, D. DRASDO, A. HOFMANN, P. L. JANSEN, J. HENGSTLER, R. REIF.*Bile Microinfarcts in Cholestasis Are Initiated by Rupture of the Apical Hepatocyte Membrane and Cause Shunting of Bile to Sinusoidal Blood*, in "Hepatology", August 2018, https://hal.inria.fr/hal-01968843

[109] L. GOSSE, N. VAUCHELET.*Hydrodynamic singular regimes in 1+1 kinetic models and spectral numerical methods*, in "Journal of Mathematical Analysis and Applications", 2016 [*DOI :* 10.1016/J.JMAA.2016.07.059], https://hal.archives-ouvertes.fr/hal-01354980

[110] J. HEMINGWAY, H. RANSON.*Insecticide resistance in insect vectors of human disease*, in "Annual review of entomology", 2000, vol. 45, n^{o} 1, p. 371–391

[111] M. HERTIG, S. B. WOLBACH.*Studies on rickettsia-like micro-organisms in insects*, in "The Journal of medical research", 1924, vol. 44, n^{o} 3, 329

[112] V. H. HOANG.*Adaptive estimation for inverse problems with applications to cell divisions*, Université de Lille 1 – Sciences et Technologies, November 2016, https://tel.archives-ouvertes.fr/tel-01417780

[113] S. HOEHME, M. BRULPORT, A. BAUER, E. BEDAWY, W. SCHORMANN, M. HERMES, V. PUPPE, R. GEBHARDT, S. ZELLMER, M. SCHWARZ, E. BOCKAMP, T. TIMMEL, J. G. HENGSTLER, D. DRASDO.*Prediction and validation of cell alignment along microvessels as order principle to restore tissue architecture in liver regeneration*, in "Proceedings of the National Academy of Sciences", 2010, vol. 107, n^{o} 23, p. 10371–10376

[114] M. HOFFMANN, A. OLIVIER.*Nonparametric estimation of the division rate of an age dependent branching process*, in "Stochastic Processes and their Applications", December 2015 [*DOI :* 10.1016/J.SPA.2015.11.009], https://hal.archives-ouvertes.fr/hal-01254203

[115] N. JAGIELLA, B. MÜLLER, M. MÜLLER, I. E. VIGNON-CLEMENTEL, D. DRASDO.*Inferring Growth Control Mechanisms in Growing Multi-cellular Spheroids of NSCLC Cells from Spatial-Temporal Image Data*, in "PLoS Computational Biology", 2016, vol. 12, n^{o} 2, e1004412 [*DOI :* 10.1371/JOURNAL.PCBI.1004412], http://hal.upmc.fr/hal-01244593

[116] F. JAMES, N. VAUCHELET.*Numerical methods for one-dimensional aggregation equations*, in "SIAM Journal on Numerical Analysis", 2015, vol. 53, n^{o} 2, p. 895-916 [*DOI :* 10.1137/140959997], https://hal.archives-ouvertes.fr/hal-00955971

[117] F. JAMES, N. VAUCHELET.*Equivalence between duality and gradient flow solutions for one-dimensional aggregation equations*, in "Discrete and Continuous Dynamical Systems - Series A", 2016, vol. 36, n^{o} 3, p. 1355-1382, https://hal.archives-ouvertes.fr/hal-00803709

[118] G. JANKOWIAK, D. PEURICHARD, A. REVERSAT, M. SIXT, C. SCHMEISER.*Modelling adhesion-independent cell migration*, in "Mathematical models and methods in applied sciences", 2019, https://doi.org/10.1142/S021820252050013X

[119] M.-J. KANG, B. PERTHAME, D. SALORT.*Dynamics of time elapsed inhomogeneous neuron network model*, in "Comptes Rendus Mathématique", September 2015, n^{o} 353, p. 1111-1115 [*DOI :* 10.1016/J.CRMA.2015.09.029], http://hal.upmc.fr/hal-01241300

[120] J. KEAN, S. RAINEY, M. MCFARLANE, C. DONALD, E. SCHNETTLER, A. KOHL, E. PONDEVILLE.*Fighting arbovirus transmission: natural and engineered control of vector competence in Aedes mosquitoes*, in "Insects", 2015, vol. 6, n^{o} 1, p. 236–278

[121] I. C. KIM, B. PERTHAME, P. E. SOUGANIDIS.*Free boundary problems for tumor growth: a viscosity solutions approach*, in "Nonlinear Analysis: Theory, Methods and Applications", 2016, vol. 138, p. 207-228 [*DOI :* 10.1016/J.NA.2016.01.019], http://hal.upmc.fr/hal-01155696

[122] M. KOLWALCZYK, B. PERTHAME, N. VAUCHELET.*Transversal instability for the thermodiffusive reaction-diffusion system*, in "Chinese Annals of Mathematics - Series B", 2015, vol. 36, n^{o} 5, p. 871-882, 13 pages, https://hal.archives-ouvertes.fr/hal-00939013

[123] P. V. LIEDEKERKE, J. NEITSCH, T. JOHANN, E. WARMT, I. GONZÀLEZ-VALVERDE, S. HOEHME, S. GROSSER, J. KAES, D. DRASDO.*A quantitative high resolution computational mechanics cell model for growing and regenerating tissues*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01956023

[124] T. LORENZI, R. H. CHISHOLM, J. CLAIRAMBAULT.*Tracking the evolution of cancer cell populations through the mathematical lens of phenotype-structured equations*, in "Biology Direct", December 2016, vol. 11, n^{o} 1, 43 [*DOI :* 10.1186/S13062-016-0143-4], https://hal.inria.fr/hal-01355314

[125] T. LORENZI, R. H. CHISHOLM, L. DESVILLETTES, B. D. HUGHES.*Dissecting the dynamics of epigenetic changes in phenotype-structured populations exposed to fluctuating environments*, in "Journal of Theoretical Biology", September 2015, vol. 386, p. 166-176 [*DOI :* 10.1016/J.JTBI.2015.08.031], https://hal.inria.fr/hal-01237890

[126] T. LORENZI, R. H. CHISHOLM, A. LORZ.*Effects of an advection term in nonlocal Lotka-Volterra equations*, December 2015, https://hal.archives-ouvertes.fr/hal-01237529

[127] A. LORZ, T. LORENZI, J. CLAIRAMBAULT, A. ESCARGUEIL, B. PERTHAME.*Modeling the effects of space structure and combination therapies on phenotypic heterogeneity and drug resistance in solid tumors*, in "Bulletin of Mathematical Biology", January 2015, vol. 77, n^{o} 1, p. 1-22 [*DOI :* 10.1007/S11538-014-0046-4], http://hal.upmc.fr/hal-00921266

[128] A. LORZ, T. LORENZI, M. E. HOCHBERG, J. CLAIRAMBAULT, B. PERTHAME.*Populational adaptive evolution, chemotherapeutic resistance and multiple anti-cancer therapies*, in "ESAIM: Mathematical Modelling and Numerical Analysis", March 2013, 23 [*DOI :* 10.1051/M2AN/2012031], https://hal.archives-ouvertes.fr/hal-00714274

[129] A. MELLET, B. PERTHAME, F. QUIROS.*A Hele-Shaw Problem for Tumor Growth*, December 2015, working paper or preprint, http://hal.upmc.fr/hal-01241309

[130] L. A. MOREIRA, I. ITURBE-ORMAETXE, J. A. JEFFERY, G. LU, A. T. PYKE, L. M. HEDGES, B. C. ROCHA, S. HALL-MENDELIN, A. DAY, M. RIEGLER, L. E. HUGO, K. N. JOHNSON, B. H. KAY, E. A. MCGRAW, A. F. VAN DEN HURK, P. A. RYAN, S. L. O'NEILL.*A Wolbachia Symbiont in Aedes aegypti Limits Infection with Dengue, Chikungunya, and Plasmodium*, in "Cell", 2009, vol. 139, n^{o} 7, p. 1268 - 1278

[131] L. NEVES DE ALMEIDA, F. BUBBA, B. PERTHAME, C. POUCHOL.*Energy and implicit discretization of the Fokker-Planck and Keller-Segel type equations*, October 2018, https://arxiv.org/abs/1803.10629 - pp:1-19 [*DOI :* 10.3934/XX.XX.XX.XX], https://hal.archives-ouvertes.fr/hal-01745769

[132] A. OLIVIER.*Statistical analysis of growth-fragmentation models*, Université Paris Dauphine - Paris IX, November 2015, https://hal.archives-ouvertes.fr/tel-01235239

[133] A. OLIVIER.*How does variability in cells aging and growth rates influence the malthus parameter?*, in "Kinetic and Related Models ", June 2017, vol. 10, n^{o} 2, p. 481-512 [*DOI :* 10.3934/KRM.2017019], https://hal.archives-ouvertes.fr/hal-01274529

[134] K. PAKDAMAN, B. PERTHAME, D. SALORT.*Adaptation and Fatigue Model for Neuron Networks and Large Time Asymptotics in a Nonlinear Fragmentation Equation*, in "Journal of Mathematical Neuroscience", 2014, vol. 4, n^{o} 1, 14 [*DOI :* 10.1186/2190-8567-4-14], https://hal.inria.fr/hal-01054561

[135] B. PERTHAME.*Transport equations in biology*, Frontiers in Mathematics, Birkhäuser Verlag, Basel, 2007, x+198

[136] B. PERTHAME, F. QUIRÓS, M. TANG, N. VAUCHELET.*Derivation of a Hele-Shaw type system from a cell model with active motion*, July 2013, http://hal.upmc.fr/hal-00906168

[137] B. PERTHAME, F. QUIRÓS, J.-L. VÁZQUEZ.*The Hele-Shaw asymptotics for mechanical models of tumor growth*, in "Archive for Rational Mechanics and Analysis", 2014, vol. 212, p. 93-127 [*DOI :* 10.1007/S00205-013-0704-Y], http://hal.upmc.fr/hal-00831932

[138] B. PERTHAME, D. SALORT.*On a voltage-conductance kinetic system for integrate and fire neural networks*, in "Kinetic and Related Models ", December 2013, vol. 6, n^{o} 4, p. 841-864 [*DOI :* 10.3934/KRM.2013.6.841], http://hal.upmc.fr/hal-00871609

[139] B. PERTHAME, D. SALORT, G. WAINRIB.*Distributed synaptic weights in a LIF neural network and learning rules*, in "Physica D: Nonlinear Phenomena", 2017, vol. 353-354, p. 20-30 [*DOI :* 10.1016/J.PHYSD.2017.05.005], http://hal.upmc.fr/hal-01541093

[140] B. PERTHAME, M. TANG, N. VAUCHELET.*Traveling wave solution of the Hele-Shaw model of tumor growth with nutrient*, in "Mathematical Models and Methods in Applied Sciences", 2014, vol. 24, n^{o} 13, p. 2601-2626, 25 pages, https://hal.archives-ouvertes.fr/hal-00931399

[141] B. PERTHAME, N. VAUCHELET.*Incompressible limit of mechanical model of tumor growth with viscosity*, in "Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences (1934–1990)", 2015, vol. 373, 20140283, 16 pages [*DOI :* 10.1098/RSTA.2014.0283], https://hal.archives-ouvertes.fr/hal-01066494

[142] B. PERTHAME, S. YASUDA.*Stiff-response-induced instability for chemotactic bacteria and flux-limited Keller-Segel equation*, in "Nonlinearity", 2018, vol. 31, n^o 9 [*DOI :* 10.1088/1361-6544/AAC760], https://hal.sorbonne-universite.fr/hal-01494963

[143] B. PERTHAME, J. ZUBELLI.*On the inverse problem for a size-structured population model*, in "Inverse Problems", 2007, vol. 23, n^o 3, p. 1037–1052

[144] D. E. A. PEURICHARD.*Simple mechanical cues could explain adipose tissue morphology*, in "J. Theor Biol", 2017, https://doi.org/10.1016/j.jtbi.2017.06.030

[145] D. E. A. PEURICHARD. *Extra-cellular matrix rigidity may dictate the fate of injury outcome*, in "J. Theor Biol", 2019, vol. 469, p. 127-136

[146] C. POUCHOL, J. CLAIRAMBAULT, A. LORZ, E. TRÉLAT.*Asymptotic analysis and optimal control of an integro-differential system modelling healthy and cancer cells exposed to chemotherapy*, December 2016, accepted in the J. Math. Pures et App., https://hal.archives-ouvertes.fr/hal-01416594

[147] S. PRIGENT, A. BALLESTA, F. CHARLES, N. LENUZZA, P. GABRIEL, L. M. TINE, H. REZAEI, M. DOUMIC.*An efficient kinetic model for assemblies of amyloid fibrils and its application to polyglutamine aggregation.*, in "PLoS ONE", 2012, vol. 7, n^o 11, e43273 [*DOI :* 10.1371/JOURNAL.PONE.0043273], https://hal.archives-ouvertes.fr/hal-00778052

[148] L. ROBERT, M. HOFFMANN, N. KRELL, S. AYMERICH, J. ROBERT, M. DOUMIC.*Division in Escherichia coli is triggered by a size-sensing rather than a timing mechanism*, in "BMC Biology", 2014, vol. 12, n^o 1, 17 [*DOI :* 10.1186/1741-7007-12-17], https://hal.inria.fr/hal-00981312

[149] F. SCHLIESS, S. HOEHME, S. HENKEL, A. GHALLAB, D. DRIESCH, J. BÖTTGER, R. GUTHKE, M. PFAFF, J. HENGSTLER, R. GEBHARDT, D. HÄUSSINGER, D. DRASDO, S. ZELLMER.*Integrated metabolic spatial-temporal model for the prediction of ammonia detoxification during liver damage and regeneration*, in "Hepatology", December 2014, vol. 60, n^o 6, p. 2040–2051 [*DOI :* 10.1002/HEP.27136], https://hal.inria.fr/hal-01110646

[150] N. SFAKIANAKIS, D. PEURICHARD, A. BRUNK, C. SCHMEISER.*Modelling cell-cell collision and adhesion with the Filament Based Lamellipodium Model*, in "BIOMATH", 2019, http://dx.doi.org/10.11145/j.biomath.2018.11.097

[151] S. P. SINKINS.*Wolbachia and cytoplasmic incompatibility in mosquitoes*, in "Insect Biochemistry and Molecular Biology", 2004, vol. 34, n^o 7, p. 723 - 729, Molecular and population biology of mosquitoes

[152] M. STRUGAREK, H. BOSSIN, Y. DUMONT.*On the use of the sterile insect technique or the incompatible insect technique to reduce or eliminate mosquito populations*, May 2018, https://arxiv.org/abs/1805.10150 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01799954

[153] M. STRUGAREK, H. JI.*Sharp seasonal threshold property for cooperative population dynamics with concave nonlinearities*, April 2018, https://arxiv.org/abs/1804.07641 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01772628

[154] M. STRUGAREK.*Mathematical modeling of population dynamics, applications to vector control of Aedes spp. (Diptera:Culicidae)*, Sorbonne Université , UPMC, September 2018, https://hal.archives-ouvertes.fr/tel-01879201

[155] P. VAN LIEDEKERKE, J. NEITSCH, T. JOHANN, K. ALESSANDRI, P. NASSOY, D. DRASDO.*Quantitative modeling identifies robust predictable stress response of growing CT26 tumor spheroids under variable conditions*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01421179

[156] T. WALKER, S. P. SINKINS.*Biological control of arbovirus vectors*, in "Arboviruses: Molecular Biology, Evolution and Control. Caister Academic Press, Norfolk, UK", 2016, p. 291–302

[157] Y. YIN, O. SEDLACZEK, B. MÜLLER, A. WARTH, M. GONZÁLEZ-VALLINAS, B. LAHRMANN, N. GRABE, H.-U. KAUCZOR, K. BREUHAHN, I. VIGNON-CLEMENTEL, D. DRASDO.*Tumor cell load and heterogeneity estimation from diffusion-weighted MRI calibrated with histological data: an example from lung cancer*, in "IEEE Transactions on Medical Imaging", 2017 [*DOI :* 10.1109/TMI.2017.2698525], https://hal.inria.fr/hal-01421398

Project-Team MATHERIALS

MATHematics for MatERIALS

IN COLLABORATION WITH: Centre d'Enseignement et de Recherche en Mathématiques et Calcul Scientifique (CERMICS)

IN PARTNERSHIP WITH:
Ecole des Ponts ParisTech

RESEARCH CENTER
Paris

THEME
Numerical schemes and simulations

Table of contents

Project-Team MATHERIALS

Creation of the Team: 2014 January 01, updated into Project-Team: 2015 April 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
A6.2.2. - Numerical probability
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.2.7. - High performance computing
A6.3.1. - Inverse problems
A6.3.4. - Model reduction
A6.4.1. - Deterministic control

Other Research Topics and Application Domains:

B1.1.2. - Molecular and cellular biology
B4.3.4. - Solar Energy
B5.3. - Nanotechnology
B5.5. - Materials
B9.5.2. - Mathematics
B9.5.3. - Physics
B9.5.4. - Chemistry

1. Team, Visitors, External Collaborators

Research Scientists

Claude Le Bris [Team leader, Ecole Nationale des Ponts et Chaussées, Senior Researcher, HDR]
Sébastien Boyaval [Ecole Nationale des Ponts et Chaussées, Senior Researcher, HDR]
Éric Cancès [Ecole Nationale des Ponts et Chaussées, Senior Researcher, HDR]
Virginie Ehrlacher [Ecole Nationale des Ponts et Chaussées, Researcher]
Frédéric Legoll [Ecole Nationale des Ponts et Chaussées, Senior Researcher, HDR]
Tony Lelièvre [Ecole Nationale des Ponts et Chaussées, Senior Researcher, HDR]
Antoine Levitt [Inria, Researcher]
Gabriel Stoltz [Ecole Nationale des Ponts et Chaussées, Senior Researcher, HDR]

Faculty Member

Xavier Blanc [Univ Denis Diderot, Professor, *en délégation* from Sep 2019, HDR]

PhD Students

Robert Benda [Ecole Nationale des Ponts et Chaussées]
Qiming Du [Inria, from Oct 2019]
Grégoire Ferré [Ecole Nationale des Ponts et Chaussées, until Sep 2019]
Rémi Goudey [Ecole Nationale des Ponts et Chaussées, from Sep 2019]

Gaspard Kemlin [Ecole Nationale des Ponts et Chaussées, from Sep 2019]
Adrien Lesage [Ecole Nationale des Ponts et Chaussées]
Mouad Ramil [Ecole Nationale des Ponts et Chaussées]
Pierre-Loïk Rothé [Ecole Nationale des Ponts et Chaussées]
Laura Silva Lopes [Ecole Nationale des Ponts et Chaussées until Sep 2019, Inria Oct-Dec 2019]
Sami Siraj-Dine [Univ Paris-Est Marne La Vallée]

Post-Doctoral Fellows

Manon Baudel [Ecole Nationale des Ponts et Chaussées, until Jun 2019]
Olga Gorynina [Ecole Nationale des Ponts et Chaussées]
Michael Herbst [Sorbonne Université]
Geneviève Robin [Inria, from Oct 2019]

2. Overall Objectives

2.1. Overall Objectives

The MATHERIALS project-team has been created jointly by the École des Ponts ParisTech (ENPC) and Inria in 2015. It is the follow-up and an extension of the former project-team MICMAC originally created in October 2002. It is hosted by the CERMICS laboratory (Centre d'Enseignement et de Recherches en Mathématiques et Calcul Scientifique) at École des Ponts. The permanent research scientists of the project-team have positions at CERMICS and at two other laboratories of École des Ponts: Institut Navier and Laboratoire Saint-Venant. The scientific focus of the project-team is to analyze and improve the numerical schemes used in the simulation of computational chemistry at the microscopic level and to create simulations coupling this microscopic scale with meso- or macroscopic scales (possibly using parallel algorithms). Over the years, the project-team has accumulated an increasingly solid expertise on such topics, which are traditionally not well known by the community in applied mathematics and scientific computing. One of the major achievements of the project-team is to have created a corpus of literature, authoring books and research monographs on the subject [1], [2], [3], [4], [5], [6], [7] that other scientists may consult in order to enter the field.

3. Research Program

3.1. Research Program

Our group, originally only involved in electronic structure computations, continues to focus on many numerical issues in quantum chemistry, but now expands its expertise to cover several related problems at larger scales, such as molecular dynamics problems and multiscale problems. The mathematical derivation of continuum energies from quantum chemistry models is one instance of a long-term theoretical endeavour.

3.1.1. Electronic structure of large systems

Quantum Chemistry aims at understanding the properties of matter through the modelling of its behavior at a subatomic scale, where matter is described as an assembly of nuclei and electrons. At this scale, the equation that rules the interactions between these constitutive elements is the Schrödinger equation. It can be considered (except in few special cases notably those involving relativistic phenomena or nuclear reactions) as a universal model for at least three reasons. First it contains all the physical information of the system under consideration so that any of the properties of this system can in theory be deduced from the Schrödinger equation associated to it. Second, the Schrödinger equation does not involve any empirical parameters, except some fundamental constants of Physics (the Planck constant, the mass and charge of the electron, ...); it can thus be written for any kind of molecular system provided its chemical composition, in terms of natures of nuclei and number of electrons, is known. Third, this model enjoys remarkable predictive capabilities, as confirmed by comparisons with a large amount of experimental data of various types. On the other hand, using this high quality model requires working with space and time scales which are both very tiny: the typical size of the electronic cloud of an isolated atom is the Angström (10^{-10} meters), and the size of the nucleus embedded in it is 10^{-15} meters; the typical vibration period of a molecular bond is the femtosecond (10^{-15} seconds), and the characteristic relaxation time for an electron is 10^{-18} seconds. Consequently, Quantum Chemistry calculations concern very short time (say 10^{-12} seconds) behaviors of very small size (say 10^{-27} m^3) systems. The underlying question is therefore whether information on phenomena at these scales is useful in understanding or, better, predicting macroscopic properties of matter. It is certainly not true that *all* macroscopic properties can be simply upscaled from the consideration of the short time behavior of a tiny sample of matter. Many of them derive from ensemble or bulk effects, that are far from being easy to understand and to model. Striking examples are found in solid state materials or biological systems. Cleavage, the ability of minerals to naturally split along crystal surfaces (e.g. mica yields to thin flakes), is an ensemble effect. Protein folding is also an ensemble effect that originates from the presence of the surrounding medium; it is responsible for peculiar properties (e.g. unexpected acidity of some reactive site enhanced by special interactions) upon which vital processes are based. However, it is undoubtedly true that *many* macroscopic phenomena originate from elementary processes which take place at the atomic scale. Let us mention for instance the fact that the elastic constants of a perfect crystal or the color of a chemical compound (which is related to the wavelengths absorbed or emitted during optic transitions between electronic levels) can be evaluated by atomic scale calculations. In the same fashion, the lubricative properties of graphite are essentially due to a phenomenon which can be entirely modeled at the atomic scale. It is therefore reasonable to simulate the behavior of matter at the atomic scale in order to understand what is going on at the macroscopic one. The journey is however a long one. Starting from the basic principles of Quantum Mechanics to model the matter at the subatomic scale, one finally uses statistical mechanics to reach the macroscopic scale. It is often necessary to rely on intermediate steps to deal with phenomena which take place on various *mesoscales*. It may then be possible to couple one description of the system with some others within the so-called *multiscale* models. The sequel indicates how this journey can be completed focusing on the first smallest scales (the subatomic one), rather than on the larger ones. It has already been mentioned that at the subatomic scale, the behavior of nuclei and electrons is governed by the Schrödinger equation, either in its time-dependent form or in its time-independent form. Let us only mention at this point that

- both equations involve the quantum Hamiltonian of the molecular system under consideration; from a mathematical viewpoint, it is a self-adjoint operator on some Hilbert space; *both* the Hilbert space and the Hamiltonian operator depend on the nature of the system;
- also present into these equations is the wavefunction of the system; it completely describes its state; its L^2 norm is set to one.

The time-dependent equation is a first-order linear evolution equation, whereas the time-independent equation is a linear eigenvalue equation. For the reader more familiar with numerical analysis than with quantum mechanics, the linear nature of the problems stated above may look auspicious. What makes the numerical simulation of these equations extremely difficult is essentially the huge size of the Hilbert space: indeed, this space is roughly some symmetry-constrained subspace of $L^2(\mathbb{R}^d)$, with $d = 3(M+N)$, M and N respectively denoting the number of nuclei and the number of electrons the system is made of. The parameter d is already 39 for a single water molecule and rapidly reaches 10^6 for polymers or biological molecules. In addition, a consequence of the universality of the model is that one has to deal at the same time with several energy scales. In molecular systems, the basic elementary interaction between nuclei and electrons (the two-body Coulomb interaction) appears in various complex physical and chemical phenomena whose characteristic energies cover several orders of magnitude: the binding energy of core electrons in heavy atoms is 10^4 times as large as a typical covalent bond energy, which is itself around 20 times as large as the energy of a hydrogen bond. High precision or at least controlled error cancellations are thus required to reach chemical accuracy when starting from the Schrödinger equation. Clever approximations of the Schrödinger problems are therefore needed. The main two approximation strategies, namely the Born-Oppenheimer-Hartree-Fock and the Born-Oppenheimer-Kohn-Sham strategies, end up with large systems of coupled *nonlinear* partial differential equations, each of these equations being posed on $L^2(\mathbb{R}^3)$. The size of the underlying functional space is thus reduced at the cost of a dramatic increase of the mathematical complexity of the problem:

nonlinearity. The mathematical and numerical analysis of the resulting models has been the major concern of the project-team for a long time. In the recent years, while part of the activity still follows this path, the focus has progressively shifted to problems at other scales.

As the size of the systems one wants to study increases, more efficient numerical techniques need to be resorted to. In computational chemistry, the typical scaling law for the complexity of computations with respect to the size of the system under study is N^3, N being for instance the number of electrons. The Holy Grail in this respect is to reach a linear scaling, so as to make possible simulations of systems of practical interest in biology or material science. Efforts in this direction must address a large variety of questions such as

- how can one improve the nonlinear iterations that are the basis of any *ab initio* models for computational chemistry?
- how can one more efficiently solve the inner loop which most often consists in the solution procedure for the linear problem (with frozen nonlinearity)?
- how can one design a sufficiently small variational space, whose dimension is kept limited while the size of the system increases?

An alternative strategy to reduce the complexity of *ab initio* computations is to try to couple different models at different scales. Such a mixed strategy can be either a sequential one or a parallel one, in the sense that

- in the former, the results of the model at the lower scale are simply used to evaluate some parameters that are inserted in the model for the larger scale: one example is the parameterized classical molecular dynamics, which makes use of force fields that are fitted to calculations at the quantum level;
- while in the latter, the model at the lower scale is concurrently coupled to the model at the larger scale: an instance of such a strategy is the so called QM/MM coupling (standing for Quantum Mechanics/Molecular Mechanics coupling) where some part of the system (typically the reactive site of a protein) is modeled with quantum models, that therefore accounts for the change in the electronic structure and for the modification of chemical bonds, while the rest of the system (typically the inert part of a protein) is coarse grained and more crudely modeled by classical mechanics.

The coupling of different scales can even go up to the macroscopic scale, with methods that couple a microscopic representation of matter, or at least a mesoscopic one, with the equations of continuum mechanics at the macroscopic level.

3.1.2. Computational Statistical Mechanics

The orders of magnitude used in the microscopic representation of matter are far from the orders of magnitude of the macroscopic quantities we are used to: The number of particles under consideration in a macroscopic sample of material is of the order of the Avogadro number $\mathcal{N}_A \sim 6 \times 10^{23}$, the typical distances are expressed in Å (10^{-10} m), the energies are of the order of $k_{\rm B}T \simeq 4 \times 10^{-21}$ J at room temperature, and the typical times are of the order of 10^{-15} s.

To give some insight into such a large number of particles contained in a macroscopic sample, it is helpful to compute the number of moles of water on earth. Recall that one mole of water corresponds to 18 mL, so that a standard glass of water contains roughly 10 moles, and a typical bathtub contains 10^5 mol. On the other hand, there are approximately 10^{18} m^3 of water in the oceans, *i.e.*7×10^{22} mol, a number comparable to the Avogadro number. This means that inferring the macroscopic behavior of physical systems described at the microscopic level by the dynamics of several millions of particles only is like inferring the ocean's dynamics from hydrodynamics in a bathtub...

For practical numerical computations of matter at the microscopic level, following the dynamics of every atom would require simulating $\mathcal{N}_A$ atoms and performing $\mathrm{O}(10^{15})$ time integration steps, which is of course impossible! These numbers should be compared with the current orders of magnitude of the problems that can be tackled with classical molecular simulation, where several millions of atoms only can be followed over time scales of the order of a few microseconds.

Describing the macroscopic behavior of matter knowing its microscopic description therefore seems out of reach. Statistical physics allows us to bridge the gap between microscopic and macroscopic descriptions of matter, at least on a conceptual level. The question is whether the estimated quantities for a system of N particles correctly approximate the macroscopic property, formally obtained in the thermodynamic limit $N \to +\infty$ (the density being kept fixed). In some cases, in particular for simple homogeneous systems, the macroscopic behavior is well approximated from small-scale simulations. However, the convergence of the estimated quantities as a function of the number of particles involved in the simulation should be checked in all cases.

Despite its intrinsic limitations on spatial and timescales, molecular simulation has been used and developed over the past 50 years, and its number of users keeps increasing. As we understand it, it has two major aims nowadays.

First, it can be used as a *numerical microscope*, which allows us to perform "computer" experiments. This was the initial motivation for simulations at the microscopic level: physical theories were tested on computers. This use of molecular simulation is particularly clear in its historic development, which was triggered and sustained by the physics of simple liquids. Indeed, there was no good analytical theory for these systems, and the observation of computer trajectories was very helpful to guide the physicists' intuition about what was happening in the system, for instance the mechanisms leading to molecular diffusion. In particular, the pioneering works on Monte-Carlo methods by Metropolis *et al.*, and the first molecular dynamics simulation of Alder and Wainwright were performed because of such motivations. Today, understanding the behavior of matter at the microscopic level can still be difficult from an experimental viewpoint (because of the high resolution required, both in time and in space), or because we simply do not know what to look for! Numerical simulations are then a valuable tool to test some ideas or obtain some data to process and analyze in order to help assessing experimental setups. This is particularly true for current nanoscale systems.

Another major aim of molecular simulation, maybe even more important than the previous one, is to compute macroscopic quantities or thermodynamic properties, typically through averages of some functionals of the system. In this case, molecular simulation is a way to obtain *quantitative* information on a system, instead of resorting to approximate theories, constructed for simplified models, and giving only qualitative answers. Sometimes, these properties are accessible through experiments, but in some cases only numerical computations are possible since experiments may be unfeasible or too costly (for instance, when high pressure or large temperature regimes are considered, or when studying materials not yet synthesized). More generally, molecular simulation is a tool to explore the links between the microscopic and macroscopic properties of a material, allowing one to address modelling questions such as "Which microscopic ingredients are necessary (and which are not) to observe a given macroscopic behavior?"

3.1.3. Homogenization and related problems

Over the years, the project-team has developed an increasing expertise on how to couple models written at the atomistic scale with more macroscopic models, and, more generally, an expertise in multiscale modelling for materials science.

The following observation motivates the idea of coupling atomistic and continuum representation of materials. In many situations of interest (crack propagation, presence of defects in the atomistic lattice, ...), using a model based on continuum mechanics is difficult. Indeed, such a model is based on a macroscopic constitutive law, the derivation of which requires a deep qualitative and quantitative understanding of the physical and mechanical properties of the solid under consideration. For many solids, reaching such an understanding is a challenge, as loads they are subjected to become larger and more diverse, and as experimental observations helping designing such models are not always possible (think of materials used in the nuclear industry). Using an atomistic model in the whole domain is not possible either, due to its prohibitive computational cost. Recall indeed that a macroscopic sample of matter contains a number of atoms on the order of 10^{23}. However, it turns out that, in many situations of interest, the deformation that we are looking for is not smooth in *only a small part* of the solid. So, a natural idea is to try to take advantage of both models, the continuum mechanics one and the atomistic one, and to couple them, in a domain decomposition spirit. In most of the domain, the deformation is expected to be smooth, and reliable continuum mechanics models are then available. In the rest of the domain, the expected deformation is singular, so that one needs an atomistic model to describe it properly, the cost of which remains however limited as this region is small.

From a mathematical viewpoint, the question is to couple a discrete model with a model described by PDEs. This raises many questions, both from the theoretical and numerical viewpoints:

- first, one needs to derive, from an atomistic model, continuum mechanics models, under some regularity assumptions that encode the fact that the situation is smooth enough for such a macroscopic model to provide a good description of the materials;
- second, couple these two models, e.g. in a domain decomposition spirit, with the specificity that models in both domains are written in a different language, that there is no natural way to write boundary conditions coupling these two models, and that one would like the decomposition to be self-adaptive.

More generally, the presence of numerous length scales in material science problems represents a challenge for numerical simulation, especially when some *randomness* is assumed on the materials. It can take various forms, and includes defects in crystals, thermal fluctuations, and impurities or heterogeneities in continuous media. Standard methods available in the literature to handle such problems often lead to very costly computations. Our goal is to develop numerical methods that are more affordable. Because we cannot embrace all difficulties at once, we focus on a simple case, where the fine scale and the coarse-scale models can be written similarly, in the form of a simple elliptic partial differential equation in divergence form. The fine scale model includes heterogeneities at a small scale, a situation which is formalized by the fact that the coefficients in the fine scale model vary on a small length scale. After homogenization, this model yields an effective, macroscopic model, which includes no small scale. In many cases, a sound theoretical groundwork exists for such homogenization results. The difficulty stems from the fact that the models generally lead to prohibitively costly computations. For such a case, simple from the theoretical viewpoint, our aim is to focus on different practical computational approaches to speed-up the computations. One possibility, among others, is to look for specific random materials, relevant from the practical viewpoint, and for which a dedicated approach can be proposed, that is less expensive than the general approach.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Grants

- Eric Cancès, Laura Grigori (ALPINES, Inria Paris), Yvon Maday (Sorbonne Université), and Jean-Philip Piquemal (Sorbonne Université) are the PIs of the ERC Synergy project EMC2 (Extreme-scale Mathematically-based Computational Chemistry) launched in September 2019 (grant agreement No 810367).
- Virginie Ehrlacher is the PI of the ANR JCJC COMODO (CrOss-diffusion systems on MOving DOmains) project, to start in January 2020.

4.1.2. Awards

- C. Le Bris was a plenary speaker at ICIAM 2019, July 2019, Valencia, Spain.
- T. Lelièvre has received a Visiting professorship from the Leverhulme Trust, for his sabbatical leave at Imperial College London (Sep 2019-Jul 2020).

5. New Software and Platforms

5.1. DFTK

KEYWORDS: Molecular simulation - Quantum chemistry - Materials

FUNCTIONAL DESCRIPTION: DFTK, short for the density-functional toolkit, is a Julia library implementing plane-wave density functional theory for the simulation of the electronic structure of molecules and materials. It aims at providing a simple platform for experimentation and algorithm development for scientists of different backgrounds.

- Contact: Antoine Levitt
- URL: http://dftk.org

5.2. gen.parRep

KEYWORDS: Molecular simulation - MPI - HPC - C++

SCIENTIFIC DESCRIPTION: Metastability is one of the major encountered obstacle when performing long molecular dynamics simulations, and many methods were developed to address this challenge. The "Parallel Replica" (ParRep) dynamics is known for allowing to simulate very long trajectories of metastable Langevin dynamics in the materials science community, but it relies on assumptions that can hardly be transposed to the world of biochemical simulations. The later developed "Generalized ParRep" variant solves those issues, but it was not applied to significant systems of interest so far.

In a recent article article, we presented the program gen.parRep, the first publicly available implementation of the Generalized Parallel Replica method (BSD 3-Clause license), targeting frequently encountered metastable biochemical systems, such as conformational equilibria or dissociation of protein-ligand complexes. It was shown that the resulting C++ implementation exhibits a strong linear scalability, providing up to 70 % of the maximum possible speedup on several hundreds of CPUs.

FUNCTIONAL DESCRIPTION: gen.parRep is the first publicly available implementation of the Generalized Parallel Replica method (BSD 3-Clause license), targeting frequently encountered metastable biochemical systems, such as conformational equilibria or dissociation of protein-ligand complexes.

It was shown (https://hal.archives-ouvertes.fr/hal-01832823) that the resulting C++/MPI implementation exhibits a strong linear scalability, providing up to 70 % of the maximum possible speedup on several hundreds of CPUs.

RELEASE FUNCTIONAL DESCRIPTION: The software was modified in order to allow reproducibility in some cases (the limiting factor is OpenMM which does not always provides deterministic output even when using the same seeds (!), see http://docs.openmm.org/latest/userguide/library.html#determinism).

The main executable now has 2 more command line options, '–inp-seeds [fname]' or '–out-seeds [fname]' for respectively loading seeds or writing seeds from/to a unique binary file. See rand.hpp/rand.cpp, od the Doxygen doc for more details.

These modifications now allow Continuous Integration (CI) on the infrastructure provided by Inria : in mol/ci a small test case will be executed at each commit to the repository and compared to reference results.

The two other minor modifications concern the Lua scripts:

"get_minimised_energy_crdvels" was added to the set of functions that the user can call from the Lua script : it simply combines in one call what "get_minimised_energy" and "get_minimised_crdvels" already provided.

a extra optional parameter is available for the "simulation" parameters block when "simulation.algorithm" is "PARREP_FV" , this parameter is "simulation.minAccumulatedObs" : it will enforce that at least minAccumulatedObs observations of an observable have already been accumulated before the convergence test is performed , this may be useful if there is a risk of early pseudo-convergence for some of the observables when only a few samples have been accumulated.

Download sources:

https://gitlab.inria.fr/parallel-replica/gen.parRep/tags/v1.2.0

or

https://github.com/FHedin/gen.parRep/releases/tag/v1.2.0

NEWS OF THE YEAR: Corresponding publication : https://hal.archives-ouvertes.fr/hal-01832823

- Participants: Florent Hedin and Tony Lelièvre
- Partner: Ecole des Ponts ParisTech
- Contact: Florent Hedin
- Publication: gen.parRep: a first implementation of the Generalized Parallel Replica dynamics for the long time simulation of metastable biochemical systems
- URL: https://gitlab.inria.fr/parallel-replica/gen.parRep

6. New Results

6.1. Electronic structure calculations and related problems

Participants: Robert Benda, Éric Cancès, Virginie Ehrlacher, Luca Gorini, Gaspard Kemlin, Claude Le Bris, Antoine Levitt, Sami Siraj-Dine, Gabriel Stoltz.

6.1.1. Mathematical analysis

The members of the team have continued their systematic study of the properties of materials in the reduced Hartree-Fock (rHF) approximation, a model striking a good balance between mathematical tractability and the ability to reproduce qualitatively complex effects.

In collaboration with L. Cao, E. Cancès and G. Stoltz have studied the nuclear dynamics of infinite crystals with local defects within the Born-Oppenheimer approximation, using the reduced Hartree-Fock model to compute the electronic ground state. In this model, nuclei obey an autonomous classical Hamiltonian dynamics on a potential energy surface obtained by rHF electronic ground-state calculations. One of the main motivations for this work is to study the *nonlinear* collective excitations of nuclei in a crystal, in order to go beyond the simple harmonic approximation of non-interacting phonons. rHF ground states associated with generic nuclear displacements with respect to the periodic configuration are not mathematically well-defined at the time of writing. However, by relying on results by Cancès, Deleurence and Lewin for the rHF ground states of crystals with local defects, it is possible to study the fully nonlinear rHF Born-Oppenheimer dynamics of nuclei in the neighborhood of an equilibrium periodic configuration of a crystal. A Hilbert space of admissible nuclear displacements, and an infinite-dimensional Hamiltonian describing the dynamics of nuclei can then be defined. For small initial data, it is proved that the Cauchy problem associated with this Hamiltonian dynamics is well posed for short times (see the PhD thesis of Lingling Cao). The existence and uniqueness for arbitrary initial data, and/or long times requires a perturbation analysis of the rHF model when the Fermi level is occupied, which is work in progress.

6.1.2. Numerical analysis

E. Cancès has pursued his long-term collaboration with Y. Maday (Sorbonne Université) on the numerical analysis of linear and nonlinear eigenvalue problems. Together with G. Dusson (Besançon), B. Stamm (Aachen, Germany), and M. Vohralìk (Inria SERENA), they have designed *a posteriori* error estimates for conforming numerical approximation of eigenvalue clusters of second-order self-adjoint operators on bounded domains [44]. Given a cluster of eigenvalues, they have estimated the error in the sum of the eigenvalues, as well as the error in the eigenvectors represented through the density matrix, i.e. the orthogonal projector on the associated eigenspace. This allows them to deal with degenerate (multiple) eigenvalues within this framework. The bounds are guaranteed and converge at the same rate as the exact error. They can be turned into fully computable bounds as soon as an estimate on the dual norm of the residual is available, which is notably the case (i) for the Laplace eigenvalue problem discretized with conforming finite elements, and (ii) for a Schrödinger operator with periodic boundary conditions discretized with plane waves.

R. Benda, E. Cancès and B. Lebental (Ecole Polytechnique) have initiated the design and analysis of multiscale models for the electrical conductivity of networks of functionalized carbon nanotubes. Such devices are used as nanosensors, for instance to monitor the quality of water. In [11], they study by means of Monte-Carlo numerical simulations the resistance of two-dimensional random percolating networks of stick, widthless nanowires. They use the multi-nodal representation (MNR) to model a nanowire network as a graph. They derive numerically from this model the expression of the total resistance as a function of all meaningful parameters, geometrical and physical, over a wide range of variation for each. They justify their choice of non-dimensional variables applying Buckingham π-theorem. The effective resistance of 2D random percolating networks of nanowires is found to have a nice expression in terms of the geometrical parameters (number of wires, aspect ratio of electrode separation over wire length) and the physical parameters (nanowire linear resistance per unit length, nanowire/nanowire contact resistance, metallic electrode/nanowire contact resistance). The dependence of the resistance on the geometry of the network, on the one hand, and on the physical parameters (values of the resistances), on the other hand, is thus clearly separated thanks to this expression, much simpler than the previously reported analytical expressions. In parallel, atomic scale models based on electronic structure theory are being developed to parameterize these mesoscale models (PhD thesis of R. Benda).

C. Le Bris has pursued his long term collaboration with Pierre Rouchon (Ecole des Mines de Paris and Inria QUANTIC) on the study of high dimensional Lindblad type equations at play in the modelling of open quantum systems. They have co-supervised the M2 internship of Luca Gorini, that was focused on the simulation of some simple quantum gates, and has investigated several discretization strategies based upon the choice of suitable basis sets.

V. Ehrlacher, L. Grigori (Inria ALPINES), D. Lombardi (Inria COMMEDIA) and H. Song (Inria ALPINES) have designed a new numerical method for the compression of high-order tensors [49]. The principle of the algorithm consists in constructing an optimal partition of the set of indices of the tensor, and construct an approximation of the tensor on each indices subdomain by means of an adapted High-Order Singular Value Decomposition. This method was used, among other examples, for the reduction of the solution of the Vlasov-Poisson system, and enabled to reach very significant compression factors. They also obtained very encouraging results on the compression of the Coulomb potential, which could be very interesting with a view to the resolution of the time-dependent Schrödinger equation in high dimension, which is currently work in progress.

A. Alfonsi, R. Coyaud (Ecole des Ponts), V. Ehrlacher and D. Lombardi (Inria COMMEDIA) studied a different approach for discretizing optimal transport problems, which relies in relaxing the marginal constraints in a finite number of marginal moment constraints, while keeping an infinite state space [40]. The advantage of such an approach is that the approximate solution of the multi-marginal optimal transport problem with Coulomb cost, which is the semi-classical limit of the so-called Lévy-Lieb functional, can be represented as a discrete measure charging a low number of points, thus avoiding the curse of dimensionality when the number of electrons is large.

M. Herbst and his collaborators have developed the `adcc` Python/C++ software package for performing excited state calculations based on algebraic-diagrammatic construction methods. It connects to four SCF packages (pyscf, psifour, molsturm and veloxchem), allows the inclusion of environmental effects through implicit or explicit solvent models, and implements methods up to third order in perturbation theory. Its features are summarized in [54].

6.2. Computational Statistical Physics

Participants: Manon Baudel, Qiming Du, Grégoire Ferré, Frédéric Legoll, Tony Lelièvre, Mouad Ramil, Geneviève Robin, Laura Silva Lopes, Gabriel Stoltz.

The objective of computational statistical physics is to compute macroscopic properties of materials starting from a microscopic description, using concepts of statistical physics (thermodynamic ensembles and molecular dynamics). The contributions of the team can be divided into four main topics: (i) the development of methods for sampling the configuration space; (ii) the efficient computation of dynamical properties which requires to sample metastable trajectories; (iii) the simulation of nonequilibrium systems and the computation of transport coefficients; (iv) coarse-graining techniques to reduce the computational cost of molecular dynamic simulations and gain some insights on the models.

6.2.1. Sampling of the configuration space: new algorithms and applications

The work [52] by G. Ferré and G. Stoltz considers fluctuations of empirical averages for stochastic differential equations. Such averages are commonly used to compute ergodic averages in statistical physics in order to estimate macroscopic quantities, but they are subject to fluctuations. If small deviations are described by the central limit theorem, important fluctuations enter the large deviations framework. This theory is well understood when considering bounded observables of a stochastic differential equation, but quantities of interest are generally unbounded. The authors identify the class of unbounded functions which enter the "usual" regime of large deviations. The answer is not trivial, and suggests that many physical observables satisfy another type of large deviations, which leads to further works. Additionally, the influence of irreversibility on the fluctuations was studied by providing a mathematical illustration of the second law of thermodynamics, stating that irreversible dynamics generate more entropy, or more disorder, than reversible ones.

The team also pursued its endeavour to study and improve free energy biasing techniques, such as adaptive biasing force or metadynamics. The gist of these techniques is to bias the original metastable dynamics used to sample the target probability measure in the configuration space by an approximation of the free energy along well-chosen reaction coordinates. This approximation is built on the fly, using empirical measures over replicas, or occupations measures over the trajectories of the replicas. Two works have been performed on such methods

- First, in [50], V. Ehrlacher, T. Lelièvre and P. Monmarché (Sorbonne Université) have developed a new numerical method in order to compute the free energy and the biased potential given by the Adaptive Biasing Force method in the case where the number of reaction cordinates in the system is too large to apply standard grid-based approximation techniques. The algorithm uses a greedy algorithm and a tensor product approximation. Convergence proofs of both the underlying ABF technique (which uses an unbiased occupation measure) and the greedy tensor-product approximation are provided.
- Second, in [55], T. Lelièvre together with B. Jourdain (Ecole des Ponts) and P.-A. Zitt (Université Paris Est) have used a parallel between metadynamics and self interacting models for polymers to study the longtime convergence of the original metadynamics algorithm in the adiabatic setting, namely when the dynamics along the collective variables decouple from the dynamics along the other degrees of freedom. The bias which is introduced when the adiabatic assumption does not hold is also discussed.

The team has also considered new applications in terms of sampling, and the analysis of related sampling methods:

- For large scale Bayesian inference, B. Leimkuhler (Edinburgh, United Kingdom), M. Sachs (Duke, USA) and G. Stoltz have studied in [57] the convergence of Adaptive Langevin dynamics, which is a method for sampling the Boltzmann-Gibbs distribution at a prescribed temperature in cases where the potential gradient is subject to stochastic perturbation of unknown magnitude. The method replaces the friction in underdamped Langevin dynamics with a dynamical variable, updated according to a negative feedback loop control law as in the Nose-Hoover thermostat. Hypocoercive techniques allow to show that the law of Adaptive Langevin dynamics converges exponentially rapidly to the stationary distribution, with a rate that can be quantified in terms of the key parameters of the dynamics. This implies in particular that a central limit theorem holds for the time averages computed along a stochastic path.
- For the simulation of log-gases, G. Ferré and G. Stoltz have studied in [46] with D. Chafaï (Université Paris Dauphine) a follow up to a former project on the efficient simulation of Coulomb and logarithmic gases. A previous work has demonstrated the usefulness of Hybrid Monte Carlo

techniques for sampling the invariant measure of such gases. Gases under constraint are now considered. First, the algorithm proposed in [31] was used to numerically explore the situation. Then, large deviations techniques were employed to study the limiting behaviour of the conditioned gas when the number of particles gets large. For a class of constraints, the equation solved by the limiting empirical density shows in particular cases a spectacular behaviour. This work suggests to further explore some research paths, such as the limiting distribution for large constraints.

6.2.2. Sampling of dynamical properties and rare events

In the preprint [48], T. Lelièvre uses the quasi-stationary distribution approach to study the first exit point distribution from a bounded domain of the overdamped Langevin dynamics, in collaboration with G. Di Gesù (TU Wien, Austria), B. Nectoux (Université Blaise Pascal) and D. Le Peutrec (Université Paris-Sud). The quasi-stationary distribution approach has been developed by T. Lelièvre and collaborators over the past years in order to rigorously model the exit event from a metastable state by a jump Markov process, and to study this exit event in the small temperature regime. In [48], the authors prove that in the small temperature regime and under rather general assumptions on the initial conditions and on the potential function, the support of the distribution of the first exit point concentrates on some points realizing the minimum of the potential on the boundary. The proof relies on tools to study tunnelling effects in semi-classical analysis. This preprint has been divided into two separate articles for publication: the first one [22] has been accepted for publication; the second one is currently under review.

6.2.3. Nonequilibrium systems and computation of transport coefficients

Stemming from the IHP trimester "Stochastic Dynamics Out of Equilibrium" held at Institut Henri Poincare in April-July 2017, a collection of contributions has been grouped in a volume of proceedings [38], focusing on aspects of nonequilibrium dynamics and its ongoing developments. This volume has been edited by G. Giacomin (Universite Paris Diderot), S. Olla (Université Paris Dauphine), E. Saada (CNRS and Université Paris Descartes), H. Sophn (TU Munich, Germany) and G. Stoltz. It includes contributions from various events relating to three domains: (i) transport in non-equilibrium statistical mechanics; (ii) the design of more efficient simulation methods; (iii) life sciences.

In addition, P. Plechac (University of Delaware, USA), T. Wang (Army Research Lab, USA) and G. Stoltz have considered in [61] numerical schemes for computing the linear response of steady-state averages of stochastic dynamics with respect to a perturbation of the drift part of the stochastic differential equation. The schemes are based on Girsanov's change-of-measure theory to reweight trajectories with factors derived from a linearization of the Girsanov weights. Both the discretization error and the finite time approximation error have been investigated. The designed numerical schemes have been shown to be of bounded variance with respect to the integration time, which is a desirable feature for long time simulation. The discretization error has been shown to be improved to second order accuracy in the time step by modifying the weight process in an appropriate way.

6.2.4. Coarse-graining

Two works have been done to explore new methods to define "good" reaction coordinates:

- The estimation of the Poincaré constant of a given probability measure allows to quantify the typical convergence rate of reversible diffusions to their equilibrium measure. Loucas Pillaud-Vivien, F. Bach and A. Rudi (Inria SIERRA), together with T. Lelièvre and G. Stoltz, have shown in [58], both theoretically and experimentally how to estimate the Poincaré constant given sufficiently many samples of the probability measure under consideration, using reproducing Hilbert kernel spaces. As a by-product of this estimation, they have also derived an algorithm that captures a low dimensional representation of the data by finding directions which are difficult to sample – reaction coordinates in the language of molecular dynamics. This amounts to finding the marginal of the high dimensional sampled measure for which the Poincaré constant is the largest possible.
- In [60], T. Lelièvre together with B. Leimkuhler and Z. Trstanova (University of Edinburgh, Scotland) has explored numerically the interest of using diffusion maps to define reaction coordinates or

metastable states. Diffusion maps approximate the generator of Langevin dynamics from simulation data, and the idea is thus to use the eigenvalues and eigenvectors to build coarse-grained variables. They have also discussed the use of diffusion maps to define local reaction coordinates within the metastable sets, formalising the locality via the concept of quasi-stationary distribution and justifying the convergence of diffusion maps applied to samples within a metastable set.

Another coarse-graining procedure was considered to justify the approximation of an infinite system of ordinary differential equations (the Becker-Doring equations, describing coagulation/fragmentation processes of species of integer sizes) in terms of a partial differential equation. Formal Taylor expansions motivate that the dynamics at large sizes should be dictated by an advection-diffusion equation, called Fokker-Planck equation. P. Terrier and G. Stoltz rigorously proved in [59] the link between these two descriptions for evolutions on finite times rather than in some hydrodynamic limit, motivated by the results of numerical simulations and the construction of dedicated algorithms based on splitting strategies. In fact, the Becker-Doring equations and the Fokker-Planck equation are related through some pure diffusion with unbounded diffusion coefficient. The crucial point in the analysis is to obtain decay estimates for the solution of this pure diffusion and its derivates to control remainders in the Taylor expansions. The small parameter in this analysis is the inverse of the minimal size of the species.

6.3. Homogenization

Participants: Xavier Blanc, Virginie Ehrlacher, Olga Gorynina, Rémi Goudey, Claude Le Bris, Frédéric Legoll, Adrien Lesage, Pierre-Loïk Rothé.

In homogenization theory, members of the project-team have pursued their ongoing systematic study of perturbations of periodic problems (by local and nonlocal defects). This has been done in several different directions.

6.3.1. Deterministic non-periodic systems

For linear elliptic equations with highly oscillating coefficients, X. Blanc and C. Le Bris have recently developed, in collaboration with P.-L. Lions (Collège de France), a theory in the case of periodic problems with local defects. In particular, the existence of a corrector function for such problems has been shown. More details on the quality of approximation achieved by their theory have been recently provided. The fact that a corrector exists with suitable properties indeed allows one to quantify the rate of convergence of the two-scale expansion (which uses that corrector) to the actual exact solution, as the small homogenization parameter ε vanishes. In that spirit, some of these works by X. Blanc and C. Le Bris, in collaboration with M. Josien (former PhD student in the team, now at MPI Leipzig, Germany), have been presented in [12], [13].

Also in the context of homogenization theory, O. Gorynina, C. Le Bris and F. Legoll have explored the question of how to determine the homogenized coefficient of heterogeneous media without explicitly performing an homogenization approach. This work is a follow-up on earlier works by C. Le Bris and F. Legoll in collaboration with K. Li and next S. Lemaire over the years. During the year, O. Gorynina, C. Le Bris and F. Legoll have mathematically studied a computational approach initially introduced by R. Cottereau (CNRS Marseille). This approach combines, in the Arlequin framework, the original fine-scale description of the medium (modelled by an oscillatory coefficient) with an effective description (modelled by a constant coefficient) and optimizes upon the coefficient of the effective medium to best fit the response of a purely homogeneous medium. In the limit of asymptotically infinitely fine structures, the approach yields the value of the homogenized coefficient. The aim is to mathematically study the problem and to investigate how to improve on the practical algorithm, in order to obtain a procedure as efficient as possible. Results will be presented in a couple of manuscripts in preparation.

6.3.2. Stochastic homogenization

The project-team has pursued its efforts in the field of stochastic homogenization of elliptic equations, aiming at designing numerical approaches that are practically relevant and keep the computational workload limited.

Using standard homogenization theory, one knows that the homogenized tensor, which is a deterministic matrix, depends on the solution of a stochastic equation, the so-called corrector problem, which is posed on the whole space $\mathbb{R}^d$. This equation is therefore delicate and expensive to solve, and the team has proposed, over the past years, many approaches to improve on the computation of the homogenized tensor.

Besides the averaged behavior of the oscillatory solution u_ε on large space scales (which is given by its homogenized limit), a question of interest is to describe how u_ε fluctuates. This question has been investigated in the PhD thesis of P.-L. Rothé, both from a theoretical and a numerical viewpoints. First, theoretical results have been obtained for a weakly stochastic setting (where the coefficient is the sum of a periodic coefficient and a small random perturbation). It has been shown that, at the first order and when ε is small, the localized fluctuations (characterized by a test function g) of u_ε are Gaussian. The corresponding variance depends on the localization function g, on the right-hand side f of the problem satisfied by u_ε, and on a fourth order tensor Q which is defined in terms of the corrector. Since the corrector function is challenging to compute, so is Q. A numerical approach (based on using the standard truncated corrector problem) has been designed to approximate Q and its convergence has been proven, again in a weakly stochastic setting. All these theoretical results critically depend on detailed properties of the Green function associated to the periodic operator. Second, numerical experiments in more general settings (i.e. full stochastic case) following the same approach have been performed, in order to investigate the generality of the obtained results. First, the convergence of the approximation of Q has been monitored. Second, it has been checked that the localized fluctuations of u_ε indeed become Gaussian when ε decreases, and that their variance can be related to Q. These promising numerical results, which are consistent with the theoretical results obtained in the weakly stochastic setting, are presented in a manuscript in preparation.

6.3.3. Multiscale Finite Element approaches

From a numerical perspective, the Multiscale Finite Element Method (MsFEM) is a classical strategy to address the situation when the homogenized problem is not known (e.g. in difficult nonlinear cases), or when the scale of the heterogeneities, although small, is not considered to be zero (and hence the homogenized problem cannot be considered as a sufficiently accurate approximation).

During the year, several research tracks have been pursued in this general direction.

The MsFEM approach uses a Galerkin approximation on a pre-computed basis, obtained by solving local problems mimicking the problem at hand at the scale of mesh elements, with carefully chosen right-hand sides and boundary conditions. The initially proposed version of MsFEM uses as basis functions the solutions to these local problems, posed on each mesh element, with null right-hand sides and with the coarse P1 elements as Dirichlet boundary conditions. Various improvements have next been proposed, such as the *oversampling* variant, which solves local problems on larger domains and restricts their solutions to the considered element. In collaboration with U. Hetmaniuk (University of Washington in Seattle, USA), C. Le Bris, F. Legoll and P.-L. Rothé have completed the study of a MsFEM method improved differently. They have considered a variant of the classical MsFEM approach with enrichments based on Legendre polynomials, both in the bulk of the mesh elements and on their interfaces. A convergence analysis of this new variant has been performed. In addition, residue type a posteriori error estimators have been proposed and certified, leading to a numerical strategy where the degree of enrichment is *locally* adapted in order to reach, at the smallest computational cost, a given error. The promising numerical results are currently being collected in a manuscript in preparation.

Many numerical analysis studies of the MsFEM are focused on obtaining a priori error bounds. In collaboration with L. Chamoin, who was on leave in the project-team a few years ago from ENS Cachan, members of the project-team have been working on a posteriori error analysis for MsFEM approaches, with the aim of developing error estimation and adaptation tools. They have extended to the MsFEM case an approach that is classical in the computational mechanics community for single scale problems, and which is based on the so-called Constitutive Relation Error (CRE). Once a numerical solution u_h has been obtained, the approach needs additional computations in order to determine a divergence-free field as close as possible to the exact flux $k\nabla u$. In the context of the MsFEM, it is important to be able to perform all expensive computations in an offline stage, independently of the right-hand side. The standard CRE approach has thus been adapted to that context. In the recent work [47], the approach has also been adapted towards the design of adaptive algorithms for specific quantities of interest (in the so-called "goal-oriented" setting). It provides an accurate estimation of the error, and leads to a discretization which is efficiently tailored to the specific quantity under consideration.

One of the perspectives of the team, through the PhD thesis of A. Lesage, is the development of Multiscale Finite Element Methods for thin heterogeneous plates. The fact that one of the dimension of the domain of interest scales as the typical size of the heterogeneities within the material induces theoretical and practical difficulties that have to be carefully taken into account (see [37]). The first steps of the work of V. Ehrlacher, F. Legoll and A. Lesage, in collaboration with A. Lebée (Ecole des Ponts) have consisted in studying the homogenized limit (and the two-scale expansion) of problems posed on thin heterogeneous plates. After having considered the case of a diffusion equation, the more challenging case of elasticity has been studied. In the so-called membrane case (that is, when the loading is in the in-plane directions), an approximation result for the two-scale expansion has been obtained. Several MsFEM variants have been proposed and compared numerically. The results will be presented in a forthcoming manuscript.

6.4. Various topics

Participants: Sébastien Boyaval, Virginie Ehrlacher.

A new mathematical framework has been identified for the modelling of complex fluids in [42]. It allows one to incorporate rheological features of a real (non-ideal, visous and compressible) fluid and, at the same time, to compute flows as solution to a *hyperbolic system of conservation laws* complemented by an initial value. In [42], the framework is specified for 2D hydrostatic flows of *Maxwell fluids*, with a numerical finite-volume scheme preserving the positivity of mass and a Clausius-Duem inequality. Formally, the (macroscopic) model has a (microscopic) molecular justification using a generalized Langevin equation for the distortion of the fluid texture.

On the other hand, recall that stochastic models are also used for the numerical simulation of hydrodynamical turbulence. In particular, a generalized Langevin equation can be used to model the "thermostated" velocity fluctuations in a "stationary" turbulent flow modelled as an invariant measure. But the interest for the effective numerical simulation of turbulent flows is not fully understood yet. In [43], S. Boyaval with S. Martel and J. Reygner (Ecole des Ponts) have studied the convergence of a discretization of a 1D stochastic scalar viscous conservation laws (a toy-model), for the numerical simulation of its invariant measure.

A. Benaceur (EDF), V. Ehrlacher and A. Ern (École des Ponts and Inria SERENA) developped a new EIM/reduced-basis method [10] for the reduction of parametrized variational inequalities with nonlinear constraints, and applied this method to the reduction of contact mechanics problems with non-coincident meshes.

V. Ehrlacher, D. Lombardi (Inria COMMEDIA), O. Mula (Université Paris-Dauphine) and F-X. Vialard (Université Paris-Est) developped new model-order reduction techniques based on the use of Wasserstein spaces for transport-dominated problems [51], which gives very encouraging results on several classes of conservative transport problems like the Burger's equation. Theoretical convergence rates are proved on some particular test cases.

J. Berendsen (Chemnitz, Germany), Martin Burger (Erlangen, Germany), V. Ehrlacher, J-F. Pietschmann (Chemnitz, Germany) proved the existence and uniqueness of strong solutions and weak-strong stability in a particular system of cross-diffusion equations [41]. It is in general very difficult to obtain such kind of results for general cross-diffusion systems. The proof for the particular system studied here relies on the fact that, when all the cross-diffusion coefficients of the system are equal to the same constant, the system boils down to a set of independent heat equations, for which uniqueness of strong solutions is trivial. The uniqueness of strong solutions was proved under the assumption that the cross-diffusion coefficients should not be close enough to one another.

7. Bilateral Contracts and Grants with Industry

7.1. Contracts and grants with Industry

Many research activities of the project-team are conducted in close collaboration with private or public companies: CEA, SANOFI, EDF. The project-team is also supported by the Office of Naval Research and the European Office of Aerospace Research and Development, for multiscale simulations of random materials. All these contracts are operated at and administrated by the École des Ponts.

8. Partnerships and Cooperations

8.1. National Initiatives

The project-team is involved in several ANR projects:

- S. Boyaval is the PI of the ANR JCJC project SEDIFLO (2016-2021) to investigate new numerical models of solid transport in rivers.
- V. Ehrlacher is the PI of the ANR project COMODO (2020-2024) which focuses on the development of efficient numerical methods to simulate cross-diffusion systems on moving domains, with application to the simulation of the fabrication process of thin film solar cells. It includes research teams from Inria Lille, Inria Sophia-Antipolis and Germany.
- V. Ehrlacher is a member of the ANR project ADAPT (2018-2022), PI: D. Lombardi, Inria REO team-project. This project is concerned with the parallelization of tensor methods for high-dimensional problems.
- F. Legoll is a member of the ANR project CINE-PARA (2015-2020), PI: Y. Maday, Sorbonne Université. This project is concerned with parallel-in-time algorithms.
- T. Lelièvre is responsible of the node "Ecole des Ponts" of the ANR QuAMProcs (2019-2023), to which G. Stoltz also participates, PI: L. Michel, Université de Bordeaux.
- G. Stoltz is the PI of the ANR project COSMOS (2014-2019) which focuses on the development of efficient numerical techniques to simulate high-dimensional systems in molecular dynamics and computational statistics. It includes research teams from Institut Mines-Telecom, Inria Rennes and IBPC Paris.

Members of the project-team are participating in the following GdR:

- AMORE (Advanced Model Order REduction),
- CORREL (correlated methods in electronic structure computations),
- DYNQUA (time evolution of quantum systems, with applications to transport problems, nonequilibrium systems, etc.),
- EGRIN (gravity flows),
- MANU (MAthematics for NUclear applications),
- MASCOT-NUM (stochastic methods for the analysis of numerical codes),
- MEPHY (multiphase flows),
- NBODY (electronic structure),
- REST (theoretical spectroscopy),
- CHOCOLAS (experimental and numerical study of shock waves).

The project-team is involved in two Labex: the Labex Bezout (2011-) and the Labex MMCD (2012-).

We have invited the following national researchers to visit our team:

- A. Lozinski (University of Besançon): repeated visits during the year 2019.

8.2. European Initiatives

The ERC consolidator Grant MSMATH (ERC Grant Agreement number 614492, PI T. Lelièvre) ended in June 2019.

The ERC Synergy Grant EMC2 (ERC Grant Agreement number 810367 , PI E. Cancès, L. Grigori, Y. Maday, J-P. Piquemal) has started in September 2019.

8.3. International Initiatives

T. Lelièvre, G. Stoltz and F. Legoll participate in the Laboratoire International Associé (LIA) CNRS / University of Illinois at Urbana-Champaign on complex biological systems and their simulation by high performance computers. This LIA involves French research teams from Université de Nancy, Institut de Biologie Structurale (Grenoble) and Institut de Biologie Physico-Chimique (Paris). The LIA has been renewed for 4 years, starting January 1st, 2018.

9. Dissemination

9.1. Promoting Scientific Activities

E. Cancès

- has been director (until August 2019) and is now co-director (starting September 2019) of CERMICS, the Applied Mathematics department at École des Ponts,
- is a member of the editorial boards of Mathematical Modelling and Numerical Analysis (2006-), SIAM Journal of Scientific Computing (2008-), SIAM Multiscale Modeling and Simulation (2012-), and the Journal of Computational Mathematics (2017-),
- has co-organized the 4th international conference on Mathematical and Numerical Analysis of Electronic Structure Models (Suzhou, China, June), an interdisciplinary summer school on Mathematical Methods for Molecular Simulation at Jussieu (Paris, June), a series of minisymposia at ICIAM 2019 (Valencia, Spain, July), and the FSMP Horizon Maths 2019 conference (Paris, December).

V. Ehrlacher

- is a member of the "Conseil d'Enseignement et de Recherche" of Ecole des Ponts,
- co-organizes the colloquium of the CERMICS lab,
- co-organized a minisymposium at the ICIAM conference in Valencia, July.

C. Le Bris is a managing editor of Networks and Heterogeneous Media. He is a member of the editorial boards of Annales mathématiques du Québec (2013-), Archive for Rational Mechanics and Analysis (2004-), Calcolo (2019-), COCV (Control, Optimization and Calculus of Variations) (2003-), Mathematics in Action (2008-), Nonlinearity (2005-), Journal de Mathématiques Pures et Appliquées (2009-), Pure and Applied Analysis (2018-). He is a member of the editorial boards of the monograph series Mathématiques & Applications, Series, Springer (2008-), Modelling, Simulations and Applications, Series, Springer (2009-), Springer Monographs in Mathematics, Springer (2016-). He is a member of

- the "Conseil de la Faculté des sciences et ingénierie", Sorbonne Université,
- the "Conseil scientifique" of SMAI.

He is the president of the scientific advisory board of the Institut des Sciences du calcul et des données, Sorbonne Université. He holds a regular position of Visiting Professor at the University of Chicago.

T. Lelièvre

- is editor-in-chief of ESAIM: Proceedings and Surveys (with D. Chafai, C. Imbert and P. Lafitte),
- is a member of the editorial boards of IMA: Journal of Numerical Analysis, SIAM/ASA Journal of Uncertainty Quantification, Communications in Mathematical Sciences and Journal of Computational Physics,
- is a member of the “Conseil d’Administration” of SMAI and École des Ponts,
- has co-organized the CECAM workshop "Learning the collective variables of biomolecular processes" at Inria Paris July 10-12th 2019 (with L. Delemotte, J. Hénin and G. Stock),
- has co-organized the ICL/CNRS workshop on "Interacting Particle Systems and applications" at Imperial College of London, December 9-10th 2019 (with G. Pavliotis),
- holds a visiting professorship position at Imperial College of London (from August 2019).

F. Legoll

- is a member of the editorial board of SIAM MMS (2012-) and of ESAIM: Proceedings and Surveys (2012-),
- is a member of the ANR committee CES-40 "Mathématiques et informatique",
- has co-organized, with U. Hetmaniuk (U. of Washington, USA), the mini-symposium "Multiscale and domain decomposition approaches for PDEs with rough coefficients" within the SIAM CSE 2019 conference (Spokane, USA, February),
- has co-organized, with C. Le Bris, the mini-symposium "Computational approaches for multiscale, possibly random problems" within the ICIAM 2019 conference (Valencia, Spain, July).

A. Levitt co-organises the applied mathematics seminar of the CERMICS lab.

G. Robin

- co-organizes the working group “Machine learning and optimization” of the Labex Bezout (with G. Stoltz, as well as W. Hachel and R. Elie),
- is the president of the group "Jeunes de la Société Française de Statistique" (SFdS).

G. Stoltz

- is a member of the scientific council of UNIT (Université Numérique Ingénierie et Technologie),
- is a member of the "Conseil d’Enseignement et de Recherche" of Ecole des Ponts.

9.2. Teaching - Supervision - Juries

The members of the project-team have taught the following courses.

At École des Ponts 1st year (equivalent to L3):

- Analyse et calcul scientifique, 30h (V. Ehrlacher, O. Gorynina, R. Goudey, A. Lesage, S. Siraj-Dine, G. Stoltz),
- Équations aux dérivées partielles et éléments finis, 15h (F. Legoll, P.-L. Rothé),
- Hydraulique numérique, 15h (S. Boyaval),
- Mécanique quantique, 10h (E. Cancès, A. Levitt),
- Méthodes numériques pour les problèmes en grande dimension, 17h30 (V. Ehrlacher, S. Boyaval),
- Optimisation, 15h, L3 (G. Kemlin, A. Lesage), Outils mathématiques pour l’ingénieur, 15h (E. Cancès, G. Ferré, F. Legoll, T. Lelièvre, P-L. Rothé),
- Projet de première année, 15h (G. Ferré).

At École des Ponts 2nd year (equivalent to M1):

- Analyse de Fourier, 15h (A. Levitt),
- Problèmes d'évolution, 36h (F. Legoll, V. Ehrlacher),
- Contrôle de systèmes dynamiques et équations aux dérivées partielles, 18h (E. Cancès),
- Projet du département IMI, 12h (G. Ferré, M. Ramil, J. Roussel, L. Silva Lopes),
- Projets Modéliser Programmer Simuler (V. Ehrlacher, G. Robin),
- Statistics and data sciences, 24h (G. Stoltz),
- Techniques de développement logiciel, 18h (M. Herbst).

At École des Ponts 3rd year (equivalent to M2):

- Méthodes de quantification des incertitudes en ingénierie, 18h (V. Ehrlacher),
- Remise à niveau: outils mathématiques, 9h (A. Lesage).

At the M2 "Mathématiques de la modélisation" of Sorbonne Université:

- Introduction à la physique statistique computationnelle, 20h (G. Stoltz),
- Méthodes numériques probabilistes, 24h (T. Lelièvre),
- Problèmes multiéchelles, aspects théoriques et numériques, 24h (F. Legoll),
- Théorie spectrale et variationnelle, 10h (E. Cancès).

At other institutions:

- Homogenization theory and multiscale problems, 21h, University of Chicago (C. Le Bris).
- Maths 2, 3h, L3, École des Mines (G. Stoltz),
- Probabilités de 1ère année, 27h, L3, Ecole des Mines (M. Ramil),
- Supervision of M2 projects, Master Mathématiques pour les sciences du vivant (MathSV) Université Paris-Saclay (G. Robin).

The following PhD theses supervised by members of the project-team have been defended:

- Ling-Ling Cao, Mathematical analysis of models of electronic structure for defected materials, Université Paris-Est, École des Ponts, defended on October 29th, 2019, supervised by E. Cancès and G. Stoltz,
- Grégoire Ferré, Large deviations theory in statistical physics: some theoretical and numerical aspects, Université Paris-Est, École des Ponts, defended on November 27th, 2019, supervised by G. Stoltz,
- Pierre-Loik Rothé, Numerical methods for the estimation of fluctuations in multi-scale materials and related problems, Université Paris-Est, École des Ponts, defended on December 12th, 2019, supervised by F. Legoll,
- Laura Silva Lopes, Rare event simulation and applications to biological systems, Université Paris-Est, Ecole des Ponts, defended on December 19th, 2019, supervised by J. Hénin (IBPC) and T. Lelièvre.

The following PhD theses supervised by members of the project-team are ongoing:

- Zineb Belkacemi, thèse CIFRE SANOFI, Machine learning for reaction coordinates in molecular dynamics, Université Paris-Est, since November 2018, supervised by T. Lelièvre and G. Stoltz,
- Robert Benda, Multiscale modeling of functionalized nanotube networks for sensor applications, École Polytechnique, started September 1st, 2018, supervised by E. Cancès and B. Lebental (École Polytechnique),
- Raed Blel, Monte Carlo methods and model reduction, started October 1st, 2018, supervised by V. Ehrlacher and T. Lelièvre,

- Rafaël Coyaud, Méthodes déterministes et stochastiques pour le transport optimal, Université Paris-Est, École des Ponts, started october 2017, supervised by A. Alfonsi (CERMICS), co-supervised by V. Ehrlacher,
- Qiming Du, Mathematical analysis of splitting methods, Ecole Doctorale Sciences Mathématiques de Paris Centre, started September 1st, 2016, supervised by A. Guyader (UPMC) and T. Lelièvre,
- Rémi Goudey, Problèmes d'homogénéisation en présence de défauts, Université Paris-Est, started in September 2019, supervised by C. Le Bris.
- Gaspard Kemlin, Mathematical and numerical analysis for electronic structures, École des Ponts, started September 1st, 2019, supervised by E. Cancès and co-supervised by A. Levitt.
- Adrien Lesage, Multi-scale methods for calculation and optimization of thin structures, started October 1st, 2017, supervised by F. Legoll, co-supervised by V. Ehrlacher and A. Lebée (Ecole des Ponts),
- Mouad Ramil, Metastability for interacting particle systems, started October 1st 2017, supervised by T. Lelièvre and J. Reygner (CERMICS),
- Lise Maurin, Non reversible and adaptive biasing processes for sampling, started 1st October 2018, supervised by T. Lelièvre and J.-P. Piquemal (Sorbonne Université), together with P. Monmarché (Sorbonne Université),
- Idrissa Niakh, Réduction de modèles pour les inégalités variationnelles, Université Paris-Est, École des Ponts, thèse CIFRE EDF, started november 2019, supervised by A. Ern (CERMICS), co-supervised by V. Ehrlacher,
- Inass Sekkat, Large scale Bayesian inference, Université Paris-Est, since March 2018, supervised by G. Stoltz,
- Sami Siraj-Dine, Modélisation mathématique des matériaux 2D, École des Ponts, started October 1st, 2017, supervised by E. Cancès, C. Fermanian and co-supervised by A. Levitt.

Project-team members have participated in the following PhD juries:

- S. Boyaval, PhD of Olivier Ozenda ("Continuous modelisation of suspension rheology and migration processes"), defended at Université Grenoble Alpes in Mars 2019,
- S. Boyaval, PhD of Sofiane Martel ("Numerical and theoretical analysis of invariant measures of scalar stochastic viscous conservation laws"), defended at Ecole des Ponts in December 2019,
- E. Cancès, PhD of Amaury Hayat ("Stabilisation de systèmes hyperboliques non-linéaires en dimension un d'espace"), defended at Sorbonne University in May 2019,
- V. Ehrlacher, PhD of Nadia Jbili ("Design and analysis of optimization schemes for nuclear magnetic resonance"), defended at Université Paris-Dauphine in December 2019,
- V. Ehrlacher, PhD of Charles Paillet ("Nouvelles démarches de réduction de modèles pour le traitement des problèmes à très grand nombre de paramètres"), defended at Ecole Normale Supérieure Paris-Saclay in June 2019,
- F. Legoll, referee for the PhD of Qingqing Feng ("Développement d'une méthode d'éléments finis multi-échelles pour les écoulements incompressibles dans un milieu hétérogène"), defended at École Polytechnique in September 2019,
- T. Lelièvre, referee for the PhD of Augustin Chevallier ("Random walks for estimating densities of states and the volume of convex bodies in high dimensional spaces"), defended at Université Côte d'Azur in April 2019,
- T. Lelièvre, referee for the PhD of Oleg Balabanov ("Randomized linear algebra for model order reduction"), defended at Université Bretagne Loire in October 2019,
- T. Lelièvre, referee for the PhD of Lara Neureither ("Irreversible multi-scale diffusions: time scales and model reduction"), defended at Brandenburgische Technische Universität in November 2019,

- G. Stoltz, referee for the PhD of Nicolas Brosse ("Around the Langevin algorithm in high dimension: extensions and applications"), defended at Ecole polytechnique in June 2019,
- G. Stoltz, PhD thesis of Laurent Laflèche ("Large particle dynamical systems"), defended at Université Paris-Dauphine Université Paris-Dauphine in June 2019,
- G. Stoltz, referee for the PhD of Nada Cvtekovic ("Convergent discretization schemes for transition path theory for diffusion processes"), defended at FU Berlin in Fall 2019,
- G. Stoltz, referee for the PhD of Nadia Jbili ("Design and analysis of optimization schemes for nuclear magnetic resonance"), defended at Université Paris-Dauphine in December 2019.

Project-team members have participated in the following habilitation juries:

- S. Boyaval, HdR of Sophie Ricci ("Uncertainties quantification and reduction in the computational geosciences – Application to free-surface hydraulics"), defended at CERFACS in April 2019,
- T. Lelièvre, president of the HDR jury of Denis Villemonais ("Exponential convergence to quasistationary distributions and applications"), defended at Université de Lorraine in November 2019.

9.3. Conference participation

Members of the project-team have delivered lectures in the following seminars, workshops and conferences:

- R. Benda, Informal Scientific Discussion (ISD) seminar of LSI laboratory, Ecole Polytechnique, Palaiseau, March,
- S. Boyaval, ICIAM minisymposia, Valencia (Spain), July,
- S. Boyaval, Workshop in computational hydraulics at UM6P, Ben Guerir (Morocco), September,
- S. Boyaval, 19th day of scientific computing and mathematical modelling in Amiens (France), June,
- E. Cancès University of Strasbourg, colloquium of the mathematics department, January,
- E. Cancès, WONAPDE, Conception, Chile, January (plenary lecture),
- E. Cancès, BIRS workhop, Banff, Canada, January,
- E. Cancès, Sorbonne Université, EMC2 seminar, Paris, February,
- E. Cancès, University of Chicago, Computational and applied mathematics seminar, USA, March,
- E. Cancès, ICMS workshop, Edinburgh, United Kingdom, March,
- E. Cancès, Fields Institute workshop, Toronto, Canada, April,
- E. Cancès, CECAM workshop, Lausanne, Switzerland, May,
- E. Cancès, AMMCS, Waterloo, Canada, August (plenary lecture),
- E. Cancès, IMS workshop, Singapore, September,
- E. Cancès, Sorbonne University, ERC EMC2 kick-off meeting, October,
- E. Cancès, Campus des Cordeliers, laboratoire J.-L. Lions 50th anniversary, Paris, November,
- V. Ehrlacher, GAMM meeting (keynote lecture), Vienna, February,
- V. Ehrlacher, Inria-LJLL seminar, Paris, April,
- V. Ehrlacher, Laboratoire Paul Painlevé seminar, Lille, April,
- V. Ehrlacher, Workshop Cambridge-MMCD, Marne-la-Vallée, April,
- V. Ehrlacher, workshop on "Optimal Transport: from Geometry to Numerics", Erwin Schrödinger Institut, Vienna, April,
- V. Ehrlacher, workshop on "Scientific Computing Across Scales: Quantum Systems in Cold-matter Physics and Chemistry", Fields Institut, Toronto, April,
- V. Ehrlacher, ICIAM, Valencia, July,
- V. Ehrlacher, MFO workshop: "Computational Multiscale Methods", Oberwolfach, August,
- V. Ehrlacher, MORTECH, Paris, November,

- G. Ferré, Young researchers' seminar, Université Paris-Dauphine, France, February,
- G. Ferré, Probability seminar at Lille University, France, March,
- G. Ferré, Probability seminar at Marseille University, France, May,
- G. Ferré, Congrès SMAI 2019, Lorient, France, May,
- G. Ferré, Journées de Probabilités, Dourdan, France, June,
- G. Ferré, ICIAM 2019, Valencia, Spain, July,
- G. Ferré, SciCADE 2019, Innsbruck, Austria, July
- G. Ferré, GAMM MoAnSi 2019 meeting, Munich, Germany, September,
- G. Ferré, Probability seminar, University Saint-Quentin en Yvelines, France, November,
- O. Gorynina, Congrès SMAI, Guidel, France, May,
- O. Gorynina, ICIAM 2019 conference, Valencia, Spain, July,
- O. Gorynina, Applied Mathematics Seminar, Surgut State University, Russia, December,
- M. Herbst, Université de Lorraine Laboratoire de Physique et Chimie Théoriques seminar, Metz, May,
- M. Herbst, Université de Lille, Laboratoire de Physique des Lasers, Atomes et Molécules seminar, Lille, May,
- M. Herbst, Technische Universität München, seminar Domcke group, München, September,
- M. Herbst, GAMM moansi annual meeting, München, September,
- M. Herbst, Université Paul Sabatier Toulouse, Laboratoire de Chimie et Physique Quantiques seminar, Toulouse, November,
- G. Kemlin, GAMM - MOANSI meeting, München, Germany, September
- C. Le Bris, ICIAM 2019 Invited speaker, July 2019, Valencia, Spain,
- F. Legoll, University of Chicago, CAMP seminar, Chicago, USA, February,
- F. Legoll, SIAM CSE 2019 conference, Spokane, USA, February,
- F. Legoll, Séminaire "Probabilités, Statistiques, Contrôle", ENSTA, Palaiseau, March,
- F. Legoll, CECAM workshop on "Big data and Uncertainty Quantification", Lausanne, Switzerland, March,
- F. Legoll, Colloque national en calcul des structures 2019, Giens, May,
- F. Legoll, ADMOS 2019 conference, Alicante, Spain, May,
- F. Legoll, Coupled problems 2019 conference, Barcelona, Spain, June,
- F. Legoll, MAFELAP (Mathematics of Finite Elements and Applications) 2019 conference, London, United Kingdom, June,
- F. Legoll, Séminaire du Laboratoire de Mécanique et d'Acoustique, Marseille, June,
- F. Legoll, ICIAM 2019 conference, Valencia, Spain, July,
- F. Legoll, Complas 2019 conference, Barcelona, Spain, September,
- F. Legoll, Workshop on "New trends in asymptotic methods for multiscale PDEs", Karlstad, Sweden, October,
- F. Legoll, MORTech 2019 workshop, Paris, November,
- T. Lelièvre, LIA CNRS - UIUC Meeting, Hauteluce, January,
- T. Lelièvre, Séminaire LPCT, Nancy, February,
- T. Lelièvre, Computational and Applied Mathematics / PDE seminar, University of Chicago, February,
- T. Lelièvre, CECAM Workshop, CIB-EPFL, Lausanne, Switzerland, March,

- T. Lelièvre, CECAM Workshop, CIB-EPFL, Lausanne, Switzerland, May,
- T. Lelièvre, CNLS Seminar, Los Alamos National Laboratory, USA, June,
- T. Lelièvre, IPAM, Lake Arrowhead, USA, June,
- T. Lelièvre, CECAM Workshop, Paris, July,
- T. Lelièvre, ICIAM, Valencia, Spain, July,
- T. Lelièvre, HetSys launch event, Warwick, England, September,
- T. Lelièvre, Numerical analysis in Bielefeld, Germany, September,
- T. Lelièvre, ANR QuAMProcs, Bordeaux, November,
- T. Lelièvre, AMMP Colloquium, Imperial College London, November,
- T. Lelièvre, "Fluids and Materials" seminar, Bristol, England, November,
- T. Lelièvre, Newton Institute, Cambridge, England, November,
- T. Lelièvre, Materials Research Society Fall Meeting, Boston, USA, December,
- A. Lesage, Congrès SMAI, Guidel, France, May,
- A. Lesage, ICIAM 2019 conference, Valencia, Spain, July,
- A. Levitt, BIRS workhop, Banff, Canada, January,
- A. Levitt, mathematical physics seminar, Copenhagen, February
- A. Levitt, Workshop Cambridge-MMCD, Marne-la-Vallée, April,
- A. Levitt, Congrès SMAI, Guidel, France, May,
- A. Levitt, Precision quantification in density functional theory, Louvain, Belgique, May,
- A. Levitt, Numerical Analysis of Electronic Structure Models, Suzhou, China, June,
- A. Levitt, ICIAM 2019 conference, Valencia, Spain, July,
- A. Levitt, GAMM - MOANSI meeting, München, Germany, September ,
- A. Levitt, weekly seminar, Aachen, Germany, October,
- A. Levitt, lunchtime seminar, Warwick UK, November,
- A. Levitt, Julia Meetup, Paris, December,
- G. Robin, seminar IACM, FORTH, Heraklion, October,
- G. Robin, seminar "Systèmes Complexes", CAMS, EHESS, November,
- P.-L. Rothé, Arbeitsgemeinschaft Applied Analysis, Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany, April,
- P.-L. Rothé, Groupe de travail des thésards du Laboratoire Jacques-Louis Lions, Paris, May,
- P.-L. Rothé, Congrès SMAI, Guidel, France, May,
- P.-L. Rothé, ADMOS 2019 conference, Alicante, Spain, May,
- P.-L. Rothé, ICIAM 2019 conference, Valencia, Spain, July,
- P.-L. Rothé, Journée "Approches probabilistes en mécanique" de la Fédération Francilienne de Mécanique (F2M), Paris, November,
- L. Silva Lopes, LIA CNRS - UIUC annual Meeting, Hauteluce, France, January,
- L. Silva Lopes, Lorentz Workshop, Leiden, Netherlands, March,
- L. Silva Lopes, IPAM, Lakearrowhead, USA, June,
- L. Silva Lopes, ICIAM, Valencia, Spain, July,
- L. Silva Lopes, MOANSI annual meeting, Munich, Germany, October,
- G. Stoltz, Mathematical Physics Seminar, Institut Henri Poincaré, March,

- G. Stoltz, CECAM workshop “Microscopic simulations: forecasting the next two decades”, Toulouse, France, April,
- G. Stoltz, Workshop Cambridge/Labex MMCD, Champs-sur-Marne, France, April,
- G. Stoltz, CIB workshop “Computational mathematics for model reduction and predictive modelling in molecular and complex systems”, Lausanne, Switzerland, May,
- G. Stoltz, MAP5 Colloquium, Université Paris Descartes, June,
- G. Stoltz, Rencontres prospectives RFCT, Nantes, France, June,
- G. Stoltz, Maths/chemistry seminar EMC2, Sorbonne-Université, June,
- G. Stoltz, DEFI/MEDISIM/POEMS seminar, Inria Saclay, July,
- G. Stoltz, ICIAM 2019, Valencia, Spain, July,
- G. Stoltz, SciCADE 2019, Innsbruck, Austria, July,
- G. Stoltz, seminar CEREMADE, Université Paris-Dauphine, October,
- G. Stoltz, Applied PDEs seminar, Imperial College London, October,
- G. Stoltz, QuAMProcs meeting, Bordeaux, France, November.

Members of the project-team have delivered the following series of lectures:

- M. Herbst, Introduction to the Julia programming language, 6h, Julia Day, Paris, December,
- A. Levitt, Complex analysis and applications in quantum physics and chemistry, 9h, GDR CORREL spring school, Paris, April,
- G. Stoltz, Sampling high-dimensional probability distributions and Bayesian learning, 6h, doctoral school UM6P, Morocco, November,

Members of the project-team have presented posters in the following seminars, workshops and international conferences:

- R. Benda, Ab Init School, CEA Bruyères-Le-Châtel (DAM), January.
- R. Benda, Journées Théorie, Modélisation et Simulations (JTMS) , Institut de Biologie Physico-chimique (IBPC), Paris, June.
- R. Benda, PhD day LPICM, Ecole Polytechnique, Palaiseau, April.
- R. Benda, LPICM Congress, Cap Ferret, October.
- R. Benda, GDR Graphene, Graphene & Co Meeting 2019, Bad Herrenalb, Germany, October.
- M. Herbst, 9th Molecular Quantum Mechanics Conference, Heidelberg, Germany, July,
- G. Ferré, CIB-CECAM meeting, Lausanne, Switzerland, May
- A. Lesage, Workshop “New trends and challenges in the mathematics of optimal design”, Cambridge, United Kingdom, June,
- G. Robin, doctoral school UM6P, Morocco, November,
- L. Silva Lopes, GRC Liquids in Complex Environments, Driving Reactions, Assembling and Pushed to Their Limits, Holderness, USA, August,

Members of the team have benefited from long-term stays in institutions abroad:

- T. Lelièvre, University of Chicago, USA, February 2019 (three weeks), Members of the project-team have participated (without giving talks nor presenting posters) in the following seminars, workshops and international conferences:

- R. Benda, Mini-school on mathematics for theoretical chemistry and physics, Paris, June,
- O. Gorynina, Oberwolfach summer school "Beyond Numerical Homogenization", Oberwolfach, Germany, June,
- M. Herbst, Mini-school on mathematics for theoretical chemistry and physics, Paris, June
- M. Herbst, Kick-off meeting extreme-scale mathematically-based computational chemistry, Paris, October,
- G. Kemlin, Mini-school on mathematics for theoretical chemistry and physics, Paris, June,
- G. Kemlin, Une giornata con Alessio - Quelques mathématiques autour d'Alessio Figali, Saclay, July,
- G. Kemlin, Kick-off meeting of the ERC project EMC2, Paris, October,
- G. Kemlin, Laboratoire J.-L. Lions 50th anniversary, Paris, November,
- G. Kemlin, Horizon Maths 2019 - Mathématiques et Chimie, Paris, December
- M. Ramil, ANR QuAMProcs, Bordeaux, November 2019,
- P.-L. Rothé, Workshop "Industrial problems solving", Montréal, Canada, August,
- L. Silva Lopes, CECAM Workshop, Paris, France July 2019,
- S. Siraj-Dine, Quantum Transport and Universality, Accademia Nazionale dei Lincei, September,

9.4. Popularization

9.4.1. Internal or external Inria responsibilities

- A. Levitt is a member of the editorial board of Interstices, Inria's popularization website,
- C. Le Bris has presented the project-team at "Mon équipe en 180 secondes", on November 7, 2019.

9.4.2. Articles and contents

- V. Ehrlacher has been interviewed for the online magazine “DigiSchool”
- G. Robin has been interviewed for the online magazine “Usbek et Rica” following her “L'Oréal-UNESCO For Women in Science” prize

9.4.3. Internal actions

- C. Le Bris has organized a research day for the students at École des Ponts on November 6, 2019.
- V. Ehrlacher has participated to a meeting “Mathématiques, nom féminin?” organized at Ecole des Ponts with pupils from middle school.

10. Bibliography

Major publications by the team in recent years

[1] E. CANCÈS, M. DEFRANCESCHI, W. KUTZELNIGG, C. LE BRIS, Y. MADAY.*Computational Quantum Chemistry: A Primer*, 2003, Le Bris, Claude (ed.), Special Volume: Computational Chemistry. Amsterdam: North-Holland. Handb. Numer. Anal. 10, 3-270 (2003)

[2] E. CANCÈS, C. LE BRIS, Y. MADAY.*Mathematical Methods in Quantum Chemistry. An Introduction. (Méthodes mathématiques en chimie quantique. Une introduction.)*, Mathématiques et Applications (Berlin) 53. Berlin: Springer. xvi, 409 p. , 2006

[3] I. CATTO, C. LE BRIS, P.-L. LIONS.*The Mathematical Theory of Thermodynamic Limits: Thomas-Fermi Type Models*, Oxford Mathematical Monographs. Oxford: Clarendon Press. xiii, 277 p., 1998

[4] J.-F. GERBEAU, C. LE BRIS, T. LELIÈVRE.*Mathematical Methods for the Magnetohydrodynamics of Liquid Metals*, Numerical Mathematics and Scientific Computation. Oxford: Oxford University Press., 324 p., 2006

[5] C. LE BRIS.*Multi-scale Analysis. Modeling and Simulation. (Systèmes multi-échelles. Modélisation et simulation.)*, Mathématiques et Applications (Berlin) 47. Berlin: Springer. xi, 212 p., 2005

[6] C. LE BRIS, P.-L. LIONS.*Parabolic Equations with Irregular Data and Related Issues: Applications to Stochastic Differential Equations*, De Gruyter Series in Applied and Numerical Mathematics, 2019, vol. 4

[7] T. LELIÈVRE, M. ROUSSET, G. STOLTZ.*Free Energy Computations: A Mathematical Perspective*, Imperial College Press, 458 p., 2010

Publications of the year

Doctoral Dissertations and Habilitation Theses

[8] G. FERRÉ.*Large Deviations Theory in Statistical Physics: Some Theoretical and Numerical Aspects*, Université Marne La Vallée, November 2019, https://hal.archives-ouvertes.fr/tel-02441535

[9] P.-L. ROTHÉ.*Numerical methods for the study of fluctuations in multi-scale materials and related problems*, Université Paris-Est Marne la Vallée, December 2019, https://tel.archives-ouvertes.fr/tel-02447725

Articles in International Peer-Reviewed Journal

[10] A. BENACEUR, A. ERN, V. EHRLACHER.*A reduced basis method for parametrized variational inequalities applied to contact mechanics*, in "International Journal for Numerical Methods in Engineering", October 2019 [*DOI :* 10.1002/NME.6261], https://hal.archives-ouvertes.fr/hal-02081485

[11] R. BENDA, E. CANCÈS, B. LEBENTAL.*Effective resistance of random percolating networks of stick nanowires: Functional dependence on elementary physical parameters*, in "Journal of Applied Physics", July 2019, vol. 126, n^{o} 4, 044306 [*DOI :* 10.1063/1.5108575], https://hal.archives-ouvertes.fr/hal-02281362

[12] X. BLANC, M. JOSIEN, C. LE BRIS.*Local precised approximation in multiscale problems with local defects*, in "Comptes Rendus Mathématique", 2019, https://arxiv.org/abs/1901.09669 [*DOI :* 10.1016/J.CRMA.2018.12.005], https://hal.archives-ouvertes.fr/hal-01893991

[13] X. BLANC, M. JOSIEN, C. LE BRIS.*Precised approximations in elliptic homogenization beyond the periodic setting*, in "Asymptotic Analysis", January 2020, vol. 116, n^{o} 2, p. 93-137, https://arxiv.org/abs/1812.07220 [*DOI :* 10.3233/ASY-191537], https://hal.archives-ouvertes.fr/hal-01958207

[14] T. BOIVEAU, V. EHRLACHER, A. ERN, A. NOUY.*Low-rank approximation of linear parabolic equations by space-time tensor Galerkin methods*, in "ESAIM: Mathematical Modelling and Numerical Analysis", May 2019, vol. 53, n^{o} 2, p. 635-658, https://arxiv.org/abs/1712.07256 [*DOI :* 10.1051/M2AN/2018073], https://hal.archives-ouvertes.fr/hal-01668316

[15] C.-E. BRÉHIER, T. LELIÈVRE.*On a new class of score functions to estimate tail probabilities of some stochastic processes with Adaptive Multilevel Splitting*, in "Chaos", 2019, vol. 29, 033126 [*DOI :* 10.1063/1.5081440], https://hal.archives-ouvertes.fr/hal-01923385

[16] E. CANCÈS, L.-L. CAO, G. STOLTZ.*A reduced Hartree–Fock model of slice-like defects in the Fermi sea*, in "Nonlinearity", January 2020, vol. 33, n^{o} 1, p. 156-195, https://arxiv.org/abs/1807.06960 [*DOI :* 10.1088/1361-6544/AB4C7D/META], https://hal.archives-ouvertes.fr/hal-01891488

[17] E. CANCÈS, V. EHRLACHER, F. LEGOLL, B. STAMM, S. XIANG.*An embedded corrector problem for homogenization. Part II: Algorithms and discretization*, in "Journal of Computational Physics", 2020, vol. 407, 109254, https://arxiv.org/abs/1810.09885 [*DOI :* 10.1016/J.JCP.2020.109254], https://hal.archives-ouvertes.fr/hal-01903486

[18] P. CARDALIAGUET, C. LE BRIS, P. E. SOUGANIDIS.*Perturbation problems in homogenization of hamilton-jacobi equations*, in "Journal de Mathématiques Pures et Appliquées", 2019, vol. 117, p. 221-262, https://arxiv.org/abs/1701.05440 , forthcoming [*DOI :* 10.1016/J.MATPUR.2018.03.005], https://hal.archives-ouvertes.fr/hal-01435744

[19] D. CHAFAÏ, G. FERRÉ.*Simulating Coulomb gases and log-gases with hybrid Monte Carlo algorithms*, in "Journal of Statistical Physics", February 2019, vol. 174, n^{o} 3, p. 692–714, https://arxiv.org/abs/1806.05985 [*DOI :* 10.1007/S10955-018-2195-6], https://hal.archives-ouvertes.fr/hal-01818268

[20] H. D. CORNEAN, D. GONTIER, A. LEVITT, D. MONACO.*Localised Wannier functions in metallic systems*, in "Annales Henri Poincaré", April 2019, vol. 20, n^{o} 4, p. 1367–1391, https://arxiv.org/abs/1712.07954 - 18 pages, 4 figures [*DOI :* 10.1007/S00023-019-00767-6], https://hal.archives-ouvertes.fr/hal-01671848

[21] G. DI GESÙ, T. LELIÈVRE, D. LE PEUTREC, B. NECTOUX.*Sharp asymptotics of the first exit point density*, in "Annals of PDE", June 2019, vol. 5, n^{o} 1, 5, https://arxiv.org/abs/1706.08728 - 146 pages [*DOI :* 10.1007/S40818-019-0059-2], https://hal.archives-ouvertes.fr/hal-01548737

[22] G. DI GESÙ, T. LELIÈVRE, D. LE PEUTREC, B. NECTOUX.*The exit from a metastable state: Concentration of the exit point distribution on the low energy saddle points, part 1*, in "Journal de Mathématiques Pures et Appliquées", June 2019 [*DOI :* 10.1016/J.MATPUR.2019.06.003], https://hal.archives-ouvertes.fr/hal-02383232

[23] G. FERRÉ, G. STOLTZ.*Error estimates on ergodic properties of discretized Feynman-Kac semigroups*, in "Numerische Mathematik", July 2019, vol. 143, n^{o} 2, p. 261–313, https://arxiv.org/abs/1712.04013 [*DOI :* 10.1007/S00211-019-01059-1], https://hal.archives-ouvertes.fr/hal-01690532

[24] D. GONTIER, A. LEVITT, S. SIRAJ-DINE.*Numerical construction of Wannier functions through homotopy*, in "Journal of Mathematical Physics", March 2019, https://arxiv.org/abs/1812.06746 [*DOI :* 10.1063/1.5085753], https://hal.archives-ouvertes.fr/hal-01955842

[25] T. J. HUDSON, F. LEGOLL, T. LELIÈVRE.*Stochastic homogenization of a scalar viscoelastic model exhibiting stress-strain hysteresis*, in "ESAIM: Mathematical Modelling and Numerical Analysis", 2019, https://arxiv.org/abs/1802.05549 - 35 pages, 4 figures, forthcoming [*DOI :* 10.1051/M2AN/2019081], https://hal.archives-ouvertes.fr/hal-01710772

[26] F. Hédin, T. Lelièvre.*gen.parRep: a first implementation of the Generalized Parallel Replica dynamics for the long time simulation of metastable biochemical systems*, in "Computer Physics Communications", 2019, vol. 239, p. 311-324, https://arxiv.org/abs/1807.02431 , forthcoming [*DOI :* 10.1016/j.cpc.2019.01.005], https://hal.archives-ouvertes.fr/hal-01832823

[27] A. Iacobucci, S. Olla, G. Stoltz.*Convergence rates for nonequilibrium Langevin dynamics*, in "Annales mathématiques du Québec", April 2019, vol. 43, n° 1, p. 73-98, https://arxiv.org/abs/1702.03685 [*DOI :* 10.1007/s40316-017-0091-0], https://hal.archives-ouvertes.fr/hal-01467131

[28] M. Josien.*Decomposition and pointwise estimates of periodic Green functions of some elliptic equations with periodic oscillatory coefficients*, in "Asymptotic Analysis", 2019, vol. 112, n° 3-4, p. 227–246, https://hal.archives-ouvertes.fr/hal-01848268

[29] C. Le Bris, F. Legoll, F. Madiot.*Multiscale Finite Element methods for advection-dominated problems in perforated domains*, in "Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal", 2019, vol. 17, n° 2, p. 773-825, https://arxiv.org/abs/1710.09331 [*DOI :* 10.1137/17M1152048], https://hal.archives-ouvertes.fr/hal-01624049

[30] F. Legoll, T. Lelièvre, U. Sharma.*Effective dynamics for non-reversible stochastic differential equations: a quantitative study*, in "Nonlinearity", 2019, vol. 32, n° 12, p. 4779-4816, https://arxiv.org/abs/1809.10498 [*DOI :* 10.1088/1361-6544/ab34bf], https://hal.archives-ouvertes.fr/hal-01883330

[31] T. Lelièvre, M. Rousset, G. Stoltz.*Hybrid Monte Carlo methods for sampling probability measures on submanifolds*, in "Numerische Mathematik", June 2019, vol. 143, n° 2, p. 379–421, https://arxiv.org/abs/1807.02356 [*DOI :* 10.1007/s00211-019-01056-4], https://hal.archives-ouvertes.fr/hal-01832820

[32] T. Lelièvre, G. Samaey, P. Zieliński.*Analysis of a micro-macro acceleration method with minimum relative entropy moment matching*, in "Stochastic Processes and their Applications", October 2019, https://arxiv.org/abs/1801.01740 - 40 pages [*DOI :* 10.1016/j.spa.2019.10.008], https://hal.archives-ouvertes.fr/hal-01680761

[33] T. Lelièvre, W. Zhang.*Pathwise Estimates for Effective Dynamics: The Case of Nonlinear Vectorial Reaction Coordinates*, in "Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal", July 2019, vol. 17, n° 3, p. 1019-1051, https://arxiv.org/abs/1805.01928 [*DOI :* 10.1137/18M1186034], https://hal.archives-ouvertes.fr/hal-02383239

[34] L. Lopes, T. Lelièvre.*Analysis of the adaptive multilevel splitting method on the isomerization of alanine dipeptide*, in "Journal of Computational Chemistry", April 2019, vol. 40, n° 11, p. 1198-1208, https://arxiv.org/abs/1707.00950 [*DOI :* 10.1002/jcc.25778], https://hal.archives-ouvertes.fr/hal-02383173

[35] J. Roussel, G. Stoltz.*A perturbative approach to control variates in molecular dynamics*, in "Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal", March 2019, vol. 17, n° 1, p. 552–591, https://arxiv.org/abs/1712.08022 [*DOI :* 10.1137/18M1171047], https://hal.archives-ouvertes.fr/hal-01709265

[36] B. Stamm, L. Lagardère, G. Scalmani, P. Gatto, E. Cancès, J.-P. Piquemal, Y. Maday, B. Mennucci, F. Lipparini.*How to make continuum solvation incredibly fast in a few simple steps: a practical guide to the domain decomposition paradigm for the Conductor-like Screening Model Continuum Solvation, Linear Scaling, Domain Decomposition*, in "International Journal of Quantum Chemistry", January 2019, vol. 119, e25669 [*DOI :* 10.1002/qua.25669], https://hal.archives-ouvertes.fr/hal-01793770

Conferences without Proceedings

[37] S. BRISARD, V. EHRLACHER.*Méthodes MsFEM pour les plaques hétérogènes*, in "CFM 2019 - 24ème Congrès Français de Mécanique", Brest, France, August 2019, https://hal-enpc.archives-ouvertes.fr/hal-02274963

Scientific Books (or Scientific Book chapters)

[38] G. GIACOMIN, S. OLLA, E. SAADA, H. SPOHN, G. STOLTZ.*Stochastic Dynamics Out of Equilibrium : Institut Henri Poincaré, Paris, France, 2017*, Springer, July 2019, vol. 282 [*DOI :* 10.1007/978-3-030-15096-9], https://hal.archives-ouvertes.fr/hal-02378069

[39] T. LELIÈVRE, D. LE PEUTREC, B. NECTOUX.*Exit event from a metastable state and Eyring-Kramers law for the overdamped Langevin dynamics*, in "Stochastic Dynamics Out of Equilibrium. IHPStochDyn 2017", Springer Proceedings in Mathematics & Statistics, July 2019, vol. 282, p. 331-363, https://arxiv.org/abs/1811.06786 [*DOI :* 10.1007/978-3-030-15096-9_9], https://hal.archives-ouvertes.fr/hal-01696187

Other Publications

[40] A. ALFONSI, R. COYAUD, V. EHRLACHER, D. LOMBARDI.*Approximation of Optimal Transport problems with marginal moments constraints*, May 2019, https://arxiv.org/abs/1905.05663 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02128374

[41] J. BERENDSEN, M. BURGER, V. EHRLACHER, J.-F. PIETSCHMANN.*Uniqueness of strong solutions and weak-strong stability in a system of cross-diffusion equations*, January 2019, https://arxiv.org/abs/1812.10711 - working paper or preprint [*DOI :* 10.1007/s00028-019-00534-4], https://hal.archives-ouvertes.fr/hal-01967899

[42] S. BOYAVAL.*Viscoelastic flows with conservation laws*, August 2019, https://arxiv.org/abs/1908.03344 - working paper or preprint, https://hal.inria.fr/hal-02262298

[43] S. BOYAVAL, S. MARTEL, J. REYGNER.*Finite-Volume approximation of the invariant measure of a viscous stochastic scalar conservation law*, September 2019, https://arxiv.org/abs/1909.08899 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02291253

[44] E. CANCÈS, G. DUSSON, Y. MADAY, B. STAMM, M. VOHRALÍK.*Guaranteed a posteriori bounds for eigenvalues and eigenvectors: multiplicities and clusters*, May 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02127954

[45] E. CANCÈS, G. DUSSON, Y. MADAY, B. STAMM, M. VOHRALÍK.*Post-processing of the planewave approximation of Schrödinger equations. Part I: linear operators*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01908039

[46] D. CHAFAÏ, G. FERRÉ, G. STOLTZ.*Coulomb gases under constraint: some theoretical and numerical results*, July 2019, https://arxiv.org/abs/1907.05803 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02184896

[47] L. CHAMOIN, F. LEGOLL.*Goal-oriented error estimation and adaptivity in MsFEM computations*, August 2019, https://arxiv.org/abs/1908.00367 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02262423

[48] G. DI GESÙ, T. LELIÈVRE, D. LE PEUTREC, B. NECTOUX.*The exit from a metastable state: concentration of the exit point distribution on the low energy saddle points*, February 2019, https://arxiv.org/abs/1902.03270 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02015493

[49] V. EHRLACHER, L. GRIGORI, D. LOMBARDI, H. SONG.*Adaptive hierarchical subtensor partitioning for tensor compression*, September 2019, working paper or preprint, https://hal.inria.fr/hal-02284456

[50] V. EHRLACHER, T. LELIÈVRE, P. MONMARCHÉ.*Adaptive force biasing algorithms: new convergence results and tensor approximations of the bias*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02314426

[51] V. EHRLACHER, D. LOMBARDI, O. MULA, F.-X. VIALARD.*Nonlinear model reduction on metric spaces. Application to one-dimensional conservative PDEs in Wasserstein spaces*, September 2019, https://arxiv.org/abs/1909.06626 - working paper or preprint, https://hal.inria.fr/hal-02290431

[52] G. FERRÉ, G. STOLTZ.*Large deviations of the empirical measure of diffusions in fine topologies with applications*, June 2019, https://arxiv.org/abs/1906.09411 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02164793

[53] S. GAUCHER, O. KLOPP, G. ROBIN.*Link Prediction in the Stochastic Block Model with Outliers*, November 2019, https://arxiv.org/abs/1911.13122 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02386940

[54] M. F. HERBST, M. SCHEURER, T. FRANSSON, D. R. REHN, A. DREUW.*adcc: A versatile toolkit for rapid development of algebraic-diagrammatic construction methods*, October 2019, https://arxiv.org/abs/1910.07757 - 30 pages, 7 figures, https://hal.archives-ouvertes.fr/hal-02319517

[55] B. JOURDAIN, T. LELIÈVRE, P.-A. ZITT.*Convergence of metadynamics: discussion of the adiabatic hypothesis*, April 2019, https://arxiv.org/abs/1904.08667 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02104961

[56] F. LEGOLL, T. LELIÈVRE, K. MYERSCOUGH, G. SAMAEY.*Parareal computation of stochastic differential equations with time-scale separation: a numerical study*, December 2019, https://arxiv.org/abs/1912.09240 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02421240

[57] B. LEIMKUHLER, M. SACHS, G. STOLTZ.*Hypocoercivity properties of adaptive Langevin dynamics*, August 2019, https://arxiv.org/abs/1908.09363 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02273261

[58] L. PILLAUD-VIVIEN, F. BACH, T. LELIÈVRE, A. RUDI, G. STOLTZ.*Statistical Estimation of the Poincaré constant and Application to Sampling Multimodal Distributions*, November 2019, https://arxiv.org/abs/1910.14564 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02327453

[59] G. STOLTZ, P. TERRIER.*A mathematical justification of the finite time approximation of Becker-Doring equations by a Fokker-Planck dynamics*, March 2019, https://arxiv.org/abs/1810.01462 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02065570

[60] Z. TRSTANOVA, B. LEIMKUHLER, T. LELIÈVRE.*Local and Global Perspectives on Diffusion Maps in the Analysis of Molecular Systems*, April 2019, https://arxiv.org/abs/1901.06936 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02104963

[61] T. WANG, G. STOLTZ, P. PLECHAC.*Convergence of the likelihood ratio method for linear response of non-equilibrium stationary states*, November 2019, https://arxiv.org/abs/1910.02479 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02385582

Project-Team MATHRISK

Mathematical Risk handling

IN COLLABORATION WITH: Centre d'Enseignement et de Recherche en Mathématiques et Calcul Scientifique (CERMICS)

IN PARTNERSHIP WITH:
CNRS

Ecole des Ponts ParisTech

Université Paris-Est Marne-la-Vallée

RESEARCH CENTER
Paris

THEME
Stochastic approaches

Table of contents

Project-Team MATHRISK

Creation of the Team: 2012 January 01, updated into Project-Team: 2013 January 01

Keywords:

Computer Science and Digital Science:

A6. - Modeling, simulation and control
A6.1. - Methods in mathematical modeling
A6.1.2. - Stochastic Modeling
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.2. - Numerical probability
A6.2.3. - Probabilistic methods
A6.4.2. - Stochastic control

Other Research Topics and Application Domains:

B3.1. - Sustainable development
B9.6.3. - Economy, Finance
B9.11. - Risk management

1. Team, Visitors, External Collaborators

Research Scientists

Agnès Bialobroda Sulem [Team leader, Inria, Senior Researcher, HDR]
Aurélien Alfonsi [École Nationale des Ponts et Chaussées, Senior Researcher, HDR]
Benjamin Jourdain [École Nationale des Ponts et Chaussées, Senior Researcher, HDR]
Bernard Lapeyre [École Nationale des Ponts et Chaussées, Senior Researcher, HDR]

Faculty Members

Vlad Bally [Univ Paris-Est Marne La Vallée, Professor, HDR]
Damien Lamberton [Univ Paris-Est Marne La Vallée, Professor, HDR]

External Collaborators

Antonino Zanette [University of Udine, HDR]
Ahmed Kebaier [Univ Paris-Nord]
Céline Labart [Univ de Savoie]
Jérôme Lelong [ENSIMAG, HDR]

Technical Staff

Pierre-Guillaume Raverdy [Inria, Engineer, until Mar 2019]

PhD Students

Oumaima Bencheikh [École Nationale des Ponts et Chaussées, PhD Student]
Rui Chen [Inria, PhD Student]
Anas Bentaleb [École Nationale des Ponts et Chaussées, PhD Student]
Adel Cherchali [École Nationale des Ponts et Chaussées, PhD Student]
Rafael Coyaud [École Nationale des Ponts et Chaussées, PhD Student]
Ezechiel Kahn [École Nationale des Ponts et Chaussées, PhD Student]
William Margheriti [École Nationale des Ponts et Chaussées, PhD Student]
Sophian Mehalla [École Nationale des Ponts et Chaussées, PhD Student]

Post-Doctoral Fellow

Alvin Tse [École Nationale des Ponts et Chaussées, Postdoc from November 2019]

Visiting Scientists

Justin Kirkby [Georgia Institute of Technology, from Jul 2019 until Aug 2019]
Oleg Kudryavtsev [Roskov University, from Jul 2019 until Aug 2019]
Xiao Wei [Beijing University, from Jul 2019 until Aug 2019]

Administrative Assistant

Derya Gok [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

The Inria project team **MathRisk** team was created in 2013. It is the follow-up of the MathFi project team founded in 2000. MathFi was focused on financial mathematics, in particular on computational methods for pricing and hedging increasingly complex financial products. The 2007 global financial crisis and its "aftermath crisis" has abruptly highlighted the critical importance of a better understanding and management of risk. The project **MathRisk** has been reoriented towards mathematical handling of risk, and addresses broad research topics embracing risk measurement and risk management, modeling and optimization in quantitative finance, but also in other related domains where risk control is paramount. The project team **MathRisk** aims both at producing mathematical tools and models in these domains, and developing collaborations with various institutions involved in risk control. Quantitative finance remains for the project an important source of mathematical problems and applications. Indeed, the pressure of new legislation leads to a massive reorientation of research priorities, and the interest of analysts shifted to risk control preoccupation.

The scientific issues related to quantitative finance we consider include systemic risk and contagion modeling, robust finance, market frictions, counterparty and liquidity risk, assets dependence modeling, market microstructure modeling and price impact. In this context, models must take into account the multidimensional feature and various market imperfections. They are much more demanding mathematically and numerically, and require the development of risk measures taking into account incompleteness issues, model uncertainties, interplay between information and performance and various defaults.

Besides, financial institutions, submitted to more stringent regulatory legislations such as FRTB or XVA computation, are facing practical implementation challenges which still need to be solved. Research focused on numerical efficiency remains strongly needed in this context, renewing the interest for the numerical platform Premia (http://www.premia.fr) that Mathrisk is developing in collaboration with a consortium of financial institutions.

While these themes arise naturally in the world of quantitative finance, a number of these issues and mathematical tools are also relevant to the treatment of risk in other areas as economy, social insurance and sustainable development, of fundamental importance in today's society. In these contexts, the management of risk appears at different time scales, from high frequency data to long term life insurance management, raising challenging renewed modeling and numerical issues.

The **MathRisk** project is strongly involved in the development of new mathematical methods and numerical algorithms. Mathematical tools include stochastic modeling, stochastic analysis, in particular stochastic (partial) differential equations and various aspects of stochastic control and optimal stopping of these equations, nonlinear expectations, Malliavin calculus, stochastic optimization, dynamic game theory, random graphs, martingale optimal transport (especially in relation to numerical considerations), long time behavior of Markov processes (with applications to Monte-Carlo methods) and generally advanced numerical methods for effective solutions.

3. Research Program

3.1. Risk management: modeling and optimization

3.1.1. Contagion modeling and systemic risk

After the recent financial crisis, systemic risk has emerged as one of the major research topics in mathematical finance. Interconnected systems are subject to contagion in time of distress. The scope is to understand and model how the bankruptcy of a bank (or a large company) may or not induce other bankruptcies. By contrast with the traditional approach in risk management, the focus is no longer on modeling the risks faced by a single financial institution, but on modeling the complex interrelations between financial institutions and the mechanisms of distress propagation among these.

The mathematical modeling of default contagion, by which an economic shock causing initial losses and default of a few institutions is amplified due to complex linkages, leading to large scale defaults, can be addressed by various techniques, such as network approaches (see in particular R. Cont et al. [45] and A. Minca [80]) or mean field interaction models (Garnier-Papanicolaou-Yang [73]).

We have contributed in the last years to the research on the control of contagion in financial systems in the framework of random graph models : In [46], [81], [5], A. Sulem with A. Minca and H. Amini consider a financial network described as a weighted directed graph, in which nodes represent financial institutions and edges the exposures between them. The distress propagation is modeled as an epidemics on this graph. They study the optimal intervention of a lender of last resort who seeks to make equity infusions in a banking system prone to insolvency and to bank runs, under complete and incomplete information of the failure cluster, in order to minimize the contagion effects. The paper [5] provides in particular important insight on the relation between the value of a financial system, connectivity and optimal intervention.

The results show that up to a certain connectivity, the value of the financial system increases with connectivity. However, this is no longer the case if connectivity becomes too large. The natural question remains how to create incentives for the banks to attain an optimal level of connectivity. This is studied in [58], where network formation for a large set of financial institutions represented as nodes is investigated. Linkages are source of income, and at the same time they bear the risk of contagion, which is endogeneous and depends on the strategies of all nodes in the system. The optimal connectivity of the nodes results from a game. Existence of an equilibrium in the system and stability properties is studied. The results suggest that financial stability is best described in terms of the mechanism of network formation than in terms of simple statistics of the network topology like the average connectivity.

3.1.2. Liquidity risk and Market Microstructure

Liquidity risk is the risk arising from the difficulty of selling (or buying) an asset. Usually, assets are quoted on a market with a Limit Order Book (LOB) that registers all the waiting limit buy and sell orders for this asset. The bid (resp. ask) price is the most expensive (resp. cheapest) waiting buy or sell order. If a trader wants to sell a single asset, he will sell it at the bid price, but if he wants to sell a large quantity of assets, he will have to sell them at a lower price in order to match further waiting buy orders. This creates an extra cost, and raises important issues. From a short-term perspective (from few minutes to some days), it may be interesting to split the selling order and to focus on finding optimal selling strategies. This requires to model the market microstructure, i.e. how the market reacts in a short time-scale to execution orders. From a long-term perspective (typically, one month or more), one has to understand how this cost modifies portfolio managing strategies (especially delta-hedging or optimal investment strategies). At this time-scale, there is no need to model precisely the market microstructure, but one has to specify how the liquidity costs aggregate.

For rather liquid assets, liquidity risk is usually taken into account via price impact models which describe how a (large) trader influences the asset prices. Then, one is typically interested in the optimal execution problem: how to buy/sell a given amount of assets optimally within a given deadline. This issue is directly related to the existence of statistical arbitrage or Price Manipulation Strategies (PMS). Most of price impact models deal with single assets. A. Alfonsi, F. Klöck and A. Schied [44] have proposed a multi-assets price impact model that extends previous works. Price impact models are usually relevant when trading at an intermediary frequency (say every hour). At a lower frequency, price impact is usually ignored while at a high frequency (every minute or second), one has to take into account the other traders and the price jumps, tick by tick. Midpoint price models are thus usually preferred at this time scale. With P. Blanc, Alfonsi [3] has proposed a model that makes a bridge between these two types of model: they have considered an (Obizhaeva and Wang) price impact model, in which the flow of market orders generated by the other traders is given by an exogeneous process. They have shown that Price Manipulation Strategies exist when the flow of order is a compound Poisson process. However, modeling this flow by a mutually exciting Hawkes process with a particular parametrization allows them to exclude these PMS. Besides, the optimal execution strategy is explicit in this model. A practical implementation is given in [40].

3.1.3. Dependence modeling

- Calibration of stochastic and local volatility models. The volatility is a key concept in modern mathematical finance, and an indicator of market stability. Risk management and associated instruments depend strongly on the volatility, and volatility modeling is a crucial issue in the finance industry. Of particular importance is the assets *dependence* modeling.

By Gyongy's theorem, a local and stochastic volatility model is calibrated to the market prices of all call options with positive maturities and strikes if its local volatility function is equal to the ratio of the Dupire local volatility function over the root conditional mean square of the stochastic volatility factor given the spot value. This leads to a SDE nonlinear in the sense of McKean. Particle methods based on a kernel approximation of the conditional expectation, as presented by Guyon and Henry-Labordère [74], provide an efficient calibration procedure even if some calibration errors may appear when the range of the stochastic volatility factor is very large. But so far, no existence result is available for the SDE nonlinear in the sense of McKean. In the particular case when the local volatility function is equal to the inverse of the root conditional mean square of the stochastic volatility factor multiplied by the spot value given this value and the interest rate is zero, the solution to the SDE is a fake Brownian motion. When the stochastic volatility factor is a constant (over time) random variable taking finitely many values and the range of its square is not too large, B. Jourdain and A. Zhou proved existence to the associated Fokker-Planck equation [21]. Thanks to results obtained by Figalli in [69], they deduced existence of a new class of fake Brownian motions. They extended these results to the special case of the LSV model called Regime Switching Local Volatility, when the stochastic volatility factor is a jump process taking finitely many values and with jump intensities depending on the spot level.

- Interest rates modeling. Affine term structure models have been popularized by Dai and Singleton [59], Duffie, Filipovic and Schachermayer [60]. They consider vector affine diffusions (the coordinates are usually called factors) and assume that the short interest rate is a linear combination of these factors. A model of this kind is the Linear Gaussian Model (LGM) that considers a vector Ornstein-Uhlenbeck diffusions for the factors, see El Karoui and Lacoste [68]. A. Alfonsi et al. [37] have proposed an extension of this model, when the instantaneous covariation between the factors is given by a Wishart process. Doing so, the model keeps its affine structure and tractability while generating smiles for option prices. A price expansion around the LGM is obtained for Caplet and Swaption prices.

3.1.4. Robust finance

- Numerical Methods for Martingale Optimal Transport problems.

The Martingale Optimal Transport (MOT) problem introduced in [57] has received a recent attention in finance since it gives model-free hedges and bounds on the prices of exotic options. The market prices of liquid call and put options give the marginal distributions of the underlying asset at each traded maturity. Under the simplifying assumption that the risk-free rate is zero, these probability measures are in increasing convex order, since by Strassen's theorem this property is equivalent to the existence of a martingale measure with the right marginal distributions. For an exotic payoff function of the values of the underlying on the time-grid given by these maturities, the model-free upper-bound (resp. lower-bound) for the price consistent with these marginal distributions is given by the following martingale optimal transport problem : maximize (resp.

minimize) the integral of the payoff with respect to the martingale measure over all martingale measures with the right marginal distributions. Super-hedging (resp. sub-hedging) strategies are obtained by solving the dual problem. With J. Corbetta, A. Alfonsi and B. Jourdain [12] have studied sampling methods preserving the convex order for two probability measures μ and ν on $\mathbf{R}^d$, with ν dominating μ.

Their method is the first generic approach to tackle the martingale optimal transport problem numerically and can also be applied to several marginals.

- Robust option pricing in financial markets with imperfections.

A. Sulem, M.C. Quenez and R. Dumitrescu have studied robust pricing in an imperfect financial market with default. The market imperfections are taken into account via the nonlinearity of the wealth dynamics. In this setting, the pricing system is expressed as a nonlinear g-expectation $\mathcal{E}^g$ induced by a nonlinear BSDE with nonlinear driver g and default jump (see [61]). A large class of imperfect market models can fit in this framework, including imperfections coming from different borrowing and lending interest rates, taxes on profits from risky investments, or from the trading impact of a large investor seller on the market prices and the default probability. Pricing and superhedging issues for American and game options in this context and their links with optimal stopping problems and Dynkin games with nonlinear expectation have been studied. These issues have also been addressed in the case of model uncertainty, in particular uncertainty on the default probability. The seller's robust price of a game option has been characterized as the value function of a Dynkin game under $\mathcal{E}^g$ expectation as well as the solution of a nonlinear doubly reflected BSDE in [9]. Existence of robust superhedging strategies has been studied. The buyer's point of view and arbitrage issues have also been studied in this context.

In a Markovian framework, the results of the paper [8] on combined optimal stopping/stochastic control with $\mathcal{E}^g$ expectation allows us to address American nonlinear option pricing when the payoff function is only Borelian and when there is ambiguity both on the drift and the volatility of the underlying asset price process. Robust optimal stopping of dynamic risk measures induced by BSDEs with jumps with model ambiguity is studied in [83].

3.2. Perspectives in Stochastic Analysis

3.2.1. Optimal transport and longtime behavior of Markov processes

The dissipation of general convex entropies for continuous time Markov processes can be described in terms of backward martingales with respect to the tail filtration. The relative entropy is the expected value of a backward submartingale. In the case of (non necessarily reversible) Markov diffusion processes, J. Fontbona and B. Jourdain [71] used Girsanov theory to explicit the Doob-Meyer decomposition of this submartingale. They deduced a stochastic analogue of the well known entropy dissipation formula, which is valid for general convex entropies, including the total variation distance. Under additional regularity assumptions, and using Itô's calculus and ideas of Arnold, Carlen and Ju [47], they obtained a new Bakry-Emery criterion which ensures exponential convergence of the entropy to 0. This criterion is non-intrinsic since it depends on the square root of the diffusion matrix, and cannot be written only in terms of the diffusion matrix itself. They provided examples where the classic Bakry Emery criterion fails, but their non-intrinsic criterion applies without modifying the law of the diffusion process.

With J. Corbetta, A. Alfonsi and B. Jourdain have studied the time derivative of the Wasserstein distance between the marginals of two Markov processes [11]. The Kantorovich duality leads to a natural candidate for this derivative. Up to the sign, it is the sum of the integrals with respect to each of the two marginals of the corresponding generator applied to the corresponding Kantorovich potential. For pure jump processes with bounded intensity of jumps, J. Corbetta, A. Alfonsi and B. Jourdain [41] proved that the evolution of the Wasserstein distance is actually given by this candidate. In dimension one, they showed that this remains true for Piecewise Deterministic Markov Processes. They applied the formula to estimate the exponential decrease rate of the Wasserstein distance between the marginals of two birth and death processes with the same generator in terms of the Wasserstein curvature.

3.2.2. Mean-field systems: modeling and control

- Mean-field limits of systems of interacting particles. In [77], B. Jourdain and his former PhD student J. Reygner have studied a mean-field version of rank-based models of equity markets such as the Atlas model introduced by Fernholz in the framework of Stochastic Portfolio Theory. They obtained an asymptotic description of the market when the number of companies grows to infinity. Then, they discussed the long-term capital distribution, recovering the Pareto-like shape of capital distribution curves usually derived from empirical studies, and providing a new description of the phase transition phenomenon observed by Chatterjee and Pal. They have also studied multitype sticky particle systems which can be obtained as vanishing noise limits of multitype rank-based diffusions (see [76]). Under a uniform strict hyperbolicity assumption on the characteristic fields, they constructed a multitype version of the sticky particle dynamics. In [78], they obtain the optimal rate of convergence as the number of particles grows to infinity of the approximate solutions to the diagonal hyperbolic system based on multitype sticky particles and on easy to compute time discretizations of these dynamics.

In [72], N. Fournier and B. Jourdain are interested in the two-dimensional Keller-Segel partial differential equation. This equation is a model for chemotaxis (and for Newtonian gravitational interaction).

- Mean field control and Stochastic Differential Games (SDGs). To handle situations where controls are chosen by several agents who interact in various ways, one may use the theory of Stochastic Differential Games (SDGs). Forward–Backward SDG and stochastic control under Model Uncertainty are studied in [84] by A. Sulem and B. Øksendal. Also of interest are large population games, where each player interacts with the average effect of the others and individually has negligible effect on the overall population. Such an interaction pattern may be modeled by mean field coupling and this leads to the study of mean-field stochastic control and related SDGs. A. Sulem, Y. Hu and B. Øksendal have studied singular mean field control problems and singular mean field two-players stochastic differential games [75]. Both sufficient and necessary conditions for the optimal controls and for the Nash equilibrium are obtained. Under some assumptions, the optimality conditions for singular mean-field control are reduced to a reflected Skorohod problem. Applications to optimal irreversible investments under uncertainty have been investigated. Predictive mean-field equations as a model for prices influenced by beliefs about the future are studied in [86].

3.2.3. Stochastic control and optimal stopping (games) under nonlinear expectation

M.C. Quenez and A. Sulem have studied optimal stopping with nonlinear expectation $\mathcal{E}^g$ induced by a BSDE with jumps with nonlinear driver g and irregular obstacle/payoff (see [83]). In particular, they characterize the value function as the solution of a reflected BSDE. This property is used in [67] to address American option pricing in markets with imperfections. The Markovian case is treated in [64] when the payoff function is continuous.

In [8], M.C. Quenez, A. Sulem and R. Dumitrescu study a combined optimal control/stopping problem under nonlinear expectation $\mathcal{E}^g$ in a Markovian framework when the terminal reward function is only Borelian. In this case, the value function u associated with this problem is irregular in general. They establish a *weak* dynamic programming principle (DPP), from which they derive that the upper and lower semi-continuous envelopes of u are the sub- and super- *viscosity solution* of an associated nonlinear Hamilton-Jacobi-Bellman variational inequality.

The problem of a generalized Dynkin game problem with nonlinear expectation $\mathcal{E}^g$ is addressed in [65]. Under Mokobodzki's condition, we establish the existence of a value function for this game, and characterize this value as the solution of a doubly reflected BSDE. The results of this work are used in [9] to solve the problem of game option pricing in markets with imperfections.
A generalized mixed game problem when the players have two actions: continuous control and stopping is studied in a Markovian framework in [66]. In this work, dynamic programming principles (DPP) are established: a strong DPP is proved in the case of a regular obstacle and a weak one in the irregular case. Using these DPPs, links with parabolic partial integro-differential Hamilton-Jacobi- Bellman variational inequalities with two obstacles are obtained.

With B. Øksendal and C. Fontana, A. Sulem has contributed on the issues of robust utility maximization [85], [86], and relations between information and performance [70].

3.2.4. Generalized Malliavin calculus

Vlad Bally has extended the stochastic differential calculus built by P. Malliavin which allows one to obtain integration by parts and associated regularity probability laws. In collaboration with L. Caramellino (Tor Vegata University, Roma), V. Bally has developed an abstract version of Malliavin calculus based on a splitting method (see [49]). It concerns random variables with law locally lower bounded by the Lebesgue measure (the so-called Doeblin's condition). Such random variables may be represented as a sum of a "smooth" random variable plus a rest. Based on this smooth part, he achieves a stochastic calculus which is inspired from Malliavin calculus [6]. An interesting application of such a calculus is to prove convergence for irregular test functions (total variation distance and more generally, distribution distance) in some more or less classical frameworks as the Central Limit Theorem, local versions of the CLT and moreover, general stochastic polynomials [53]. An exciting application concerns the number of roots of trigonometric polynomials with random coefficients [15]. Using Kac Rice lemma in this framework one comes back to a multidimensional CLT and employs Edgeworth expansions of order three for irregular test functions in order to study the mean and the variance of the number of roots. Another application concerns U statistics associated to polynomial functions. The techniques of generalized Malliavin calculus developed in [49] are applied in for the approximation of Markov processes (see [56] and [55]). On the other hand, using the classical Malliavin calculus, V. Bally in collaboration with L. Caramellino and P. Pigato studied some subtle phenomena related to diffusion processes, as short time behavior and estimates of tubes probabilities (see [51], [52], [50]).

3.3. Numerical Probability

Our project team is very much involved in numerical probability, aiming at pushing numerical methods towards the effective implementation. This numerical orientation is supported by a mathematical expertise which permits a rigorous analysis of the algorithms and provides theoretical support for the study of rates of convergence and the introduction of new tools for the improvement of numerical methods. This activity in the MathRisk team is strongly related to the development of the Premia software.

3.3.1. Simulation of stochastic differential equations

3.3.1.1. - Weak convergence of the Euler scheme in optimal transport distances.

With A. Kohatsu-Higa, A. Alfonsi and B. Jourdain [4] have proved using optimal transport tools that the Wasserstein distance between the time marginals of an elliptic SDE and its Euler discretization with N steps is not larger than $C\sqrt{\log(N)}/N$. The logarithmic factor may is removed when the uniform time-grid is replaced by a grid still counting N points but refined near the origin of times.

3.3.1.2. - Strong convergence properties of the Ninomiya Victoir scheme and multilevel Monte-Carlo estimators.

With their former PhD student, A. Al Gerbi, E. Clément and B. Jourdain [1] have proved strong convergence with order $1/2$ of the Ninomiya-Victoir scheme which is known to exhibit order 2 of weak convergence [82]. This study was aimed at analysing the use of this scheme either at each level or only at the finest level of a multilevel Monte Carlo estimator : indeed, the variance of a multilevel Monte Carlo estimator is related to the strong error between the two schemes used in the coarse and fine grids at each level. In [38], they proved that the order of strong convergence of the crude Ninomiya Victoir scheme is improved to 1 when the vector fields corresponding to each Brownian coordinate in the SDE commute, and in [39], they studied the error introduced by discretizing the ordinary differential equations involved in the Ninomiya-Victoir scheme.

3.3.1.3. - Non-asymptotic error bounds for the multilevel Monte Carlo Euler method.

A. Kebaier and B. Jourdain are interested in deriving non-asymptotic error bounds for the multilevel Monte Carlo method. As a first step, they dealt in [20] with the explicit Euler discretization of stochastic differential equations with a constant diffusion coefficient. They obtained Gaussian-type concentration. To do so, they used the Clark-Ocone representation formula and derived bounds for the moment generating functions of the squared difference between a crude Euler scheme and a finer one and of the squared difference of their Malliavin derivatives. The estimation of such differences is much more complicated than the one of a single Euler scheme contribution and explains why they suppose the diffusion coefficient to be constant. This assumption ensures boundedness of the Malliavin derivatives of both the SDE and its Euler scheme.

3.3.1.4. - Computation of sensibilities of integrals with respect to the invariant measure.

In [48], R. Assaraf, B. Jourdain, T. Lelièvre and R. Roux considered the solution to a stochastic differential equation with constant diffusion coefficient and with a drift function which depends smoothly on some real parameter λ, and admitting a unique invariant measure for any value of λ around $\lambda = 0$. Their aim was to compute the derivative with respect to λ of averages with respect to the invariant measure, at $\lambda = 0$. They analyzed a numerical method which consists in simulating the process at $\lambda = 0$ together with its derivative with respect to λ on a long time horizon. They gave sufficient conditions implying uniform-in-time square integrability of this derivative. This allows in particular to compute efficiently the derivative with respect to λ of the mean of an observable through Monte Carlo simulations.

3.3.1.5. - Approximation of doubly reflected Backward stochastic differential equations.

R. Dumitrescu and C. Labart have studied the discrete time approximation scheme for the solution of a doubly reflected Backward Stochastic Differential Equation with jumps, driven by a Brownian motion and an independent compensated Poisson process [63], [62].

3.3.1.6. - Parametrix methods.

V. Bally and A. Kohatsu-Higa have recently proposed an unbiased estimator based on the parametrix method to compute expectations of functions of a given SDE ([54]). This method is very general, and A. Alfonsi, A. Kohastu-Higa and M. Hayashi [42] have applied it to the case of one-dimensional reflected diffusions. In this case, the estimator can be obtained explicitly by using the scheme of Lépingle [79] and is quite simple to implement. It is compared to other simulation methods for reflected SDEs.

3.3.2. Estimation of the parameters of a Wishart process

A. Alfonsi, A. Kebaier and C. Rey [43] have computed the Maximum Likelihood Estimator for the Wishart process and studied its convergence in the ergodic and in some non ergodic cases. In the ergodic case, which is the most relevant for applications, they obtain the standard square-root convergence. In the non ergodic case, the analysis rely on refined results for the Laplace transform of Wishart processes, which are of independent interest.

3.3.3. Optimal stopping and American options

In joint work with A. Bouselmi, D. Lamberton studied the asymptotic behavior of the exercise boundary near maturity for American put options in exponential Lévy models. In [7], they deal with jump-diffusion models, and establish that, in some cases, the behavior differs from the classical Black and Scholes setting. D. Lamberton has also worked on the binomial approximation of the American put. The conjectured rate of convergence is $O(1/n)$ where n is the number of time periods. He was able to derive a $O((\ln n)^{\alpha}/n)$ bound, where the exponent α is related to the asymptotic behavior of the exercise boundary near maturity.

4. Application Domains

4.1. Financial Mathematics, Insurance

The domains of application are quantitative finance and insurance with emphasis on risk modeling and control. In particular, Mathrisk focuses on dependence modeling, systemic risk, market microstructure modeling and risk measures.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

Aurélien Alfonsi : Award for the Best Young Researcher in Finance and Insurance, Europlace Institute of Finance and SCOR Corporate Foundation.

5.1.2. Publications

B. Øksendal and A. Sulem. *Applied Stochastic Control of Jump Diffusions*. 3rd edition (436 pages) 2019, Universitext, Springer Verlag, Berlin, Heidelberg, New York [23].

6. New Software and Platforms

6.1. PREMIA

KEYWORDS: Financial products - Computational finance - Option pricing

SCIENTIFIC DESCRIPTION: The Premia project keeps track of the most recent advances in the field of computational finance in a well-documented way. It focuses on the implementation of numerical analysis techniques for both probabilistic and deterministic numerical methods. An important feature of the platform Premia is the detailed documentation which provides extended references in option pricing.

Premia is thus a powerful tool to assist Research and Development professional teams in their day-to-day duty. It is also a useful support for academics who wish to perform tests on new algorithms or pricing methods without starting from scratch.

Besides being a single entry point for accessible overviews and basic implementations of various numerical methods, the aim of the Premia project is: 1 - to be a powerful testing platform for comparing different numerical methods between each other, 2 - to build a link between professional financial teams and academic researchers, 3 - to provide a useful teaching support for Master and PhD students in mathematical finance.

FUNCTIONAL DESCRIPTION: Premia is a software designed for option pricing, hedging and financial model calibration.

- Participants: Agnes Sulem, Antonino Zanette, Aurélien Alfonsi, Benjamin Jourdain, Jérôme Lelong and Bernard Lapeyre
- Partners: Inria - Ecole des Ponts ParisTech - Université Paris-Est
- Contact: Agnes Sulem
- URL: http://www.premia.fr

6.2. Premia

6.2.1. Development of the quantitative platform Premia in 2019

Premia 21 has been delivered to the Premia Consortium on March 21th 2019.
In this version, the following algorithms have been implemented:

6.2.1.1. Machine Learning, Risk Management, Risk model, Insurance

- Machine Learning for Quantitative Finance: Fast Derivative Pricing, Hedging and Fitting. J. De Spiegeleer, D. B. Madan, S. Reyners, W. Schoutens
Quantitative Finance 2018
- Gaussian Process Regression for Pricing Variable Annuities with Stochastic Volatility and Interest Rate. L.Goudenege A.Molent A.Zanette 2019
- Machine Learning for Pricing American Options in High Dimension. L.Goudenege A.Molent A.Zanette 2019
- Neural networks for American options. L.Goudenege
- Sampling of probability measures in the convex order and approximation of Martingale Optimal Transport problems. A. Alfonsi, J. Corbetta, B. Jourdain
- Computing Credit Valuation Adjustment solving coupled PIDEs in the Bates model. L. Goudenege, A. Molent, A. Zanette
Computational Management Science, to appear 2019
- Valuation of variable annuities with Guaranteed Minimum Withdrawal Benefit under stochastic interest rate. P.V.Shevchenko, X.Luo
Insurance: Mathematics and Economics, Volume 76, 2017

6.2.1.2. Equity Derivatives

- An adjoint method for the exact calibration of Stochastic Local Volatility models. M.Wins, K.J.in 't Hout.
Journal of Computational Science, Volume 24, 2018
- Discretization of class of diffusions nonlinear in the sense of McKean including the calibrated LVSV model. B. Jourdain, A. Zhou
- On an efficient multiple time-step Monte Carlo simulation of the SABR model. A. Leitao, L.A. Grzelak, C.W. Oosterlee
Applied Mathematics and Computation, Vol. 293, 2017
- Robust Barrier Option Pricing by Frame Projection under Exponential Levy Dynamics. J.L. Kirkby
Applied Mathematical Finance, Volume 24, Issue 4, 2017
- Ultra-Fast Pricing Barrier Options and CDSs. S.Levendorskii
International Journal of Theoretical and Applied Finance, 20(5), 2017
- Efficient Binomial tree for the discretization of the CEV model. L.Caramellino, E.Lombardo.
- Pricing path-dependent Bermudan options using Wiener chaos expansion: an embarrassingly parallel approach. J.Lelong

We benefited from the help of the engineer Pierre-Guillaume Raverdy.

7. New Results

7.1. Control of systemic risk in a dynamic framework

Agnès Sulem, Andreea Minca (Cornell University), Hamed Amini (J. Mack Robinson College of Business, Georgia State University) and Rui Chen have studied a Dynamic Contagion Risk Model With Recovery Features [27]. In this paper, they introduce threshold growth in the classical threshold contagion model, in which nodes have downward jumps when there is a failure of a neighboring node. Choosing the configuration model as underlying graph, they prove fluid limits for the baseline model, as well as extensions to the directed case, state-dependent inter-arrival times and the case of growth driven by upward jumps. They obtain explicit ruin probabilities for the nodes according to their characteristics: initial threshold and in- (and out-) degree. They then allow nodes to choose their connectivity by trading off link benefits and contagion risk. They define a rational equilibrium concept in which nodes choose their connectivity according to an expected failure probability of any given link, and then impose condition that the expected failure probability coincides with the actual failure probability under the optimal connectivity. Existence of an asymptotic equilibrium is shown as well as convergence of the sequence of equilibria on the finite networks. In particular, these results show that systems with higher overall growth may have higher failure probability in equilibrium.

The results have been presented in Lisbon at the COMPLEX NETWORKS 2019 conference. Rui Chen has defended his thesis in July 2019 on this topic [10].

7.2. Mean-field BSDEs and systemic risk measures

Agnès Sulem with her PhD student Rui Chen, Andreea Minca and Roxana Dumitrescu have studied mean-field BSDEs with a generalized mean-field operator that can capture the average intensity in an inhomogeneous random graph. Comparison and strict comparison results have been obtained. Based on these, they interpret the BSDE solution as a global dynamic risk measure that can account for the intensity of system interactions and therefore incorporate systemic risk. Using Fenchel-Legendre transforms, they establish a dual representation for the expectation of the risk measure, and exhibit its dependence on the mean-field operator [31].

7.3. Risk management in finance and insurance

7.3.1. Option pricing in a non-linear incomplete market model with default

Agnès Sulem has studied with Miryana Grigorova (University of Leeds) and Marie-Claire Quenez (Université Paris Denis Diderot) superhedging prices and the associated superhedging strategies for both European and American options (see [33] and [32] in a non-linear incomplete market model with default. The underlying market model consists of a risk-free asset and a risky asset driven by a Brownian motion and a compensated default martingale. The portfolio processes follow non-linear dynamics with a non-linear driver f.

7.3.2. Neural network regression for Bermudan option pricing

The pricing of Bermudan options amounts to solving a dynamic programming principle, in which the main difficulty, especially in high dimension, comes from the conditional expectation involved in the computation of the continuation value. These conditional expectations are classically computed by regression techniques on a finite dimensional vector space. In [36], Bernard Lapeyre and Jérôme Lelong study neural networks approximations of conditional expectations. They prove the convergence of the well-known Longstaff and Schwartz algorithm when the standard least-square regression is replaced by a neural network approximation. They illustrate the numerical efficiency of neural networks as an alternative to standard regression methods for approximating conditional expectations on several numerical examples.

7.3.3. Hybrid numerical method for option pricing

With Giulia Terenzi, Lucia Caramellino (Tor Vegata University), and Maya Briani (CNR Roma), Antonino Zanette develop and study stability properties of a hybrid approximation of functionals of the Bates jump model with stochastic interest rate that uses a tree method in the direction of the volatility and the interest rate and a finite-difference approach in order to handle the underlying asset price process. They also propose hybrid simulations for the model, following a binomial tree in the direction of both the volatility and the interest rate, and a space-continuous approximation for the underlying asset price process coming from a Euler–Maruyama type scheme. They test their numerical schemes by computing European and American option prices [17].

7.3.4. American options

With his PhD student Giulia Terenzi, Damien Lamberton has been working on American options in Heston's model [22]. He is currently preparing his contribution to a winter school on "Theory and practice of optimal stopping and free boundary problems" (cf. https://conferences.leeds.ac.uk/osfbp/).

7.3.5. Solvency Capital Requirement in Insurance

A. Alfonsi has obtained a grant from AXA Foundation on a Joint Research Initiative with a team of AXA France working on the strategic asset allocation. This team has to make recommendations on the investment over some assets classes as, for example, equity, real estate or bonds. In order to do that, each side of the balance sheet (assets and liabilities) is modeled in order to take into account their own dynamics but also their interactions. Given that the insurance products are long time contracts, the projections of the company's margins have to be done considering long maturities. When doing simulations to assess investment policies, it is necessary to take into account the SCR which is the amount of cash that has to be settled to manage the portfolio. Typically, the computation of the future values of the SCR involve expectations under conditional laws, which is greedy in computation time.

A. Alfonsi and his PhD student A. Cherchali have developed a model of the ALM management of insurance companies that takes into account the regulatory constraints on life-insurance [25]. We now focus on developing Multilevel Monte-Carlo methods to approximate the SCR (Solvency Capital Requirement).

7.3.6. Pricing and hedging variable annuities of GMWB type in advanced stochastic models

Antonino Zanette with Ludovic Goudenège (Ecole Centrale de Paris) and Andrea Molent (University of Udine) study the valuation of a particular type of variable annuity called GMWB when advanced stochastic models are considered. As remarked by Yang and Dai (Insur Math Econ 52(2):231–242, 2013), and Dai et al. (Insur Math Econ 64:364–379, 2015), the Black–Scholes framework seems to be inappropriate for such a long maturity products. Also Chen et al. (Insur Math Econ 43(1):165–173, 2008) show that the price of GMWB variable annuities is very sensitive to the interest rate and the volatility parameters. They propose here to use a stochastic volatility model (the Heston model) and a Black–Scholes model with stochastic interest rate (the Black–Scholes Hull–White model). For this purpose, they consider four numerical methods: a hybrid tree-finite difference method, a hybrid tree-Monte Carlo method, an ADI finite difference scheme and a Standard Monte Carlo method. These approaches are employed to determine the no-arbitrage fee for a popular version of the GMWB contract and to calculate the Greeks used in hedging. Both constant withdrawal and dynamic withdrawal strategies are considered. Numerical results are presented, which demonstrate the sensitivity of the no-arbitrage fee to economic and contractual assumptions as well as the different features of the proposed numerical methods [18].

7.4. Stochastic Analysis and probabilistic numerical methods

7.4.1. Particles approximation of mean-field SDEs

O. Bencheikh and Benjamin Jourdain analysed the rate of convergence of a system of N interacting particles with mean-field rank based interaction in the drift coefficient and constant diffusion coefficient [16], [30]. They first adapted arguments by Kolli and Shkolnikhov to check trajectorial propagation of chaos with optimal rate $N^{-1/2}$ to the associated stochastic differential equations nonlinear in the sense of McKean. They next relaxed the assumption needed by Bossy to check convergence in $L^1(\mathbf{R})$ of the empirical cumulative distribution function of the Euler discretization with step h of the particle system to the solution of a one dimensional viscous scalar conservation law with rate $\mathcal{O}\left(\frac{1}{\sqrt{N}}+h\right)$. Last, they proved that the bias of this stochastic particle method behaves in $\mathcal{O}\left(\frac{1}{N}+h\right)$, which is confirmed by numerical experiments.

7.4.2. Abstract Malliavin calculus and convergence in total variation

In collaboration with L. Caramellino (University Tor Vergata) and with G. Poly (University of Rennes), V. Bally has settled a Malliavin type calculus for a general class of random variables, which are not supposed to be Gaussian (as it is the case in the standard Malliavin calculus). This is an alternative to the Γ calculus settled by Bakry, Gentile and Ledoux. The main application is the estimate in total variation distance of the error in general convergence theorems. This is done in [29].

7.4.3. Invariance principles

As an application of the methodology mentioned above, V. Bally and coauthors have studied several limit theorems of Central Limit type - (see [14] and [15]. In particular they have estimate the total variation distance between random polynomials on one hand, and proved an universality principle for the variance of the number of roots of trigonometric polynomials with random coefficients, on the other hand.

7.4.4. Regularity of the low of the solution of jump type equations

V. Bally, L. Caramellino and G. Poly obtained some new regularity results for the solution of the 2 dimensional Bolzmann equation (see [13]). Moreover, in collaboration with L. Caramellino and A. Kohatsu Higa, V. Bally has started a research program on the regularity of the solutions of jump type equations. A first result in this sense is contained in [28].

7.4.5. Approximation of ARCH models

Benjamin Jourdain and Gilles Pagès (LPSM) are interested in proposing approximations of a sequence of probability measures in the convex order by finitely supported probability measures still in the convex order [35]. They propose to alternate transitions according to a martingale Markov kernel mapping a probability measure in the sequence to the next and dual quantization steps. In the case of ARCH models and in particular of the Euler scheme of a driftless Brownian diffusion, the noise has to be truncated to enable the dual quantization step. They analyze the error between the original ARCH model and its approximation with truncated noise and exhibit conditions under which the latter is dominated by the former in the convex order at the level of sample-paths. Last, they analyse the error of the scheme combining the dual quantization steps with truncation of the noise according to primal quantization.

7.4.6. Convergence of metadynamics

By drawing a parallel between metadynamics and self interacting models for polymers, B. Jourdain, T. Leliëvre (Cermics / ENPC) and P.-A. Zitt (LAMA) study the longtime convergence of the original metadynamics algorithm in the adiabatic setting, namely when the dynamics along the collective variables decouples from the dynamics along the other degrees of freedom. They also discuss the bias which is introduced when the adiabatic assumption does not holds [34].

7.4.7. Optimal transport

With V. Ehrlacher, D. Lombardi and R. Coyaud, Aurelien Alfonsi isworking on numerical a approximations of the optimal transport between two (or more) probability measures [26].

7.4.8. Generic approximation schemes for Markov semigroups.

A. Alfonsi and V. Bally have produced a general approximation scheme for Markov semigroups, based on random grids. This is a new approach to approximation schemes which is an alternative to the multi level method and the Romberg method [24].

7.4.9. Approximation with rough paths

A. Alfonsi and A. Kebaier are working on the approximation of some processes with rough paths.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Consortium PREMIA, Natixis - Inria
- Consortium PREMIA, Crédit Agricole Corporate Investment Bank (CA - CIB) - Inria
- AXA Joint Research Initiative on Numerical methods for the ALM, from September 2017 to August 2020. PhD grant of Adel Cherchali, Supervisor: A. Alfonsi.
- CIFRE agreement Milliman company/Ecole des Ponts (http://fr.milliman.com),

  PhD thesis of Sophian Mehalla (started November 2017) on "Interest rate risk modeling for insurance companies", Supervisor: Bernard Lapeyre.
- Collaboration with IRT Systemx

  PhD grant of Adrien Touboul (started November 2017) on "Uncertainty computation in a graph of physical simulations", Supervisors: Bernard Lapeyre and Julien Reygner.

8.2. Grants with Industry

Chair X-ENPC-SU-Société Générale "Financial Risks" of the Risk fondation : A. Alfonsi, B. Jourdain, B. Lapeyre

9. Partnerships and Cooperations

9.1. National Initiatives

- ANR Cosmos 2015-2018, Participant: B. Jourdain ; Partners : Ecole des Ponts, Telecom, INIRIA Rennes and IBPC
- Labex Bezout
  http://bezout.univ-paris-est.fr

9.1.1. Competitivity Clusters

Pôle Finance Innovation

9.2. International Initiatives

9.2.1. Inria International Partners

9.2.1.1. Informal International Partners

- Center of Excellence program in Mathematics and Life Sciences at the Department of Mathematics, University of Oslo, Norway, (B. Øksendal).
- Cornell University, ORIE department (Andreea Minca)
- Roma Tor Vergata University (Lucia Caramellino)
- Ritsumeikan University (A. Kohatsu-Higa).

9.3. International Research Visitors

9.3.1. Visits of International Scientists

- Oleg Kudryavtsev (Rostov University, Russia)
- B. Stemper (Weierstrass Institute, Berlin)
- A. Kohatsu Higa (Ritsumeikan University)
- Justin Kirkby (Georgia Institute of Technology, Atlanta)
- Xiao Wei (Beijing University)
- Anton Arnold (TU Vienna)

9.3.1.1. Internships

- Baba Abdel Hamid, Inria
- Asma Sassi, Inria

9.3.2. Visits to International Teams

9.3.2.1. Research Stays Abroad

In the period 15.05 - 15. 06. 2019 Vlad Bally was an invited professor at the University Tor Vergata, Roma. Here he gave a course of 20h entitled "Integration by Parts and Convergence in Total Variation".

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. Member of the Organizing Committees

- A. Alfonsi

  Co-organizer of the working group seminar of MathRisk "Méthodes stochastiques et finance". http://cermics.enpc.fr/~alfonsi/GTMSF.html
- V. Bally

  Organizer of the seminar of the LAMA laboratory of Université Paris-Est.
- A. Sulem

  Co-organizer of the seminar Inria-MathRisk /Université Paris Diderot LPSM "Numerical probability and mathematical finance". https://www.lpsm.paris/mathfipronum/gt

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

- B. Jourdain

  Associate editor of:

  - ESAIM : Proceedings and Surveys
  - Stochastic Processes and their Applications (SPA)
- D. Lamberton

  Associate editor of:

  - Mathematical Finance,
  - Associate editor of ESAIM Probability & Statistics
- A. Sulem

  Associate editor of:

  - Journal of Mathematical Analysis and Applications (JMAA)
  - International Journal of Stochastic Analysis (IJSA)
  - SIAM Journal on Financial Mathematics (SIFIN)

10.1.2.2. Reviewer - Reviewing Activities

- B. Jourdain: Reviewer for *Mathematical Reviews*
- A. Sulem: Reviewer for *Mathematical Reviews*

10.1.3. Invited Talks

- Aurélien Alfonsi
  - 4th of April, 2019: "Lifted and Geometric Differentiability of the Squared Quadratic Wasserstein Distance", Séminaire MAD-Stat, Toulouse School of Economics.
  - 16th of May, 2019: "Approximation of OT problems with marginal moments contraints", Séminaire de probabilités et mathématiques financières, Evry.
  - 20th of June, 2019: "A generic construction for high order approximation schemes of semigroups using random grids", Mathematical and Computational Finance Seminar, Oxford.

- 27th of September, 2019: "A generic construction for high order approximation schemes of semigroups using random grids", Séminaire Bachelier.
- 24th of October, 2019: "A generic construction for high order approximation schemes of semigroups using random grids", GT Finance mathématique, probabilités numériques et statistique des processus, LPSM.

- Vlad Bally
  - The 22nd Conference of the Romanian Society of Probability and Statistics, Bucharest, May 10-11, Talk "Convergence in the CLT in distribution norms"
  - Analyse stochastique et thèmes connexes. 6-9 mai 2019, Bucarest, Roumanie. Talk "Regularity and estimates for the function solution of the 2D Bolzmann equation"
  - Perturbation Techniques in Stochastic Analysis and its Applications. Luminy 11-15 March. Talk "Regularity and estimates for the function solution of the 2D Bolzmann equation"
  - Conference Unirandom on random nodal sets, Rennes 9-13 September.
  - Conference on asymptotic expansions and Malliavin calculus. Talk:"Regularization lemmas and convergence in total variation".
- Benjamin Jourdain
  - Vienna Seminar in Mathematical Finance and Probability, 10 October 2019 : Differentiability of the squared quadratic Wasserstein distance
  - Inria-CWI workshop, Amsterdam, 17-18 September 2019 : Sampling of probability measures in the convex order and computation of robust option price bounds
  - MCMC 2019, Sydney, 8-12 July 2019 : Weak error analysis for some mean-field SDEs
  - Nine talks on contemporary optimal transport problems, Strasbourg, 4-5 July 2019: The inverse transform martingale coupling
  - 9th general AMAMEF conference, Paris, 11-14 June 2019 : The inverse transform martingale coupling
  - Nancy probability and statistics seminar, 9 May 2019 : A new family of martingale couplings in dimension one
  - Conference Perturbation Techniques in Stochastic Analaysis and its Applications, Marseille 11-15 March 2019 : 3h lectures entitled Stochastic Differential Equations with mean-field rank-based coefficients
- Agnès Sulem
  - Conference on Stochastic Analysis and Applications Risør, Norway, August 26-30 2019. *Nonlinear pricing in incomplete markets with default.*
  - Complex Networks 2919, December 10-12, 2019, Lisbon. (Conference with referee process), *A Dynamic Contagion Risk Model With Recovery Features.*

10.1.4. Research Administration

- A. Alfonsi
  - Deputy director of CERMICS laboratory until August,
  - Director of the CERMICS since September.
  - In charge of the Master “Finance and Data” at the Ecole des Ponts.
- V. Bally
  - Member of the LAMA committee
  - Responsible of the Master 2, option finance.
- Benjamin Jourdain
  - Head of the doctoral school MSTIC, University Paris-Est until February 2019

- D. Lamberton
  - Vice-president for research at Université Paris-Est Marne-la-Vallée
- A. Sulem
  - Member of the Scientific Committee of AMIES
  - Corresponding member of the Operational Committee for the assesment of Legal and Ethical risks (COERLE) at Inria Paris research center
  - Member of the Committee for Inria international Chairs

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence :

- A. Alfonsi: "Probabilités”, first year course at Ecole des Ponts.
- V. Bally : "Analyse Hilbertienne", Course L3, UPEMLV

Master :

- Aurélien Alfonsi

- “Données Haute Fréquence en finance”, lecture for the Master at UPEMLV.
- “Mesures de risque”, Master course of UPEMLV and Sorbonne Université.
- Professeur chargé de cours at Ecole Polytechnique.

- Vlad Bally

- "Taux d’Intêret", M2 Finance.
- "Calcul de Malliavin et applications en finance", M2 Finance
- "Analyse du risque" M2 Actuariat,
- "Processus Stochastiques" M2 Recherche

- Benjamin Jourdain

- course "Mathematical finance", 2nd year ENPC
- course "Monte-Carlo Markov chain methods and particle algorithms", Research Master Probabilités et Modèles Aléatoires, Sorbonne Université

- B. Jourdain, B. Lapeyre

- course "Monte-Carlo methods", 3rd year ENPC and Research Master Mathématiques et Application, university of Marne-la-Vallée

- J.-F. Delmas, B.Jourdain

- course "Jump processes with applications to energy markets", 3rd year ENPC and Research Master Mathématiques et Application, University of Marne-la-Vallée

- D. Lamberton

- "Calcul stochastique pour la finance", master 1 course, Université Paris-Est Marne-la-Vallée

- B. Lapeyre

- Monte-Carlo methods in quantitative finance, Master of Mathematics, University of Luxembourg,

- A. Sulem

- "PDE methods in Finance", Master of Mathematics, University of Luxembourg, 22 h lectures and responsible of the module "Numerical Methods in Finance".

10.2.2. Supervision

- PhD: Rui Chen, "Dynamic Optimal Control for Distress in Large Financial Networks and Mean field Systems with Jumps", Université Paris-Dauphine, defended on July 19th 2019, Supervisor: Agnès Sulem [10].
- PhD in progress :
  - Anas Bentaleb (started February 2018) : Mathematical techniques for expected exposure evaluation, Supervisor: B. Lapeyre.
  - Adel Cherchali, “Numerical methods for the ALM”, funded by Fondation AXA, started in September 2017, Supervisor: Aurélien Alfonsi
  - Rafaël Coyaud, “Deterministic ans stochastic numerical methods for multimarginal and martingale constraint optimal transport problems”, started in October 2017, Supervisor: Aurélien Alfonsi
  - Oumaima Bencheikh (started November 2017) "Acceleration of probabilistic particle methods", supervised by B.Jourdain
  - Ezechiel Kahn (started September 2018) "Functional inequalities for random matrices models", supervised by B. Jourdain and D. Chafai
  - Sophian Mehalla (started November 2017), CIFRE agreement Milliman company/Ecole des Ponts (http://fr.milliman.com, Supervisor: B. Lapeyre
  - William Margheriti (started January 2018) "Numerical methods for martingale optimal transport problems", supervised by J.-F. Delmas and B. Jourdain

10.2.3. Juries

- Aurélien Alfonsi:
  - Report and jury of the PhD thesis of Rui Chen, July 19, Paris Dauphine.
  - Report and jury of the PhD thesis of Babacar Diallo, December 9, Evry University.
- Benjamin Jourdain:
  - Report and jury of the PhD of Yating Liu, defended on December 3, Sorbonne University
  - PhD of Victor Marx, defended on October 25, University Nice Côte d’Azur
  - PhD of Nicolas Thomas, defended on June 20, Sorbonne University
- Agnès Sulem:
  - Report and jury of the HdR thesis *"Gestion optimale de Portefeuille: Du contrôle Sensible au Risque, à la Finance Comportementale et à la Science des Données* Sebastien Lleo, CNAM, Paris, 4 February 2019
  - PhD of Cyril Benezet, defended on 5 November 2019, LPSM, Université Paris-Diderot, (Chair of the Committee)
  - Member of the Committee for the recruitment of a Professor in applied mathematics, finance and numerical probability, Laboratoire de probabilités (LPSM), Université Paris-Diderot, Spring 2019.

11. Bibliography

Major publications by the team in recent years

[1] A. Al Gerbi, B. Jourdain, E. Clément.*Ninomiya-Victoir scheme: strong convergence, antithetic version and application to multilevel estimators*, in "Monte Carlo Method and Applications", July 2016, vol. 22, n^{o} 3, p. 197-228, https://arxiv.org/abs/1508.06492 , https://hal-enpc.archives-ouvertes.fr/hal-01188675

[2] A. ALFONSI.*Affine diffusions and related processes: simulation, theory and applications*, Bocconi and Springer Series, Mathematics statistics, finance and economics, Springer, 2015

[3] A. ALFONSI, P. BLANC.*Dynamic optimal execution in a mixed-market-impact Hawkes price model*, in "Finance and Stochastics", January 2016, https://arxiv.org/abs/1404.0648 [*DOI :* 10.1007/s00780-015-0282-Y], https://hal-enpc.archives-ouvertes.fr/hal-00971369

[4] A. ALFONSI, B. JOURDAIN, A. KOHATSU-HIGA.*Optimal transport bounds between the time-marginals of a multidimensional diffusion and its Euler scheme*, in "Electronic Journal of Probability", 2015, https://arxiv.org/abs/1405.7007 , https://hal-enpc.archives-ouvertes.fr/hal-00997301

[5] H. AMINI, A. MINCA, A. SULEM.*Control of interbank contagion under partial information*, in "SIAM Journal on Financial Mathematics", December 2015, vol. 6, n^o 1, 24, https://hal.inria.fr/hal-01027540

[6] V. BALLY, L. CARAMELLINO.*Convergence and regularity of probability laws by using an interpolation method*, in "Annals of Probability", 2017, vol. 45, n^o 2, p. 1110–1159, https://hal-upec-upem.archives-ouvertes.fr/hal-01109276

[7] A. BOUSELMI, D. LAMBERTON.*The critical price of the American put near maturity in the jump diffusion model*, in "SIAM Journal on Financial Mathematics", May 2016, vol. 7, n^o 1, p. 236–272, https://arxiv.org/abs/1406.6615 [*DOI :* 10.1137/140965910], https://hal-upec-upem.archives-ouvertes.fr/hal-00979936

[8] R. DUMITRESCU, M.-C. QUENEZ, A. SULEM.*A Weak Dynamic Programming Principle for Combined Optimal Stopping/Stochastic Control with E^f-Expectations*, in "SIAM Journal on Control and Optimization", 2016, vol. 54, n^o 4, p. 2090-2115 [*DOI :* 10.1137/15M1027012], https://hal.inria.fr/hal-01370425

[9] R. DUMITRESCU, M.-C. QUENEZ, A. SULEM.*Game Options in an Imperfect Market with Default*, in "SIAM Journal on Financial Mathematics", January 2017, vol. 8, n^o 1, p. 532 - 559 [*DOI :* 10.1137/16M1109102], https://hal.inria.fr/hal-01614758

Publications of the year

Doctoral Dissertations and Habilitation Theses

[10] R. CHEN.*Dynamic optimal control for distress large financial networks and Mean field systems with jumps*, Université Paris-Dauphine, July 2019, https://hal.inria.fr/tel-02434108

Articles in International Peer-Reviewed Journal

[11] A. ALFONSI, J. CORBETTA, B. JOURDAIN.*Sampling of one-dimensional probability measures in the convex order and computation of robust option price bounds*, in "International Journal of Theoretical and Applied Finance", 2019, vol. 22, n^o 3, This paper is an updated version of a part of the paper https://hal.archives-ouvertes.fr/hal-01589581 (or https://arxiv.org/pdf/1709.05287.pdf) [*DOI :* 10.1142/S021902491950002X], https://hal-enpc.archives-ouvertes.fr/hal-01963507

[12] A. ALFONSI, J. CORBETTA, B. JOURDAIN.*Sampling of probability measures in the convex order by Wasserstein projection*, in "Annales de l'IHP - Probabilités et Statistiques", 2019, https://arxiv.org/abs/1709.05287 , forthcoming, https://hal.archives-ouvertes.fr/hal-01589581

[13] V. BALLY.*Upper bounds for the function solution of the homogenuous 2D Boltzmann equation with hard potential*, in "Annals of Applied Probability", 2019, vol. 29, p. 1929 - 1961, https://arxiv.org/abs/1710.00695 [*DOI :* 10.1214/18-AAP1451], https://hal-upec-upem.archives-ouvertes.fr/hal-02429468

[14] V. BALLY, L. CARAMELLINO.*Total variation distance between stochastic polynomials and invariance principles*, in "Annals of Probability", 2019, vol. 47, p. 3762 - 3811, https://arxiv.org/abs/1705.05194 [*DOI :* 10.1214/19-AOP1346], https://hal-upec-upem.archives-ouvertes.fr/hal-02429560

[15] V. BALLY, L. CARAMELLINO, G. POLY.*Non universality for the variance of the number of real roots of random trigonometric polynomials*, in "Probability Theory and Related Fields", 2019, vol. 174, n° 3-4, p. 887-927, https://arxiv.org/abs/1711.03316 [*DOI :* 10.1007/s00440-018-0869-2], https://hal.archives-ouvertes.fr/hal-01634848

[16] O. BENCHEIKH, B. JOURDAIN.*Bias behaviour and antithetic sampling in mean-field particle approximations of SDEs nonlinear in the sense of McKean*, in "ESAIM: Proceedings and Surveys", April 2019, vol. 65, p. 219-235, https://arxiv.org/abs/1809.06838 - 14 pages [*DOI :* 10.1051/PROC/201965219], https://hal.archives-ouvertes.fr/hal-01877002

[17] M. BRIANI, L. CARAMELLINO, G. TERENZI, A. ZANETTE.*Numerical stability of a hybrid method for pricing options*, in "International Journal of Theoretical and Applied Finance", September 2019, 1950036 [*DOI :* 10.1142/S0219024919500365], https://hal.archives-ouvertes.fr/hal-02380723

[18] L. GOUDENÈGE, A. MOLENT, A. ZANETTE.*Pricing and hedging GMWB in the Heston and in the Black–Scholes with stochastic interest rate models*, in "Computational Management Science", February 2019, vol. 16, n° 1-2, p. 217-248 [*DOI :* 10.1007/s10287-018-0304-2], https://hal.archives-ouvertes.fr/hal-01940715

[19] H. GÉRARD, M. DE LARA, J.-P. CHANCELIER.*Equivalence Between Time Consistency and Nested Formula*, in "Annals of Operations Research", May 2019, p. 1-21, https://arxiv.org/abs/1711.08633 [*DOI :* 10.1007/s10479-019-03276-1], https://hal-enpc.archives-ouvertes.fr/hal-01645564

[20] B. JOURDAIN, A. KEBAIER.*Non-asymptotic error bounds for The Multilevel Monte Carlo Euler method applied to SDEs with constant diffusion coefficient*, in "Electronic Journal of Probability", 2019, vol. 24, n° 12, p. 1-34, https://arxiv.org/abs/1708.07064 [*DOI :* 10.1214/19-EJP271], https://hal.archives-ouvertes.fr/hal-01577874

[21] B. JOURDAIN, A. ZHOU.*Existence of a calibrated Regime Switching Local Volatility model*, in "Mathematical Finance", 2019, https://arxiv.org/abs/1607.00077 , forthcoming [*DOI :* 10.1111/MAFI.12231], https://hal.archives-ouvertes.fr/hal-01341212

[22] D. LAMBERTON, G. TERENZI.*Variational formulation of American option prices in the Heston Model*, in "SIAM Journal on Financial Mathematics", April 2019, vol. 10, n° 1, p. 261-368, https://arxiv.org/abs/1711.11311 , https://hal-upec-upem.archives-ouvertes.fr/hal-01649496

Scientific Books (or Scientific Book chapters)

[23] B. ØKSENDAL, A. SULEM.*Applied Stochastic Control of Jump Diffusions*, Springer, Universitext, 2019, 436, 3rd edition [*DOI :* 10.1007/978-3-030-02781-0], https://hal.archives-ouvertes.fr/hal-02411121

Other Publications

[24] A. ALFONSI, V. BALLY.*A generic construction for high order approximation schemes of semigroups using random grids*, December 2019, https://arxiv.org/abs/1905.08548 - working paper or preprint, https://hal-enpc.archives-ouvertes.fr/hal-02406433

[25] A. ALFONSI, A. CHERCHALI, J. A. I. ACEVEDO.*A full and synthetic model for Asset-Liability Management in life insurance, and analysis of the SCR with the standard formula*, December 2019, https://arxiv.org/abs/1908.00811 - working paper or preprint, https://hal-enpc.archives-ouvertes.fr/hal-02406439

[26] A. ALFONSI, R. COYAUD, V. EHRLACHER, D. LOMBARDI.*Approximation of Optimal Transport problems with marginal moments constraints*, May 2019, https://arxiv.org/abs/1905.05663 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02128374

[27] H. AMINI, R. CHEN, A. MINCA, A. SULEM.*A dynamic contagion risk model with recovery features*, 2019, working paper or preprint [*DOI :* 10.2139/SSRN.3435257], https://hal.inria.fr/hal-02421342

[28] V. BALLY, L. CARAMELLINO.*Transfer of regularity for Markov semigroups*, January 2020, working paper or preprint, https://hal-upec-upem.archives-ouvertes.fr/hal-02429530

[29] V. BALLY, L. CARAMELLINO, G. POLY.*Regularization lemmas and convergence in total variation*, January 2020, working paper or preprint, https://hal-upec-upem.archives-ouvertes.fr/hal-02429512

[30] O. BENCHEIKH, B. JOURDAIN.*Weak and strong error analysis for mean-field rank based particle approximations of one dimensional viscous scalar conservation law*, October 2019, https://arxiv.org/abs/1910.11237 - 47 pages, https://hal.archives-ouvertes.fr/hal-02332760

[31] R. CHEN, R. DUMITRESCU, A. MINCA, A. SULEM.*Mean field BSDEs and global dynamic risk measures*, December 2019, working paper or preprint [*DOI :* 10.2139/SSRN.3446360], https://hal.inria.fr/hal-02421316

[32] M. GRIGOROVA, M.-C. QUENEZ, A. SULEM.*American options in a non-linear incomplete market model with default*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02025835

[33] M. GRIGOROVA, M.-C. QUENEZ, A. SULEM.*European options in a non-linear incomplete market model with default*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02025833

[34] B. JOURDAIN, T. LELIÈVRE, P.-A. ZITT.*Convergence of metadynamics: discussion of the adiabatic hypothesis*, April 2019, https://arxiv.org/abs/1904.08667 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02104961

[35] B. JOURDAIN, G. PAGÈS.*Convex order, quantization and monotone approximations of ARCH models*, October 2019, https://arxiv.org/abs/1910.00799 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02304190

[36] B. LAPEYRE, J. LELONG.*Neural network regression for Bermudan option pricing*, December 2019, https://arxiv.org/abs/1907.06474 - working paper or preprint, http://hal.univ-grenoble-alpes.fr/hal-02183587

References in notes

[37] A. AHDIDA, A. ALFONSI, E. PALIDDA.*Smile with the Gaussian term structure model*, in "Journal of Computational Finance", 2017, https://arxiv.org/abs/1412.7412 , https://hal.archives-ouvertes.fr/hal-01098554

[38] A. AL GERBI, B. JOURDAIN, E. CLÉMENT.*Asymptotics for the normalized error of the Ninomiya-Victoir scheme*, in "Stochastic Processes and their Applications", 2018, vol. 128, n^{o} 6, p. 1889-1928, https://hal-enpc.archives-ouvertes.fr/hal-01259915

[39] A. AL GERBI, B. JOURDAIN, E. CLÉMENT.*Ninomiya-Victoir scheme : Multilevel Monte-Carlo estimators and discretization of the involved Ordinary Differential Equations*, in "ESAIM: Proceedings and Surveys", November 2017, vol. 59, p. 1-14, https://arxiv.org/abs/1612.07017 , https://hal.archives-ouvertes.fr/hal-01421337

[40] A. ALFONSI, P. BLANC.*Extension and calibration of a Hawkes-based optimal execution model*, in "Market microstructure and liquidity", August 2016, https://arxiv.org/abs/1506.08740 [*DOI :* 10.1142/S2382626616500052], https://hal-enpc.archives-ouvertes.fr/hal-01169686

[41] A. ALFONSI, J. CORBETTA, B. JOURDAIN.*Evolution of the Wasserstein distance between the marginals of two Markov processes*, in "Bernoulli", 2018, vol. 24, n^{o} 4A, p. 2461-2498, https://hal.archives-ouvertes.fr/hal-01390887

[42] A. ALFONSI, M. HAYASHI, A. KOHATSU-HIGA.*Parametrix Methods for One-Dimensional Reflected SDEs*, in "Modern Problems of Stochastic Analysis and StatisticsSelected Contributions In Honor of Valentin Konakov", Springer, November 2017, vol. Springer Proceedings in Mathematics & Statistics, n^{o} 208, https://hal-enpc.archives-ouvertes.fr/hal-01670011

[43] A. ALFONSI, A. KEBAIER, C. REY.*Maximum Likelihood Estimation for Wishart processes*, in "Stochastic Processes and their Applications", November 2016, https://arxiv.org/abs/1508.03323 [*DOI :* 10.1016/J.SPA.2016.04.026], https://hal-enpc.archives-ouvertes.fr/hal-01184310

[44] A. ALFONSI, A. SCHIED, F. KLÖCK.*Multivariate transient price impact and matrix-valued positive definite functions*, in "Mathematics of Operations Research", March 2016, https://arxiv.org/abs/1310.4471 [*DOI :* 10.1287/MOOR.2015.0761], https://hal-enpc.archives-ouvertes.fr/hal-00919895

[45] H. AMINI, R. CONT, A. MINCA.*Resilience to Contagion in Financial Networks*, in "Mathematical Finance", 2013

[46] H. AMINI, A. MINCA, A. SULEM.*Optimal equity infusions in interbank networks*, in "Journal of Financial Stability", August 2017, vol. 31, p. 1 - 17 [*DOI :* 10.1016/J.JFS.2017.05.008], https://hal.inria.fr/hal-01614759

[47] A. ARNOLD, E. CARLEN, Q. JU.*Large-time behavior of non-symmetric Fokker-Planck type equations*, in "Communications on Stochastic Analysis", 2008, vol. 2, n^{o} 1, p. 153-175

[48] R. ASSARAF, B. JOURDAIN, T. LELIÈVRE, R. ROUX.*Computation of sensitivities for the invariant measure of a parameter dependent diffusion*, in "Stochastics and Partial Differential Equations: Analysis and Computations", October 2017, p. 1-59, https://arxiv.org/abs/1509.01348 [*DOI :* 10.1007/S40072-017-0105-6], https://hal.archives-ouvertes.fr/hal-01192862

[49] V. BALLY, L. CARAMELLINO.*Asymptotic development for the CLT in total variation distance*, in "Bernoulli", 2016, vol. 22, p. 2442-2485, https://hal-upec-upem.archives-ouvertes.fr/hal-01104866

[50] V. BALLY, L. CARAMELLINO.*Regularity of Wiener functionals under a Hörmander type condition of order one*, in "Annals of Probability", 2017, vol. 45, n^o 3, p. 1488-1511, https://hal-upec-upem.archives-ouvertes.fr/hal-01413556

[51] V. BALLY, L. CARAMELLINO, P. PIGATO.*Diffusions under a local strong Hörmander condition. Part I: density estimates*, December 2016, preprint, https://hal-upec-upem.archives-ouvertes.fr/hal-01413546

[52] V. BALLY, L. CARAMELLINO, P. PIGATO.*Diffusions under a local strong Hörmander condition. Part II: tube estimates*, July 2016, preprint, https://hal.archives-ouvertes.fr/hal-01407420

[53] V. BALLY, L. CARAMELLINO, G. POLY.*Convergence in distribution norms in the CLT for non identical distributed random variables*, January 2017, preprint, https://hal-upec-upem.archives-ouvertes.fr/hal-01413548

[54] V. BALLY, A. KOHATSU-HIGA.*A probabilistic interpretation of the parametrix method*, in "Annals of Applied Probability", 2015, vol. 25, p. 3095-3138, https://hal.archives-ouvertes.fr/hal-00926479

[55] V. BALLY, V. RABIET.*Asymptotic behavior for multi-scale PDMP's*, April 2015, preprint, https://hal.archives-ouvertes.fr/hal-01144107

[56] V. BALLY, C. REY.*Approximation of Markov semigroups in total variation distance*, in "Electronic Journal of Probability", 2016, vol. 21, n^o 12, https://hal-upec-upem.archives-ouvertes.fr/hal-01110015

[57] M. BEIGLBÖCK, P.-H. LABORDÈRE, F. PENKNER.*Model-independent bounds for option prices - a mass transport approach*, in "Finance Stoch.", 2013, vol. 17, n^o 3, p. 477-501

[58] R. CHEN, A. MINCA, A. SULEM.*Optimal connectivity for a large financial network*, in "ESAIM: Proceedings and Surveys", 2017, vol. 59, p. 43 - 55, Editors : B. Bouchard, E. Gobet and B. Jourdain, https://hal.inria.fr/hal-01618701

[59] Q. DAI, K. J. SINGLETON.*Specification Analysis of Affine Term Structure Models*, in "The Journal of Finance", 2000, vol. 55, n^o 5, p. 1943–1978, http://dx.doi.org/10.1111/0022-1082.00278

[60] D. DUFFIE, D. FILIPOVIĆ, W. SCHACHERMAYER.*Affine processes and applications in finance*, in "Ann. Appl. Probab.", 2003, vol. 13, n^o 3, p. 984–1053, http://dx.doi.org/10.1214/aoap/1060202833

[61] R. DUMITRESCU, M. GRIGOROVA, M.-C. QUENEZ, A. SULEM.*BSDEs with default jump*, in "Computation and Combinatorics in Dynamics, Stochastics and Control - The Abel Symposium, Rosendal, Norway August 2016", E. CELLEDONI, G. D. NUNNO, K. EBRAHIMI-FARD, H. MUNTHE-KAAS (editors), The Abel Symposia book series, Springer, 2018, vol. 13 [*DOI :* 10.1007/978-3-030-01593-0], https://hal.inria.fr/hal-01799335

[62] R. DUMITRESCU, C. LABART.*Numerical approximation of doubly reflected BSDEs with jumps and RCLL obstacles*, in "Journal of Mathematical Analysis and applications", October 2016, vol. 442, n^o 1, p. 206-243, https://arxiv.org/abs/1406.3612 , https://hal.archives-ouvertes.fr/hal-01006131

[63] R. DUMITRESCU, C. LABART.*Reflected scheme for doubly reflected BSDEs with jumps and RCLL obstacles*, in "Journal of Computational and Applied Mathematics", April 2016, vol. 296, p. 827-839, https://arxiv.org/abs/1502.02888 , https://hal.archives-ouvertes.fr/hal-01114996

[64] R. DUMITRESCU, M.-C. QUENEZ, A. SULEM.*Optimal Stopping for Dynamic Risk Measures with Jumps and Obstacle Problems*, in "Journal of Optimization Theory and Applications", 2015, vol. 167, n^{o} 1, 23 [*DOI :* 10.1007/s10957-014-0635-2], https://hal.inria.fr/hal-01096501

[65] R. DUMITRESCU, M.-C. QUENEZ, A. SULEM.*Generalized Dynkin games and doubly reflected BSDEs with jumps*, in "Electronic Journal of Probability", 2016 [*DOI :* 10.1214/16-EJP4568], https://hal.inria.fr/hal-01388022

[66] R. DUMITRESCU, M.-C. QUENEZ, A. SULEM. *Mixed generalized Dynkin game and stochastic control in a Markovian framework*, in "Stochastics", 2016, vol. 89, n^{o} 1, 30, https://hal.inria.fr/hal-01417203

[67] R. DUMITRESCU, M.-C. QUENEZ, A. SULEM.*American Options in an Imperfect Complete Market with Default*, in "ESAIM: Proceedings and Surveys", 2018, p. 93–110 [*DOI :* 10.1051/PROC/201864093], https://hal.inria.fr/hal-01614741

[68] N. EL KAROUI, V. LACOSTE.*Multifactor models of the term structure of interest rates*, 1992, Preprint University of Paris 6

[69] A. FIGALLI.*Existence and uniqueness for martingale solutions of SDEs with rough or degenerate coefficients*, in "Journal of Functional Analysis", 2008, vol. 254, p. 109–153

[70] C. FONTANA, B. ØKSENDAL, A. SULEM.*Market viability and martingale measures under partial information*, in "Methodology and Computing in Applied Probability", 2015, vol. 17, 24 [*DOI :* 10.1007/s11009-014-9397-4], https://hal.inria.fr/hal-00789517

[71] J. FONTBONA, B. JOURDAIN.*A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations*, in "Annals of Probability", February 2016, vol. 44, n^{o} 1, p. 131-170, https://arxiv.org/abs/1107.3300 , https://hal.archives-ouvertes.fr/hal-00608977

[72] N. FOURNIER, B. JOURDAIN.*Stochastic particle approximation of the Keller-Segel equation and two-dimensional generalization of Bessel processes*, in "The Annals of Applied Probability : an official journal of the institute of mathematical statistics", November 2017, vol. 27, n^{o} 5, p. 2807-2861, https://arxiv.org/abs/1507.01087 , https://hal-enpc.archives-ouvertes.fr/hal-01171481

[73] J. GARNIER, G. PAPANICOLAOU, T. YANG.*Large deviations for a mean field model of systemic risk*, in "SIAM Journal on Financial Mathematics", 2013, vol. 41, n^{o} 1, p. 151–184

[74] J. GUYON, P. HENRY-LABORDÈRE.*Being particular about calibration*, in "Risk", January 2012

[75] Y. HU, B. ØKSENDAL, A. SULEM.*Singular mean-field control games*, in "Stochastic Analysis and Applications", June 2017, vol. 35, n^{o} 5, p. 823 - 851 [*DOI :* 10.1080/07362994.2017.1325745], https://hal.inria.fr/hal-01614747

[76] B. JOURDAIN, J. REYGNER.*The small noise limit of order-based diffusion processes*, in "Electronic Journal of Probability", March 2014, vol. 19, n° 29, p. 1-36, https://arxiv.org/abs/1307.0490 [*DOI :* 10.1214/EJP.v19-2906], https://hal-enpc.archives-ouvertes.fr/hal-00840185

[77] B. JOURDAIN, J. REYGNER.*Capital distribution and portfolio performance in the mean-field Atlas model*, in "Annals of Finance", May 2015, vol. 11, n° 2, p. 151-198, https://arxiv.org/abs/1312.5660 [*DOI :* 10.1007/s10436-014-0258-5], https://hal-enpc.archives-ouvertes.fr/hal-00921151

[78] B. JOURDAIN, J. REYGNER.*Optimal convergence rate of the multitype sticky particle approximation of one-dimensional diagonal hyperbolic systems with monotonic initial data*, in "Discrete and Continuous Dynamical Systems - Series A", September 2016, vol. 36, n° 9, p. 4963-4996, https://arxiv.org/abs/1507.01085 [*DOI :* 10.3934/DCDS.2016015], https://hal-enpc.archives-ouvertes.fr/hal-01171261

[79] D. LÉPINGLE.*Euler scheme for reflected stochastic differential equations*, in "Math. Comput. Simulation", 1995, vol. 38, n° 1-3, p. 119–126, Probabilités numériques (Paris, 1992), http://dx.doi.org/10.1016/0378-4754(93)E0074-F

[80] A. MINCA.*Modélisation mathématique de la contagion de défaut; Mathematical modeling of financial contagion*, Université Pierre et Marie Curie, Paris 6, September 5 2011

[81] A. MINCA, A. SULEM.*Optimal Control of Interbank Contagion Under Complete Information*, in "Statistics and Risk Modeling", 2014, vol. 31, n° 1, p. 1001-1026 [*DOI :* 10.1524/STRM.2014.5005], https://hal.inria.fr/hal-00916695

[82] S. NINOMIYA, N. VICTOIR.*Weak approximation of stochastic differential equations and application to derivative pricing*, in "Applied Mathematical Finance", 2008, vol. 15, p. 107-121

[83] M.-C. QUENEZ, A. SULEM.*Reflected BSDEs and robust optimal stopping for dynamic risk measures with jumps*, in "Stochastic Processes and their Applications", September 2014, vol. 124, n° 9, 23, https://arxiv.org/abs/1212.6744 , https://hal.inria.fr/hal-00773708

[84] B. ØKSENDAL, A. SULEM.*Forward–Backward Stochastic Differential Games and Stochastic Control under Model Uncertainty*, in "Journal of Optimization Theory and Applications", April 2014, vol. 161, n° 1, p. 22 - 55 [*DOI :* 10.1007/s10957-012-0166-7], https://hal.inria.fr/hal-01681150

[85] B. ØKSENDAL, A. SULEM.*Dynamic Robust Duality in Utility Maximization*, in "Applied Mathematics and Optimization", 2016, p. 1-31, https://hal.inria.fr/hal-01406663

[86] B. ØKSENDAL, A. SULEM.*Optimal control of predictive mean-field equations and applications to finance*, in "Springer Proceedings in Mathematics & Statistics", Stochastic of Environmental and Financial Economics, Springer Verlag, 2016, vol. 138, p. 301–320 [*DOI :* 10.1007/978-3-319-23425-0], https://hal.inria.fr/hal-01406649

Project-Team MIMOVE

Middleware on the Move

RESEARCH CENTER
Paris

THEME
Distributed Systems and middleware

Table of contents

Project-Team MIMOVE

Creation of the Team: 2014 July 01, updated into Project-Team: 2018 February 01

Keywords:

Computer Science and Digital Science:

A1.2.1. - Dynamic reconfiguration
A1.2.3. - Routing
A1.2.4. - QoS, performance evaluation
A1.2.5. - Internet of things
A1.2.6. - Sensor networks
A1.2.7. - Cyber-physical systems
A1.3. - Distributed Systems
A1.4. - Ubiquitous Systems
A1.5. - Complex systems
A1.5.1. - Systems of systems
A1.5.2. - Communicating systems
A2.5. - Software engineering
A2.6.2. - Middleware
A3.1.7. - Open data
A3.1.8. - Big data (production, storage, transfer)
A3.3. - Data and knowledge analysis
A3.5. - Social networks

Other Research Topics and Application Domains:

B6.3. - Network functions
B6.4. - Internet of things
B6.5. - Information systems
B8.2. - Connected city
B8.5.1. - Participative democracy

1. Team, Visitors, External Collaborators

Research Scientists

Nikolaos Georgantas [Team leader, Inria, Researcher, HDR]
Vassilis Christophides [Inria, Advanced Research Position, until Jun 2019]
Renata Cruz Teixeira [Inria, Senior Researcher, HDR]
Valérie Issarny [Inria, Senior Researcher, HDR]

External Collaborators

Shohreh Ahvar [ISEP, until Jun 2019]
Rafael Angarita Arocha [ISEP]
Sara Ayoubi [Nokia Bell Labs, from May 2019]
Zied Ben Houidi [Nokia Bell Labs, until Apr 2019]
Francesco Bronzino [Nokia Bell Labs]
Vassilis Christophides [Univ de Cergy Pontoise, from Jul 2019]
Bruno Lefevre [Univ Paris Nord]

Françoise Sailhan [CNAM]
Rachit Agarwal [IIT Kanpur]
Georgios Bouloukakis [Univ of California, Irvine, from May 2019]

Technical Staff

Ehsan Ahvar [Inria, Engineer, until Jun 2019]
Pierre-Guillaume Raverdy [Inria, Engineer, from Mar 2019]

PhD Students

William Aboucaya [Sorbonne Université, PhD Student, from Oct 2019]
Fethi Dilmi [Inria, PhD Student, until Nov 2019]
Yifan Du [Inria, PhD Student]
Israel Marquez Salinas [Univ Pierre et Marie Curie, PhD Student, until Mar 2019]
Patient Ntumba [Inria, PhD Student]
Grigorios Piperagkas [Inria, PhD Student]
Sarah Wassermann [Inria, PhD Student, until Apr 2019]

Post-Doctoral Fellows

Sara Ayoubi [Inria, Post-Doctoral Fellow, until Apr 2019]
Georgios Bouloukakis [Inria, Post-Doctoral Fellow, until Apr 2019]
Giulio Grassi [Inria, Post-Doctoral Fellow]

Visiting Scientist

Mark Crovella [Boston University, until Jul 2019]

Administrative Assistants

Nathalie Gaudechoux [Inria, Administrative Assistant]
Meriem Guemair [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

Given the prevalence of global networking and computing infrastructures (such as the Internet and the Cloud), mobile networking environments, powerful hand-held user devices, and physical-world sensing and actuation devices, the possibilities of new mobile distributed systems have reached unprecedented levels. Such systems are dynamically composed of networked resources in the environment, which may span from the immediate neighborhood of the users – as advocated by pervasive computing – up to the entire globe – as envisioned by the Future Internet and one of its major constituents, the Internet of Things. Hence, we can now talk about truly ubiquitous computing.

The resulting ubiquitous systems have a number of unique – individually or in their combination – features, such as dynamicity due to volatile resources and user mobility, heterogeneity due to constituent resources developed and run independently, and context-dependence due to the highly changing characteristics of the execution environment, whether technical, physical or social. The latter two aspects are particularly manifested through the physical but also social sensing and actuation capabilities of mobile devices and their users. More specifically, leveraging the massive adoption of smart phones and other user-controlled mobile devices, besides physical sensing – where a device's sensor passively reports the sensed phenomena – *social sensing/crowd sensing* comes into play, where the user is aware of and indeed aids in the sensing of the environment.

Mobile systems with the above specifics further push certain problems related to the Internet and user experience to their extreme: (i) Technology is too complex. Most Internet users are not tech-savvy and hence cannot fix performance problems and anomalous network behavior by themselves. The complexity of most Internet applications makes it hard even for networking experts to fully diagnose and fix problems. Users can't even know whether they are getting the Internet performance that they are paying their providers for. (ii) There is too much content. The proliferation of user-generated content (produced anywhere with mobile devices and immediately published in social media) along with the vast amount of information produced by traditional media (e.g., newspapers, television, radio) poses new challenges in achieving an effective, near real-time information awareness and personalization. For instance, users need novel filtering and recommendation tools for helping them to decide which articles to read or which movie to watch.

This challenging context raises key research questions:

- How to deal with heterogeneity and dynamicity, which create runtime uncertainty, when developing and running mobile systems in the open and constantly evolving Internet and IoT environment?
- How to enable automated diagnosis and optimization of networks and systems in the Internet and IoT environment for improving the QoE of their users?
- How to raise human centric crowd-sensing to a reliable means of sensing world phenomena?
- How to deal with combination, analysis and privacy aspects of Web/social media and IoT crowd-sensing data streams?

3. Research Program

3.1. Introduction

The research questions identified above call for radically new ways in conceiving, developing and running mobile distributed systems. In response to this challenge, MiMove's research aims at enabling next-generation mobile distributed systems that are the focus of the following research topics.

3.2. Emergent mobile distributed systems

Uncertainty in the execution environment calls for designing mobile distributed systems that are able to run in a beforehand unknown, ever-changing context. Nevertheless, the complexity of such change cannot be tackled at system design-time. Emergent mobile distributed systems are systems which, due to their automated, dynamic, environment-dependent composition and execution, *emerge* in a possibly non-anticipated way and manifest *emergent properties*, i.e., both systems and their properties take their complete form only at runtime and may evolve afterwards. This contrasts with the typical software engineering process, where a system is finalized during its design phase. MiMove's research focuses on enabling the emergence of mobile distributed systems while assuring that their required properties are met. This objective builds upon pioneering research effort in the area of *emergent middleware* initiated by members of the team and collaborators [1], [3].

3.3. Large-scale mobile sensing and actuation

The extremely large scale and dynamicity expected in future mobile sensing and actuation systems lead to the clear need for algorithms and protocols for addressing the resulting challenges. More specifically, since connected devices will have the capability to sense physical phenomena, perform computations to arrive at decisions based on the sensed data, and drive actuation to change the environment, enabling proper coordination among them will be key to unlocking their true potential. Although similar challenges have been addressed in the domain of networked sensing, including by members of the team [7], the specific challenges arising from the *extremely large scale* of mobile devices – a great number of which will be attached to people, with uncontrolled mobility behavior – are expected to require a significant rethink in this domain. MiMove's research investigates techniques for efficient coordination of future mobile sensing and actuation systems with a special focus on their dependability.

3.4. Mobile social crowd-sensing

While mobile social sensing opens up the ability of sensing phenomena that may be costly or impossible to sense using embedded sensors (e.g., subjective crowdedness causing discomfort or joyfulness, as in a bus or in a concert) and leading to a feeling of being more socially involved for the citizens, there are unique consequent challenges. Specifically, MiMove's research focuses on the problems involved in the combination of the physically sensed data, which are quantitative and objective, with the mostly qualitative and subjective data arising from social sensing. Enabling the latter calls for introducing mechanisms for incentivising user participation and ensuring the privacy of user data, as well as running empirical studies for understanding the complex social behaviors involved. These objectives build upon previous research work by members of the team on mobile social ecosystems and privacy, as well as a number of efforts and collaborations in the domain of smart cities and transport that have resulted in novel mobile applications enabling empirical studies of social sensing systems.

3.5. Active and passive probing methods

We are developing methods that actively introduce probes in the network to discover properties of the connected devices and network segments. We are focusing in particular on methods to discover properties of home networks (connected devices and their types) and to distinguish if performance bottlenecks lie within the home network versus in the different network segments outside (e.g., Internet access provider, interconnects, or content provider). Our goal is to develop adaptive methods that can leverage the collaboration of the set of available devices (including end-user devices and the home router, depending on which devices are running the measurement software).

We are also developing passive methods that simply observe network traffic to infer the performance of networked applications and the location of performance bottlenecks, as well as to extract patterns of web content consumption. We are working on techniques to collect network traffic both at user's end-devices and at home routers. We also have access to network traffic traces collected on a campus network and on a large European broadband access provider.

3.6. Inferring user online experience

We are developing hybrid measurement methods that combine passive network measurement techniques to infer application performance with techniques from HCI to measure user perception as well as methods to directly measure application quality. We later use the resulting datasets to build models of user perception of network performance based only on data that we can obtain automatically from the user device or from user's traffic observed in the network.

3.7. Real time data analytics

The challenge of deriving insights from the Internet of Things (IoT) has been recognized as one of the most exciting and key opportunities for both academia and industry. The time value of data is crucial for many IoT-based systems requiring *real-time* (or near real-time) *control* and *automation*. Such systems typically collect data continuously produced by "things" (i.e., devices), and analyze them in (sub-) seconds in order to act promptly, e.g., for detecting security breaches of digital systems, for spotting malfunctions of physical assets, for recommending goods and services based on the proximity of potential clients, etc. Hence, they require to both *ingest* and *analyze in real-time* data arriving with different velocity from various IoT data streams.

Existing incremental (online or streaming) techniques for descriptive statistics (e.g., frequency distributions, frequent patterns, etc.) or predictive statistics (e.g., classification, regression) usually assume a good enough quality dataset for mining patterns or training models. However, IoT raw data produced in the wild by sensors embedded in the environment or wearable by users are prone to errors and noise. Effective and efficient algorithms are needed for *detecting* and *repairing data impurities* (for controlling data quality) as well as *understanding data dynamics* (for defining alerts) in real-time, for collections of IoT data streams that might be geographically distributed. Moreover, supervised deep learning and data analytics techniques are challenged by the presence of sparse ground truth data in real IoT applications. Lightweight and adaptive semi-supervised or unsupervised techniques are needed to power real-time anomaly and novelty detection in IoT data streams. The effectiveness of these techniques should be able to reach a useful level through training on a relatively small amount of (preferably unlabeled) data while they can cope distributional characteristics of data evolving over time.

4. Application Domains

4.1. Mobile urban systems for smarter cities

With the massive scale adoption of mobile devices and further expected significant growth in relation with the Internet of Things, mobile computing is impacting most – if not all – the ICT application domains. One such domain is the one of "*smart cities*". The smart city vision anticipates that the whole urban space, including buildings, power lines, gas lines, roadways, transport networks, and cell phones, can all be wired together and monitored. Detailed information about the functioning of the city then becomes available to both city dwellers and businesses, thus enabling better understanding and consequently management of the city's infrastructure and resources. This raises the prospect that cities will become more sustainable environments, ultimately enhancing the citizens' well being. There is the further promise of enabling radically new ways of living in, regulating, operating and managing cities, through the increasing active involvement of citizens by ways of crowd-sourcing/sensing and social networking.

Still, the vision of what smart cities should be about has been and keeps evolving at a fast pace in close concert with the latest technology trends. It is notably worth highlighting how mobile and social network use has reignited citizen engagement, thereby opening new perspectives for smart cities beyond data analytics that have been initially one of the core foci for smart cities technologies. Similarly, open data programs foster the engagement of citizens in the city operation and overall contribute to make our cities more sustainable. The unprecedented democratization of urban data fueled by open data channels, social networks and crowd sourcing enables not only the monitoring of the activities of the city but also the assessment of their nuisances based on their impact on the citizens, thereby prompting social and political actions. However, the comprehensive integration of urban data sources for the sake of sustainability remains largely unexplored. This is an application domain that we focus on, further leveraging our research on emergent mobile distributed systems, large-scale mobile sensing & actuation, and mobile social crowd-sensing.

In particular, we concentrate on the following specialized applications:

- **Democratization of urban data for healthy cities.** We integrate the various urban data sources, especially by way of crowd-Xing, to better understand city nuisances. This goes from raw pollution sensing (e.g., sensing noise) to the sensing of its impact on citizens (e.g., how people react to urban noise and how this affects their health).
- **Social applications.** Mobile applications are being considered by sociologists as a major vehicle to actively involve citizens and thereby prompt them to become activists. We study such a vehicle from the ICT perspective and in particular elicit relevant middleware solutions to ease the development of such "*civic apps*".

4.2. Home network diagnosis

With the availability of cheap broadband connectivity, Internet access from the home has become a ubiquity. Modern households host a multitude of networked devices, ranging from personal devices such as laptops and smartphones to printers and media centers. These devices connect among themselves and to the Internet via a local-area network–a home network–that has become an important part of the "Internet experience". In fact, ample anecdotal evidence suggests that the home network can cause a wide array of connectivity impediments, but their nature, prevalence, and significance remain largely unstudied.

Our long-term goal is to assist users with concrete indicators of the quality of their Internet access, causes of potential problems and–ideally–ways to fix them. We intend to develop a set of easy-to-use home network monitoring and diagnosis tools. The development of home network monitoring and diagnosis tools brings a number of challenges. First, home networks are heterogeneous. The set of devices, configurations, and applications in home networks vary significantly from one home to another. We must develop sophisticated techniques that can learn and adapt to any home network as well as to the level of expertise of the user. Second, Internet application and services are also heterogeneous with very diverse network requirements. We must develop methods that can infer application quality solely from the observation of (often encrypted) application network traffic. There are numerous ways in which applications can fail or experience poor performance in home networks. Often there are a number of explanations for a given symptom. We must devise techniques that can identify the most likely cause(s) for a given problem from a set of possible causes. Finally, even if we can identify the cause of the problem, we must then be able to identify a solution. It is important that the output of the diagnosis tools we build is “actionable”. Users should understand the output and know what to do.

In our patternship with Princeton University (associate team HOMENET) we have deployed monitoring infrastructure within users’ homes. We are developing a mostly passive measurement system to monitor the performance of user applications, which we call Network Microscope. We are developing Network Microscope to run in a box acting as home gateway. We have deployed these boxes in 50 homes in the US and 10 in France. The US deployment was ran and financed by the Wall Street Journal. They were interested in understanding the relationship between Internet access speed and video quality. We have been discussing with Internet regulators (in particular, FCC, ACERP, and BEREC) as well as residential access ISP in how Network Microscope can help overcome the shortcomings of existing Internet quality monitoring systems.

4.3. Mobile Internet quality of experience

Mobile Internet usage has boomed with the advent of ever smarter handheld devices and the spread of fast wireless access. People rely on mobile Internet for everyday tasks such as banking, shopping, or entertainment. The importance of mobile Internet in our lives raises people’s expectations. Ensuring good Internet user experience (or Quality of Experience—QoE) is challenging, due to the heavily distributed nature of Internet services. For mobile applications, this goal is even more challenging as access connectivity is less predictable due to user mobility, and the form factor of mobile devices limits the presentation of content. For these reasons, the ability to monitor QoE metrics of mobile applications is essential to determine when the perceived application quality degrades and what causes this degradation in the chain of delivery. Our goal is to improve QoE of mobile applications.

To achieve this goal, we are working on three main scientific objectives. First, we are working on novel methods to monitor mobile QoE. Within the IPL BetterNet we are developing the HostView for Android tool that runs directly on mobile devices to monitor network and system performance together with the user perception of performance. Second, we plan to develop models to predict QoE of mobile applications. We will leverage the datasets collected with HostView for Android to build data-driven models. Finally, our goal is to develop methods to optimize QoE for mobile users. We are currently developing optimization methods for interactive video applications. We envision users walking or driving by road-side WiFi access points (APs) with full 3G/LTE coverage and patchy WiFi coverage (i.e., community Wifi or Wifi APs on Lampposts) or devices with multiple 3G/LTE links. To achieve this goal, we plan to leverage multi-path and cross-layer optimizations.

4.4. Internet Scanning

Internet-wide scanning has enabled researchers to answer a wealth of new security and measurement questions ranging from “How are authoritarian regimes spying on journalists?” to “Are security notifications effective at prompting operators to patch?” Most of these studies have used tools like ZMap, which operates naiıvely, scanning every IPv4 address once. This simplicity enables researchers to easily answer a question once, but the methodology scales poorly when continually scanning to detect changes, as networks change at dramatically different rates. Service configurations change more frequently on cloud providers like Amazon and Azure than on residential networks. Internet providers in developing regions often have extremely short DHCP windows. Some networks are unstable with host presence varying wildly between different hours and others have distinct periodic patterns, e.g., hosts are only available during regional business hours. A handful of large autonomous systems have not had hosts present in decades. Our work in collaboration with Stanford University is developing more intelligent Internet-wide scanning methods to then implement a system that can scan continuously. Such a system will allow for up-to-date analysis of Internet trends and threats with real-time alerts of important events.

5. Highlights of the Year

5.1. Highlights of the Year

The Wall Street Journal published an article, “The Truth About Faster Internet: It’s Not Worth It” (on the front page of the printed version of the paper on August 21, 2019) based on the results of the models we have developed to infer video quality from encrypted network traffic and the Network Microscope system.

6. New Software and Platforms

6.1. SocialBus

Universal Social Network Bus

KEYWORDS: Middleware - Interoperability - Social networks - Software Oriented Service (SOA)

FUNCTIONAL DESCRIPTION: Online social network services (OSNSs) have become an integral part of our daily lives. At the same time, the aggressive market competition has led to the emergence of multiple competing siloed OSNSs that cannot interoperate. As a consequence, people face the burden of creating and managing multiple OSNS accounts and learning how to use them, to stay connected. The goal of the Universal Social Network Bus (USNB) is to relieve users from such a burden, letting them use their favorite applications to communicate.

- Authors: Rafael Angarita Arocha, Nikolaos Georgantas and Valérie Issarny
- Contact: Valérie Issarny
- URL: https://gitlab.inria.fr/usnb/universal-social-network-bus

6.2. WeBrowse

KEYWORDS: Web Usage Mining - Content analysis - Recommendation systems

FUNCTIONAL DESCRIPTION: The amount of information available on the web today, and the fast rate with which new information appears, overwhelm most users. The goal of our research is to assist Web users in discovering content. One of the most powerful means today to help people discover new web content is sharing between members of online communities. In the case of communities of a place (e.g., people who live, study, or work together) people share common interests, but often fail to actively share content. To address this problem, we have developed WeBrowse, a passive crowdsourced content discovery system for communities of a place.

WeBrowse leverages the passive observation of web-clicks (i.e., the URLs users intentionally visit) as an indication of users’ interest in a piece of content. Intuitively, the more users click on a URL, the higher the interest in the content on the corresponding page. Our approach is then to leverage the collective clicks in a community to automatically discover relevant content to promote to users of the community.

To implement passive crowdsourcing, one must be in a position to observe the aggregated web-clicks of the community. Luckily, in many communities of a place, users will connect to the Internet from the same network, such as, e.g., the campus/enterprise network or the network of a residential Internet Service Provider (ISP) in a neighborhood. WeBrowse (i) observes web packets flowing through a network link, (ii) passively extracts HTTP logs (i.e., streams recording the headers of HTTP requests), and (iii) detects and decides on-the-fly the set of URLs to show to users.

- Contact: Renata Cruz Teixeira
- URL: https://team.inria.fr/muse/webrowse-info-page/

6.3. VSB

eVolution Service Bus

KEYWORDS: Service and Thing choreographies - Middleware protocol interoperability - Enterprise service bus

FUNCTIONAL DESCRIPTION: VSB is a development and runtime environment dedicated to complex distributed applications of the Future Internet. Such applications are open, dynamic choreographies of extremely heterogeneous services and Things, including lightweight embedded systems (e.g., sensors, actuators and networks of them), mobile systems (e.g., smartphone applications), and resource-rich IT systems (e.g., systems hosted on enterprise servers and Cloud infrastructures). VSB's objective is to seamlessly interconnect, inside choreographies, services and Things that employ heterogeneous interaction protocols at the middleware level, e.g., SOAP Web services, REST Web services, Things using CoAP. This is based on runtime conversions between such protocols, with respect to their primitives and data type systems, while properly mapping between their semantics. This also includes mapping between the public interfaces of services/Things, regarding their operations and data, from the viewpoint of the middleware: the latter means that operations and data are converted based on their middleware-level semantics, while their business semantics remains transparent to the conversion. VSB follows the well-known Enterprise Service Bus (ESB) paradigm. We propose a generic interface description, which we call GIDL, for application components that employ VSB. Based on GIDL, we enable automated synthesis of binding components for connecting heterogeneous services and Things onto VSB.

- Participants: Georgios Bouloukakis, Nikolaos Georgantas and Patient Ntumba
- Contact: Nikolaos Georgantas
- URL: https://gitlab.ow2.org/chorevolution/evolution-service-bus

6.4. Service traceroute

KEYWORDS: Network monitoring - Network diagnosis

FUNCTIONAL DESCRIPTION: Traceroute is often used to help diagnose when users experience issues with Internet applications or services. Unfortunately, probes issued by classic traceroute tools differ from application traffic and hence can be treated differently by middleboxes within the network. We propose a new traceroute tool, called Service traceroute. Service traceroute leverages the idea from paratrace, which passively listens to application traffic to then issue traceroute probes that pretend to be part of the application flow. We extend this idea to work for modern Internet services with support for identifying the flows to probe automatically, for tracing of multiple concurrent flows, and for UDP flows. We implement command-line and library versions of Service traceroute, which we release as open source.

- Partner: Princeton University
- Contact: Renata Cruz Teixeira
- URL: https://github.com/wontoniii/service-traceroute

6.5. Network Microscope

KEYWORDS: Quality of Experience - Network monitoring - Video analysis

FUNCTIONAL DESCRIPTION: A system that accurately infers video streaming quality metrics in real time, such as startup delay or video resolution, by using just a handful of features extracted from passive traffic measurement. Network Microscope passively collects a corpus of network features about the traffic flows of interest in the network and directs those to a real-time analytics framework that can perform more complex inference tasks. Network Microscope enables network operators to determine degradations in application quality as they happen, even when the traffic is encrypted.

- Participants: Francesco Bronzino and Renata Cruz Teixeira
- Contact: Renata Cruz Teixeira
- URL: https://netmicroscope.com/

6.6. HostView Mobile

KEYWORDS: Quality of Experience - Network monitoring

FUNCTIONAL DESCRIPTION: HostView for mobile runs on Android devices to monitor user system and network performance together with user feedback on Internet experience.

- Contact: Giulio Grassi

7. New Results

7.1. Automated Synthesis of Mediators for Middleware-Layer Protocol Interoperability in the IoT

Participants: Georgios Bouloukakis, Nikolaos Georgantas, Patient Ntumba, Valérie Issarny (MiMove)

To enable direct Internet connectivity of Things, complete protocol stacks need to be deployed on resource-constrained devices. Such protocol stacks typically build on lightweight IPv6 adaptations and may even include a middleware layer supporting high-level application development. However, the profusion of IoT middleware-layer interaction protocols has introduced technology diversity and high fragmentation in the IoT systems landscape with siloed vertical solutions. To enable the interconnection of heterogeneous Things across these barriers, advanced interoperability solutions at the middleware layer are required. In this paper, we introduce a solution for the automated synthesis of protocol mediators that support the interconnection of heterogeneous Things. Our systematic approach relies on the Data eXchange (DeX) connector model, which comprehensively abstracts and represents existing and potentially future IoT middleware protocols. Thanks to DeX, Things seamlessly interconnect through lightweight mediators. We validate our solution with respect to: (i) the support to developers when developing heterogeneous IoT applications; (ii) the runtime performance of the synthesized mediators.

7.2. Probabilistic Event Dropping for Intermittently Connected Subscribers over Pub/Sub Systems

Participants: Georgios Bouloukakis, Nikolaos Georgantas (MiMove), Ioannis Moscholios (Univ of Peloponnese)

Internet of Things (IoT) aim to leverage data from multiple sensors, actuators and devices for improving peoples' daily life and safety. Multiple data sources must be integrated, analyzed from the corresponding application and notify interested stakeholders. To support the data exchange between data sources and stakeholders, the publish/subscribe (pub/sub) middleware is often employed. Pub/sub provides additional mechanisms such as reliable messaging, event dropping, prioritization, etc. The event dropping mechanism is often used to satisfy Quality of Service (QoS) requirements and ensure system stability. To enable event dropping, basic approaches apply finite buffers or data validity periods and more sophisticated ones are information-aware. In this paper, we introduce a pub/sub mechanism for probabilistic event dropping by considering the stakeholders' intermittent connectivity and QoS requirements. We model the pub/sub middleware as a network of queues which includes a novel ON/OFF queueing model that enables the definition of join probabilities. We validate our analytical model via simulation and compare our mechanism with existing ones. Experimental results can be used as insights for developing hybrid dropping mechanisms.

7.3. Adaptive Mediation for Data Exchange in IoT Systems

Participants: Georgios Bouloukakis (MiMove & Univ of California, Irvine), Andrew Chio, Sharad Mehrotra, Nalini Venkatasubramanian (Univ of California, Irvine), Cheng-Hsin Hsu (National Tsing Hua Univ)

Messaging and communication is a critical aspect of next generation Internet-of-Things (IoT) systems where interactions among devices, software systems/services and end-users is the expected mode of operation. Given the diverse and changing communication needs of entities, the data exchange interactions may assume different protocols (MQTT, CoAP, HTTP) and interaction paradigms (point to point, multicast, unicast). In this paper, we address the issue of supporting adaptive communications in IoT systems through a mediation-based architecture for data exchange. Here, components called mediators support protocol translation to bridge the heterogeneity gap. Aiming to provide a placement of mediators to nodes, we introduce an integer linear programming solution that takes as input: a set of Edge nodes, IoT devices, and networking semantics. Our proposed solution achieves adaptive placement resulting in timely interactions between IoT devices for larger topologies of IoT spaces.

7.4. Universal Social Network Bus: Toward the Federation of Heterogeneous Online Social Network Services

Participants: Valérie Issarny, Nikolaos Georgantas, Ehsan Ahvar, Bruno Lefèvre, Shohreh Ahvar (MiMove), Rafael Angarita (ISEP Paris)

Online Social Network Services (OSNSs) are changing the fabric of our society, impacting almost every aspect of it. Over the past few decades, an aggressive market rivalry has led to the emergence of multiple competing, “closed” OSNSs. As a result, users are trapped in the walled gardens of their OSNS, encountering restrictions about what they can do with their personal data, the people they can interact with, and the information they get access to. As an alternative to the platform lock-in, “open” OSNSs promote the adoption of open, standardized APIs. However, users still massively adopt closed OSNSs to benefit from the services’ advanced functionalities and/or follow their “friends,” although the users’ virtual social sphere is ultimately limited by the OSNSs they join. Our work aims at overcoming such a limitation by enabling users to meet and interact beyond the boundary of their OSNSs, including reaching out to “friends” of distinct closed OSNSs. We specifically introduce *Universal Social Network Bus* (USNB), which revisits the “service bus” paradigm that enables interoperability across computing systems to address the requirements of “social interoperability.” USNB features synthetic profiles and personae for interaction across the boundaries of closed and open and profile- and non-profile-based OSNSs through a reference social interaction service. We ran a 1-day workshop with a panel of users who experimented with the USNB prototype to assess the potential benefits of social interoperability for social network users. Results show the positive evaluation of users for USNB, especially as an enabler of applications for civic participation. This further opens up new perspectives for future work, among which includes enforcing security and privacy guarantees.

7.5. Social Middleware for Civic Engagement

Participants: Valérie Issarny, Nikolas Georgantas, Grigoris Piperagkas (MiMove), Rafael Angarita (ISEP Paris)

Civic engagement refers to any collective action towards the identification and solving of public issues. Current civic technologies are traditional Web- or mobile-based platforms that make difficult, or just impossible, the participation of citizens via different communication technologies. Moreover, connected objects sensing physical-world data can nourish participatory processes by providing physical evidence to citizens; however, leveraging these data is not direct and still a time-consuming process for civic technologies developers. We introduce the concept of *social middleware* for civic engagement. Social middleware allows citizens to engage in participatory processes -supported by civic technologies- via their favorite communication tools, and to interact not only with other citizens but also with relevant connected objects and software platforms. The mission of social middleware goes beyond the connection of all these heterogeneous entities. It aims at easing the implementation of distributed applications oriented toward civic engagement by featuring dedicated built-in services.

7.6. Mobile Crowd-Sensing as a Resource for Contextualized Urban Public Policies

Participants: Valŕie Issarny, Bruno Lefèvre, Rachit Agarwal (MiMove), Vivien Mallet (Inria Ange)

Environmental noise is a major pollutant in contemporary cities and calls for the active monitoring of noise levels to spot the locations where it most affects the people's health and well-being. However, due to the complex relationship between environmental noise and its perception by the citizens, it is not sufficient to quantitatively measure environmental noise. We need to collect and aggregate contextualized –both quantitative and qualitative– data about the urban environmental noise so as to be able to study the objective and subjective relationships between sound and living beings. This complex knowledge is a prerequisite for making efficient territorial public policies for soundscapes that are inclined towards living beings welfare. In this paper, we investigate how Mobile Phone Sensing (MPS) –*aka* crowdsensing– enables the gathering of such knowledge, provided the implementation of sensing protocols that are customized according to the context of use and the intended exploitation of the data. Through three case studies that we carried out in France and Finland, we show that MPS is not solely a tool that contributes to sensitizing citizens and decision-makers about noise pollution; it also contributes to increasing our knowledge about the impact of the environmental noise on people's health and well-being in relation to its physical and subjective perception.

7.7. Multi-Sensor Calibration Planning in IoT-Enabled Smart Spaces

Participants: Valérie Issarny (MiMove), Françoise Sailhan (CNAM), Qiuxi Zhu, Md Yusuf Sarwar Uddin, Nalini Venkatasubramanian (University of California, Irvine)

Emerging applications in smart cities and communities require massive IoT deployments using sensors/actuators (things) that can enhance citizens' quality of life and public safety. However, budget constraints often lead to limited instrumentation and/or the use of low-cost sensors that are subject to drift and bias. This raises concerns of robustness and accuracy of the decisions made on uncertain data. To enable effective decision making while fully exploiting the potential of low-cost sensors, we propose to send mobile units (e.g., trained personnel) equipped with high-quality (more expensive) and freshly-calibrated reference sensors so as to carry out calibration in the field. We design and implement an efficient cooperative approach to solve the calibration planning problem, which aims at minimizing the cost of the recurring calibration of multiple sensor types in the long-term operation. We propose a two-phase solution that consists of a sensor selection phase that minimizes the average cost of recurring calibration, and a path planning phase that minimizes the travel cost of multiple calibrators which have load constraints. We provide fast and effective heuristics for both phases. We further build a prototype that facilitates the mapping of the deployment field and provides navigation guidance to mobile calibrators. Extensive use-case-driven simulations show that our proposed approach significantly reduces the average cost compared to naive approaches: up to 30% in a moderate-sized indoor case, and higher in outdoor cases depending on scale

7.8. User-Centric Context Inference for Mobile Crowdsensing

Participants: Yifan Du, Valérie Issarny (MiMove), Françoise Sailhan (CNAM)

Mobile crowdsensing is a powerful mechanism to aggregate hyperlocal knowledge about the environment. Indeed, users may contribute valuable observations across time and space using the sensors embedded in their smartphones. However, the relevance of the provided measurements depends on the adequacy of the sensing context with respect to the phenomena that are analyzed. Our research concentrates more specifically on assessing the sensing context when gathering observations about the physical environment beyond its geographical position in the Euclidean space, i.e., whether the phone is in-/out-pocket, in-/out-door and on-/under-ground. We introduce an online learning approach to the local inference of the sensing context so as to overcome the disparity of the classification performance due to the heterogeneity of the sensing devices as well as the diversity of user behavior and novel usage scenarios. Our approach specifically features a hierarchical algorithm for inference that requires few opportunistic feedbacks from the user, while increasing the accuracy of the context inference per user.

7.9. Let Opportunistic Crowdsensors Work Together for Resource-efficient, Quality-aware Observations

Participants: Yifan Du, Valérie Issarny (MiMove), Françoise Sailhan (CNAM)

Opportunistic crowdsensing empowers citizens carrying hand-held devices to sense physical phenomena of common interest at a large and fine-grained scale without requiring the citizens' active involvement. However, the resulting uncontrolled collection and upload of the massive amount of contributed raw data incur significant resource consumption, from the end device to the server, as well as challenge the quality of the collected observations. Our research tackles both challenges raised by opportunistic crowdsensing, that is, enabling the resource-efficient gathering of relevant observations. To achieve so, we introduce the *BeTogether* middleware fostering context-aware, collaborative crowdsensing at the edge so that co-located crowdsensors operating in the same context, group together to share the work load in a cost- and quality-effective way. Our implementation-driven evaluation of the proposed solution, which leverages a dataset embedding nearly one million entries contributed by 550 crowdsensors over a year, shows that *BeTogether* increases the quality of the collected data while reducing the overall resource cost compared to the cloud-centric approach.

7.10. Detecting Mobile Crowdsensing Context in the Wild

Participants: Rachit Agarwal, Shaan Chopra, Vassilis Christophides, Nikolaos Georgantas, Valérie Issarny (MiMove)

Understanding the sensing context of raw data is crucial for assessing the quality of large crowdsourced spatio-temporal datasets and supporting context-augmented personal trajectories. Detecting sensing contexts in the wild is a challenging task and requires features from smartphone sensors that are not always available. In this paper, we propose three heuristic algorithms for detecting sensing contexts such as in/out-pocket, under/over-ground, and in/out-door for crowdsourced spatio-temporal datasets. These are unsupervised binary classifiers with a small memory footprint and execution time. Using a segment of the Ambiciti real dataset-a feature-limited crowdsourced dataset-we report that our algorithms perform equally well in terms of balanced accuracy (within 4.3%) when compared to machine learning (ML) models reported by an AutoML tool.

7.11. Inferring Streaming Video Quality from Encrypted Traffic: Practical Models and Deployment Experience

Participants: Francesco Bronzino, Sara Ayouibi, Renata Teixeira (MiMove), Paul Schmitt (Princeton), Guilherme Martins, Nick Feamster (University of Chicago)

Inferring the quality of streaming video applications is important for Internet service providers, but the fact that most video streams are encrypted makes it difficult to do so. We develop models that infer quality metrics (i.e., startup delay and resolution) for encrypted streaming video services. Our paper builds on previous work, but extends it in several ways. First, the models work in deployment settings where the video sessions and segments must be identified from a mix of traffic and the time precision of the collected traffic statistics is more coarse (e.g., due to aggregation). Second, we develop a single composite model that works for a range of different services (i.e., Netflix, YouTube, Amazon, and Twitch), as opposed to just a single service. Third, unlike many previous models, our models perform predictions at finer granularity (e.g., the precise startup delay instead of just detecting short versus long delays) allowing to draw better conclusions on the ongoing streaming quality. Fourth, we demonstrate the models are practical through a 16-month deployment in 66 homes and provide new insights about the relationships between Internet "speed" and the quality of the corresponding video streams, for a variety of services; we find that higher speeds provide only minimal improvements to startup delay and resolution. This work was accepted for publication at the ACM SIGMETRICS conference. The models we developed in this work and the findings were the basis for a first-page story published on The Wall Street Journal ("The Truth About Faster Internet: It's Not Worth It"). [0]

[0] The article is available online at: https://www.wsj.com/graphics/faster-internet-not-worth-it/?mod=article_inline&mod=hp_lead_pos5.

7.12. Implications of User Perceived Page Load Time Multi-Modality on Web QoE Measurement

Participants: Renata Teixeira, Vassilis Christophides (MiMove), Flavia Salutari, Diego Da Hora (Telecom Paris Tech), Matteo Varvello (Brave Software), Dario Rossi (Huawei)
Web browsing is one of the most popular applications for both desktop and mobile users. A lot of effort has been devoted to speedup the Web, as well as in designing metrics that can accurately tell whether a webpage loaded fast or not. An often implicit assumption made by industrial and academic research communities is that a *single* metric is sufficient to assess whether a webpage loaded fast. In this work we collect and make publicly available a unique dataset which contains webpage features (e.g., number and type of embedded objects) along with both *objective* and *subjective* Web quality metrics. This dataset was collected by crawling over 100 websites—representative of the top 1 M websites in the Web—while crowdsourcing 6,000 user opinions on *user perceived page load time* (uPLT). In contrast to related work, we show that the uPLT distribution is often multimodal and that, in practice, no more than three modes are present. The main conclusion drawn from our analysis is that, for complex webpages, each of the different objective QoE metrics proposed in the literature (such as AFT, TTI, PLT, etc.) is suited to approximate one of the different uPLT modes.

7.13. The News We Like Are Not the News We Visit: News Categories Popularity in Usage Data

Participants: Renata Teixeira (MiMove), Giuseppe Scavo (MiMove, Nokia Bell Labs), Zied Ben-Houidi (Nokia Bell-Labs), Stefano Traverso, Marco Mellia (Politecnico di Torino)
Most of our knowledge about online news consumption comes from survey-based news market reports, partial usage data from a single editor, or what people publicly share on social networks. Our work published on the 13th International AAAI Conference on Web and Social Media (ICWSM-2019) complements these sources by presenting the first holistic study of visits across online news outlets that a population uses to read news. We monitored the entire network traffic generated by Internet users in four locations in Italy. Together these users generated 80 million visits to 5.4 million news articles in about one year and a half. This unique view allowed us to evaluate how usage data complements existing data sources. We find for instance that only 16% of news visits in our datasets came from online social networks. In addition, the popularity of news categories when considering all visits is quite different from the one when considering only news discovered on social media, or visits to a single major news outlet. Interestingly, a substantial mismatch emerges between self-reported news-category preferences (as measured by Reuters Institute in the same year and same country) and their actual popularity in terms of visits in our datasets. In particular, unlike self-reported preferences expressed by users in surveys that put “Politics”, “Science” and “International” as the most appreciated categories, “Tragedies and Weird news" and “Sport” are by far the most visited. Our paper discusses two possible causes of this mismatch and conjecture that the most plausible reason is the disassociation that may occur between individuals’ cognitive values and their cue-triggered attraction.

7.14. Classification of Load Balancing in the Internet

Participants: Renata Teixeira (MiMove), Rafael Almeida, Ítalo Cunha (Universidade Federal de Minas Gerais), Darryl Veitch (University of Technology Sydney), Christophe Diot (Google)

Recent advances in programmable data planes, software-defined networking, and the adoption of IPv6, support novel, more complex load balancing strategies. We introduce the Multipath Classification Algorithm (MCA), a probing algorithm that extends traceroute to identify and classify load balancing in Internet routes. MCA extends existing formalism and techniques to consider that load balancers may use arbitrary combinations of bits in the packet header for load balancing. We propose optimizations to reduce probing cost that are applicable to MCA and existing load balancing measurement techniques. Through large-scale measurement campaigns, we characterize and study the evolution of load balancing on the IPv4 and IPv6 Internet with multiple transport protocols. Our results that will appear in the IEEE INFOCOM 2020 conference show that load balancing is more prevalent and that load balancing strategies are more mature than previous characterizations have found.

7.15. MinoanER: Schema-Agnostic, Non-Iterative, Massively Parallel Resolution of Web Entities

Participants: Vassilis Christophides (MiMove), Vasilis Efthymiou (IBM Almaden Research Center), George Papadakis (Univ of Athens), Kostas Stefanidis (Univ of Tampere)

Entity Resolution (ER) aims to identify different descriptions in various Knowledge Bases (KBs) that refer to the same entity. ER is challenged by the Variety, Volume and Veracity of entity descriptions published in the Web of Data. To address them, we propose the MinoanER framework that simultaneously fulfills full automation, support of highly heterogeneous entities, and massive parallelization of the ER process. MinoanER leverages a token-based similarity of entities to define a new metric that derives the similarity of neighboring entities from the most important relations, as they are indicated only by statistics. A composite blocking method is employed to capture different sources of matching evidence from the content, neighbors, or names of entities. The search space of candidate pairs for comparison is compactly abstracted by a novel disjunctive blocking graph and processed by a non-iterative, massively parallel matching algorithm that consists of four generic, schema-agnostic matching rules that are quite robust with respect to their internal configuration. We demonstrate that the effectiveness of MinoanER is comparable to existing ER tools over real KBs exhibiting low Variety, but it outperforms them significantly when matching KBs with high Variety.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

"Application Performance Bottleneck Detection", Comcast Gift to R. Teixeira 2018-2019.

9. Partnerships and Cooperations

9.1. National Initiatives

"BottleNet: Understanding and Diagnosing End-to-end Communication Bottlenecks of the Internet", project funded by the French research agency (ANR), from Feb 2016 to Sep 2020.

9.1.1. Inria Support

9.1.1.1. Inria IPL BetterNet

Participants: Renata Teixeira, Vassilis Christophides, Giulio Grassi.

- **Name:** BetterNet – *An observatory to measure and improve Internet service access from user experience*
- **Period:** [2016 – 2019]
- **Inria teams:** Diana, Dionysos, Inria Chile, Madynes, MiMove, Spirals
- **URL:** https://project.inria.fr/betternet/

BetterNet aims at building and delivering a scientific and technical collaborative observatory to measure and improve the Internet service access as perceived by users. In this Inria Project Lab, we will propose new original user-centered measurement methods, which will associate social sciences to better understand Internet usage and the quality of services and networks. Our observatory can be defined as a vantage point, where:

1. tools, models and algorithms/heuristics will be provided to collect data,
2. acquired data will be analyzed, and shared appropriately with scientists, stakeholders and civil society,
3. and new value-added services will be proposed to end-users.

9.1.1.2. Inria ADT SocialBus

Participants: Valérie Issarny , Rafael Angarita, Nikolaos Georgantas, Ehsan Ahvar , Lior Diler .

- **Name:** SocialBus – *Contributing to the development of SocialBus - A Universal Social Network Bus*
- **Period:** [July 2018 – June 2019 ; November 2019 – October 2020]
- **Partners:** Inria MiMove.

Computer-mediated communication can be defined as any form of human communication achieved through computer technology. From its beginnings, it has been shaping the way humans interact with each other, and it has influenced many areas of society. There exist a plethora of social interaction services enabling computer-mediated social communication (e.g., Skype, Facebook Messenger, Telegram, WhatsApp, Twitter, Slack, etc.). Based on personal preferences, users may prefer a social interaction services rather than another. As a result, users sharing same interests may not be able to interact since they are using incompatible technologies.

To tackle the above interoperability barrier, we propose SocialBus, a middleware solution targeted to enable the interaction via heterogeneous social interaction services. The ADT specifically supports the related implementation through the funding an engineer, toward technology transfer in the mid-term.

The SocialBus software is available under the AGPL open source license at https://gitlab.inria.fr/usnb/universal-social-network-bus.

9.2. International Initiatives

9.2.1. Inria International Labs

Inria@EastCoast

Associate Team involved in the International Lab:

9.2.1.1. HOMENET

Title: Home network diagnosis and security

International Partner (Institution - Laboratory - Researcher):

Princeton (United States) - Computer Science - Nick Feamster

Start year: 2017

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

Modern households connect a multitude of networked devices (ranging from laptops and smartphones to a number of Internet of Things devices) via a home network. Most home networks, however, do not have a technically skilled network administrator for managing the network, for example to identify faulty equipment or take steps to secure end hosts such as applying security patches. Home networks represent a particularly challenging environment due to the diversity of devices, applications, and services users may connect. The goal of HOMENET is to assist users in diagnosing and securing their home networks. Our approach is based on developing new algorithms and mechanisms that will run on the home router (or in-collaboration with the router). The router connects the home network to the rest of the Internet; it is hence the ideal place to secure home devices and to distinguish problems that happen in the home from those happening elsewhere. We will address a number of research challenges for example in device discovery and fingerprinting, anomaly detection in the Internet of Things, home network diagnosis (including wireless diagnosis). HOMENET will bring together two leading research teams in the network measurement arena with successful prior collaboration. Moreover, Princeton brings an existing home router platform and expertise in security, wireless, and software-defined networks; and Muse brings an existing Web-based measurement platform, and expertise in traffic-based profiling and anomaly detection.

Inria@SiliconValley

Associate Team involved in the International Lab:

9.2.1.2. MINES

Title: Adaptive Communication Middleware for Resilient Sensing & Actuation IN Emergency Response Scenarios

International Partner (Institution - Laboratory - Researcher):

University of California, Irvine (United States) - Information and Computer Science - Nalini Venkatasubramanian

Start year: 2018

See also: http://mimove-apps.paris.inria.fr/mines/index.html

Emerging smart-city and smart-community efforts will require a massive deployment of connected entities (Things) to create focused smartspaces. Related applications will enhance citizen quality of life and public safety (e.g., providing safe evacuation routes in fires). However, supporting IoT deployments are heterogeneous and can be volatile and failure-prone as they are often built upon low-powered, mobile and inexpensive devices - the presence of faulty components and intermittent network connectivity, especially in emergency scenarios, tend to deliver inaccurate/delayed information. The MINES associate team addresses the resulting challenge of enabling interoperability and resilience in large-scale IoT systems through the design and development of a dedicated middleware. More specifically, focusing on emergency situations, the MINES middleware will: (i) enable the dynamic composition of IoT systems from any and all available heterogeneous devices; (ii) support the timely and reliable exchange of critical data within and across IoT in the enabled large-scale and dynamic system over heterogeneous networks. Finally, the team will evaluate the proposed solution in the context of emergency response scenario use cases.

9.2.2. Inria International Partners

9.2.2.1. Informal International Partners

- Northeastern University (Prof. David Choffnes): We are working on methods based on active probing to diagnose poor video quality.
- Universidade Federal do Rio de Janeiro, Brazil (Prof. Edmundo Souza e Silva): We are working on characterizing Internet bottlenecks.
- Universidade Federal de Goias, Brazil (Prof. Fabio Costa): We are working on service selection and cloud resource allocation for QoS-aware enactment of service choreographies.

9.3. International Research Visitors

9.3.1. Visits of International Scientists

Mark Crovella from Boston University was visiting professor at Inria.

9.3.2. Visits to International Teams

9.3.2.1. Sabbatical programme

Renata Teixeira is visiting scholar at the Computer Science Department at Stanford University.

9.3.2.2. Research Stays Abroad

Georgios Bouloukakis was Inria postdoctoral fellow at University of California, Irvine, in the context of the Inria@SiliconValley program.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. Member of the Organizing Committees

- R. Teixeira, member of the selection committee of the Heidelberg Laureate Forum 2019–2021.

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- V. Issarny, TPC co-chair of: The International Conference on the IoT (ICIOT) at SCF'19 (June 2019, San Diego, CA, USA); and The 5th 2CM/IEEE Conference on Internet of Things Design & Implementation (IoTDI) at CPS-IoT week (April 2020, Sydney, Australia).

10.1.2.2. Member of the Conference Program Committees

- V. Issarny, member of the TPC of the following international conferences: IEEE CIC'19, ACM EUROSYS'20 (Heavy PC), ACM ESEC/FSE'19, IEEE ICDCS'20, ICSE'20, ICSE-SEIS'19 & 20, ACM/IEEE IoTDI'19, and ACM/IFIP Middleware'19.
- R. Teixeira, member of the TPC of: USENIX Symposium on Networked Systems Design and Implementation (NSDI) 2020, Passive and Active Measurement Conference (PAM) 2020, and the Buffer Sizing Workshop 2019.
- N. Georgantas, member of the TPC of the following international conferences: SAC'19&'20, SOSE'19&'20, The Web Conference'19&'20.
- N. Georgantas, member of the TPC of the following international workshops: MRT'19, IoT-ASAP'19, ARM'19, ASYDE'19.
- V. Christophides, member of the TPC of the EDBT/ICDT 2020 Joint Conference, IEEE ICDE 2020, and the 28th ACM International Conference on Information and Knowledge Management (CIKM).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- V. Issarny, member of the following editorial boards: The ACM Transaction on Autonomous and Adaptive Systems (TAAS), The ACM Transactions on the Internet of Things (TIOT), The IEEE Transactions on Services Computing (TSC), and The IEEE Transactions on Software Engineering (TSE).
- N. Georgantas, Associate editor, International Journal of Ambient Computing and Intelligence (IJACI).
- V. Christophides, Associate editor, MDPI Open Access Journal of Internet of Things (IoT) and MDPI Open Access Journal of Future Internet.

10.1.4. Invited Talks

- R. Teixeira "Passive Measurements of Residential Internet Performance", Google Networking Summit, 2019.
- R. Teixeira, "Lightweight, General Inference of Streaming Video Quality from Encrypted Traffic", Workshop em ciencia de dados LNCC, Brazil, 2019.
- N. Georgantas, "Enabling Emergent Mobile IoT Systems: Middleware Interoperability & QoS Analysis", Eclipse IoT Day Paris Saclay 2019.
- V. Christophides, "Detecting Outliers in IoT Data Streams", IRT SystemX, 2019.

10.1.5. Leadership within the Scientific Community

- V. Christophides, Member of the Executive Board of the EDBT Association.

10.1.6. Scientific Expertise

- V. Issarny, member of: the FWP expert panel for PhD fellowships on strategic basic research, and the expert panel of the 2019 FET open call.
- V. Issarny, member of: the ARCEP Scientific Council (appointed), and The ACM Europe Council (Elected).
- N. Georgantas, member of: the EDITE Doctoral School selection committe for PhD fellowships, 2018 & 2019 .

10.2. Teaching - Supervision - Juries

10.2.1. Supervision

PhD: Sara el Aouad, “Personalized, Aspect-based Summarization of Movie Reviews”, Sorbonne Université, April 26, 2019, R. Teixeira, V. Christophides, C. Diot.

PhDs in progress:

- Yifan Du (Since October 2017): "In-network collaborative crowd-Xing", Sorbonne University, V. Issarny and F. Sailhan (CNAM)
- Grigoris Piperagkas (Since October 2018): "Leveraging universal social networking and the IoT for urban-scale participatory systems", Sorbonne University, V. Issarny and R. Angarita (ISEP).
- William Aboucaya (Since October 2019): "Version control for urban participatory systems", Sorbonne University, V. Issarny and R. Angarita (ISEP).
- Dimitris Tsolovos (Since March 2017): "A privacy-by-design middleware for urban-scale mobile crowdsensing", UVSQ, N. Anciaux (Inria PETRUS @ Saclay) and V. Issarny.
- Patient Ntumba (Since August 2018): “Dynamic management of IoT data stream analytics in the edge-fog-cloud continuum”, Sorbonne University, N. Georgantas and Vassilis Christophides.

10.2.2. Juries

- V. Issarny, member of the following juries: Christian Cabrera (11/19, PhD, Trinity College Dublin, Ireland), Gilles Tredan (06/19, HDR, University of Toulous), and Nic Volanschi (11/19, HDR, University of Bordeaux).

11. Bibliography

Major publications by the team in recent years

[1] A. BENNACEUR, V. ISSARNY.*Automated Synthesis of Mediators to Support Component Interoperability*, in "IEEE Transactions on Software Engineering", 2015, 22, https://hal.inria.fr/hal-01076176

[2] B. BILLET, V. ISSARNY.*Spinel: An Opportunistic Proxy for Connecting Sensors to the Internet of Things*, in "ACM Transactions on Internet Technology", March 2017, vol. 17, n^{o} 2, p. 1 - 21 [*DOI :* 10.1145/3041025], https://hal.inria.fr/hal-01505879

[3] G. BLAIR, A. BENNACEUR, N. GEORGANTAS, P. GRACE, V. ISSARNY, V. NUNDLOLL, M. PAOLUCCI.*The Role of Ontologies in Emergent Middleware: Supporting Interoperability in Complex Distributed Systems*, in "Big Ideas track of ACM/IFIP/USENIX 12th International Middleware Conference", Lisbon, Portugal, 2011, http://hal.inria.fr/inria-00629059/en

[4] M. CAPORUSCIO, P.-G. RAVERDY, V. ISSARNY.*ubiSOAP: A Service Oriented Middleware for Ubiquitous Networking*, in "IEEE Transactions on Services Computing", 2012, vol. 99 [*DOI :* 10.1109/TSC.2010.60], http://hal.inria.fr/inria-00519577

[5] Í. CUNHA, R. TEIXEIRA, D. VEITCH, C. DIOT.*DTrack: A System to Predict and Track Internet Path Changes*, in "IEEE/ACM Transactions on Networking", August 2014, vol. 22, n^o 4, p. 1025–1038, https://hal.inria.fr/hal-01097439

[6] O. GOGA, P. LOISEAU, R. SOMMER, R. TEIXEIRA, K. P. GUMMADI.*On the Reliability of Profile Matching Across Large Online Social Networks*, in "KDD'15: ACM SIGDD Conference on Knowledge Discovery and Data Mining", Sydeny, Australia, August 2015 [*DOI :* 10.1145/2783258.2788601], https://hal.inria.fr/hal-01162402

[7] S. HACHEM, A. PATHAK, V. ISSARNY.*Service-Oriented Middleware for Large-Scale Mobile Participatory Sensing*, in "Pervasive and Mobile Computing", 2014, http://hal.inria.fr/hal-00872407

[8] V. ISSARNY, N. GEORGANTAS, S. HACHEM, A. ZARRAS, P. VASSILIADIS, M. AUTILI, M. A. GEROSA, A. BEN HAMIDA.*Service-Oriented Middleware for the Future Internet: State of the Art and Research Directions*, in "Journal of Internet Services and Applications", May 2011, vol. 2, n^o 1, p. 23-45 [*DOI :* 10.1007/S13174-011-0021-3], http://hal.inria.fr/inria-00588753/en

[9] K. MIRYLENKA, V. CHRISTOPHIDES, T. PALPANAS, I. PEFKIANAKIS, M. MAY.*Characterizing Home Device Usage From Wireless Traffic Time Series*, in "19th International Conference on Extending Database Technology (EDBT)", Bordeaux, France, March 2016, https://hal.inria.fr/hal-01249778

Publications of the year

Doctoral Dissertations and Habilitation Theses

[10] S. EL AOUAD.*Personalized, Aspect-based Summarization of Movie Reviews*, Sorbonne universite, April 2019, https://hal.inria.fr/tel-02444980

Articles in International Peer-Reviewed Journal

[11] R. ANGARITA, B. LEFÈVRE, S. AHVAR, E. AHVAR, N. GEORGANTAS, V. ISSARNY.*Universal Social Network Bus: Towards the Federation of Heterogeneous Online Social Network Services*, in "ACM Transactions on Internet Technology", 2019, forthcoming [*DOI :* 10.1145/3323333], https://hal.inria.fr/hal-02072544

[12] G. BOULOUKAKIS, N. GEORGANTAS, P. NTUMBA, V. ISSARNY.*Automated synthesis of mediators for middleware-layer protocol interoperability in the IoT*, in "Future Generation Computer Systems", December 2019, vol. 101, p. 1271-1294 [*DOI :* 10.1016/J.FUTURE.2019.05.064], https://hal.inria.fr/hal-02304074

[13] F. BRONZINO, P. SCHMITT, S. AYOUBI, G. MARTINS, R. TEIXEIRA, N. FEAMSTER.*Inferring Streaming Video Quality from Encrypted Traffic: Practical Models and Deployment Experience*, in "Proceedings of the ACM on Measurement and Analysis of Computing Systems ", December 2019, vol. 3, n^o 3, https://hal.inria.fr/hal-02396405

[14] Y. DU, V. ISSARNY, F. SAILHAN.*When the Power of the Crowd Meets the Intelligence of the Middleware : The Mobile Phone Sensing Case*, in "Operating Systems Review", July 2019, vol. 53, n^o 1, p. 85-90 [*DOI :* 10.1145/3352020.3352033], https://hal.inria.fr/hal-02238256

[15] B. LEFÈVRE, R. AGARWAL, V. ISSARNY, V. MALLET.*Mobile Crowd-Sensing as a Resource for Contextualized Urban Public Policies: A Study using Three Use Cases on Noise and Soundscape Monitoring*, in "Cities & Health", May 2019 [*DOI :* 10.1080/23748834.2019.1617656], https://hal.inria.fr/hal-02127052

International Conferences with Proceedings

[16] R. AGARWAL, S. CHOPRA, V. CHRISTOPHIDES, N. GEORGANTAS, V. ISSARNY.*Detecting Mobile Crowdsensing Context in the Wild*, in "20th IEEE International Conference on Mobile Data Management (MDM) 2019", Hong Kong, Hong Kong SAR China, June 2019, https://hal.inria.fr/hal-02151434

[17] R. ALMEIDA, Í. CUNHA, R. TEIXEIRA, D. VEITCH, C. DIOT.*Classification of Load Balancing in the Internet*, in "IEEE INFOCOM 2020 - International Conference on Computer Communications", Beijing, China, April 2020, https://hal.inria.fr/hal-02396406

[18] R. ANGARITA, N. GEORGANTAS, V. ISSARNY.*Social Middleware for Civic Engagement*, in "ICDCS 2019 - 39th IEEE International Conference on Distributed Computing Systems - Vision Track", Dallas, United States, July 2019, https://hal.inria.fr/hal-02162736

[19] G. BOULOUKAKIS, I. MOSCHOLIOS, N. GEORGANTAS.*Probabilistic Event Dropping for Intermittently Connected Subscribers over Pub/Sub Systems*, in "ICC 2019 - IEEE International Conference on Communications", Shanghai, China, May 2019, https://hal.inria.fr/hal-02058417

[20] A. CHIO, G. BOULOUKAKIS, C.-H. HSU, S. MEHROTRA, N. VENKATASUBRAMANIAN.*Adaptive Mediation for Data Exchange in IoT Systems*, in "ARM 2019 - 18th Workshop on Adaptive and Reflexive Middleware", Davis, United States, December 2019 [*DOI :* 10.1145/3366612.3368122], https://hal.inria.fr/hal-02319447

[21] Y. DU, V. ISSARNY, F. SAILHAN.*User-centric context inference for mobile crowdsensing*, in "IoTDI 2019: ACM/IEEE International Conference on Internet of Things Design and Implementation", Montreal, Canada, ACM Press, April 2019, p. 261-266 [*DOI :* 10.1145/3302505.3310088], https://hal.inria.fr/hal-02082034

[22] V. EFTHYMIOU, G. PAPADAKIS, K. STEFANIDIS, V. CHRISTOPHIDES.*MinoanER: Schema-Agnostic, Non-Iterative, Massively Parallel Resolution of Web Entities*, in "EDBT/ICDT 2019 Joint Conference", Lisbon, Portugal, March 2019, https://hal.inria.fr/hal-01960933

[23] V. ISSARNY, B. BILLET, G. BOULOUKAKIS, D. FLORESCU, C. TOMA.*LATTICE: A Framework for Optimizing IoT System Configurations at the Edge*, in "ICDCS 2019 - 39th IEEE International Conference on Distributed Computing Systems", Dallas, Texas, United States, July 2019, https://hal.inria.fr/hal-02161795

[24] I. MORANDI, F. BRONZINO, R. TEIXEIRA, S. SUNDARESAN.*Service traceroute: Tracing Paths of Application Flows*, in "PAM 2019 - Passive and Active Measurement Conference", Puerto Varas, Chile, March 2019, https://hal.inria.fr/hal-02058307

[25] G. TEXIER, V. ISSARNY.*La puissance du crowdsensing au service de la durée de vie des réseaux IoT urbains*, in "ALGOTEL 2019 - 21èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications", Saint Laurent de la Cabrerisse, France, June 2019, https://hal.archives-ouvertes.fr/hal-02118074

[26] Q. ZHU, F. SAILHAN, Y. S. UDDIN, V. ISSARNY, N. VENKATASUBRAMANIAN.*Multi-Sensor Calibration Planning in IoT-Enabled Smart Spaces*, in "39th International Conference on Distributed Computing Systems (ICDCS)", Dallas, United States, July 2019, https://hal.inria.fr/hal-02136852

[27] B. H. ZIED, G. SCAVO, S. TRAVERSO, R. TEIXEIRA, M. MELLIA, S. GANGULY.*The News We Like Are Not the News We Visit: News Categories Popularity in Usage Data*, in "ICWSM-2019 - 13th International AAAI Conference on Web and Social Media", Munich, Germany, June 2019, https://hal.inria.fr/hal-02268409

Other Publications

[28] R. AGARWAL, S. CHOPRA, V. CHRISTOPHIDES, N. GEORGANTAS, V. ISSARNY.*Inferring Context of Mobile Data Crowdsensed in the Wild*, July 2019, NetMob 2019 - Conference on the scientific analysis of mobile phone datasets, Poster, https://hal.inria.fr/hal-02132194

Project-Team MOKAPLAN

Advances in Numerical Calculus of Variations

IN COLLABORATION WITH: CEREMADE

IN PARTNERSHIP WITH:
CNRS

Université Paris-Dauphine

RESEARCH CENTER
Paris

THEME
Numerical schemes and simulations

Table of contents

Project-Team MOKAPLAN

Creation of the Team: 2013 January 01, updated into Project-Team: 2015 December 01

Keywords:

Computer Science and Digital Science:

A5.3. - Image processing and analysis
A5.9. - Signal processing
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.6. - Optimization

Other Research Topics and Application Domains:

B1.2. - Neuroscience and cognitive science
B9.5.2. - Mathematics
B9.5.3. - Physics
B9.5.4. - Chemistry
B9.6.3. - Economy, Finance

1. Team, Visitors, External Collaborators

Research Scientists

Jean-David Benamou [Team leader, Inria, Senior Researcher, HDR]
Vincent Duval [Inria, Senior Researcher]
Thomas Gallouët [Inria, Researcher]
Irene Waldspurger [CNRS, Researcher]

Faculty Members

Guillaume Carlier [Univ de Dauphine, Professor, HDR]
Paul Pegon [Univ Paris Dauphine, Assistant Professor]
Bruno Nazaret [Univ Panthéon Sorbonne, Professor, from Sep 2019]

External Collaborators

Yohann de Castro [Ecole Centrale Lyon, Professor, HDR]
François-Xavier Vialard [Univ Gustave Eiffel, Professor, HDR]
Andrea Natale [Univ Paris Sacay, since Nov 2019]
Theo Golvet [Univ de Dauphine,until Oct 2019]
Yann Brenier [Ecole Normale Supérieure Paris]
Gwendoline de Bie [Ecole Normale Superieure Paris]
Quentin Merigot [CNRS]
Gabriel Peyré [CNRS, HDR]
Shuangjian Zhang [Ecole Normale Superieure Paris]

PhD Students

Paul Catala [Ecole Normale Supérieure Paris]
Katharina Eichinger [Univ Paris Sciences et Lettres, from Oct 2019]
Aude Genevay [Univ de Dauphine, until Mars 2019]
Lucas Martinet [Univ Paris Dauphine]
Quentin Ismael Petit [Univ Paris Dauphine]
Romain Petit [Univ Paris Sciences et Lettres, from Oct 2019]

Giorgi Rukhaia [Univ Paris Dauphine]
Gabriele Todeschi [Univ Paris Sciences et Lettres]
Miao Yu [Univ Denis Diderot]

Post-Doctoral Fellows

Guillaume Mijoule [Inria]
Andrea Natale [Ecole Normale Supérieure Paris, until Nov 2019]
Luca Tamanini [Univ Paris Sciences et Lettres, from Nov 2019]
Vincent Reverdy [Ecole Normale Supérieure Paris, from Dec 2019]

2. Overall Objectives

2.1. Introduction

The last decade has witnessed a remarkable convergence between several sub-domains of the calculus of variations, namely optimal transport (and its many generalizations), infinite dimensional geometry of diffeomorphisms groups and inverse problems in imaging (in particular sparsity-based regularization). This convergence is due to (i) the mathematical objects manipulated in these problems, namely sparse measures (e.g. coupling in transport, edge location in imaging, displacement fields for diffeomorphisms) and (ii) the use of similar numerical tools from non-smooth optimization and geometric discretization schemes. Optimal Transportation, diffeomorphisms and sparsity-based methods are powerful modeling tools, that impact a rapidly expanding list of scientific applications and call for efficient numerical strategies. Our research program shows the important part played by the team members in the development of these numerical methods and their application to challenging problems.

2.2. Static Optimal Transport and Generalizations

2.2.1. Optimal Transport, Old and New.

Optimal Mass Transportation is a mathematical research topic which started two centuries ago with Monge's work on the "Théorie des déblais et des remblais" (see [103]). This engineering problem consists in minimizing the transport cost between two given mass densities. In the 40's, Kantorovich [110] introduced a powerful linear relaxation and introduced its dual formulation. The *Monge-Kantorovich* problem became a specialized research topic in optimization and Kantorovich obtained the 1975 Nobel prize in economics for his contributions to resource allocations problems. Since the seminal discoveries of Brenier in the 90's [56], Optimal Transportation has received renewed attention from mathematical analysts and the Fields Medal awarded in 2010 to C. Villani, who gave important contributions to Optimal Transportation and wrote the modern reference monographs [150], [149], arrived at a culminating moment for this theory. Optimal Mass Transportation is today a mature area of mathematical analysis with a constantly growing range of applications. Optimal Transportation has also received a lot of attention from probabilists (see for instance the recent survey [120] for an overview of the Schrödinger problem which is a stochastic variant of the Benamou-Brenier dynamical formulation of optimal transport). The development of numerical methods for Optimal Transportation and Optimal Transportation related problems is a difficult topic and comparatively underdeveloped. This research field has experienced a surge of activity in the last five years, with important contributions of the MOKAPLAN group (see the list of important publications of the team). We describe below a few of recent and less recent Optimal Transportation concepts and methods which are connected to the future activities of MOKAPLAN :

Brenier's theorem [57] characterizes the unique optimal map as the gradient of a convex potential. As such Optimal Transportation may be interpreted as an infinite dimensional optimisation problem under "convexity constraint": i.e. the solution of this infinite dimensional optimisation problem is a convex potential. This connects Optimal Transportation to "convexity constrained" non-linear variational problems such as, for instance, Newton's problem of the body of minimal resistance. The value function of the optimal transport problem is also known to define a distance between source and target densities called the *Wasserstein distance* which plays a key role in many applications such as image processing.

2.2.2. Monge-Ampère Methods.

A formal substitution of the optimal transport map as the gradient of a convex potential in the mass conservation constraint (a Jacobian equation) gives a non-linear Monge-Ampère equation. Caffarelli [65] used this result to extend the regularity theory for the Monge-Ampère equation. In the last ten years, it also motivated new research on numerical solvers for non-linear degenerate Elliptic equations [91] [118] [42] [43] and the references therein. Geometric approaches based on Laguerre diagrams and discrete data [127] have also been developed. Monge-Ampère based Optimal Transportation solvers have recently given the first linear cost computations of Optimal Transportation (smooth) maps.

2.2.3. Generalizations of OT.

In recent years, the classical Optimal Transportation problem has been extended in several directions. First, different ground costs measuring the “physical" displacement have been considered. In particular, well posedness for a large class of convex and concave costs has been established by McCann and Gangbo [102]. Optimal Transportation techniques have been applied for example to a Coulomb ground cost in Quantum chemistry in relation with Density Functional theory [87]. Given the densities of electrons Optimal Transportation models the potential energy and their relative positions. For more than more than 2 electrons (and therefore more than 2 densities) the natural extension of Optimal Transportation is the so called Multi-marginal Optimal Transport (see [131] and the references therein). Another instance of multi-marginal Optimal Transportation arises in the so-called Wasserstein barycenter problem between an arbitrary number of densities [27]. An interesting overview of this emerging new field of optimal transport and its applications can be found in the recent survey of Ghoussoub and Pass [130].

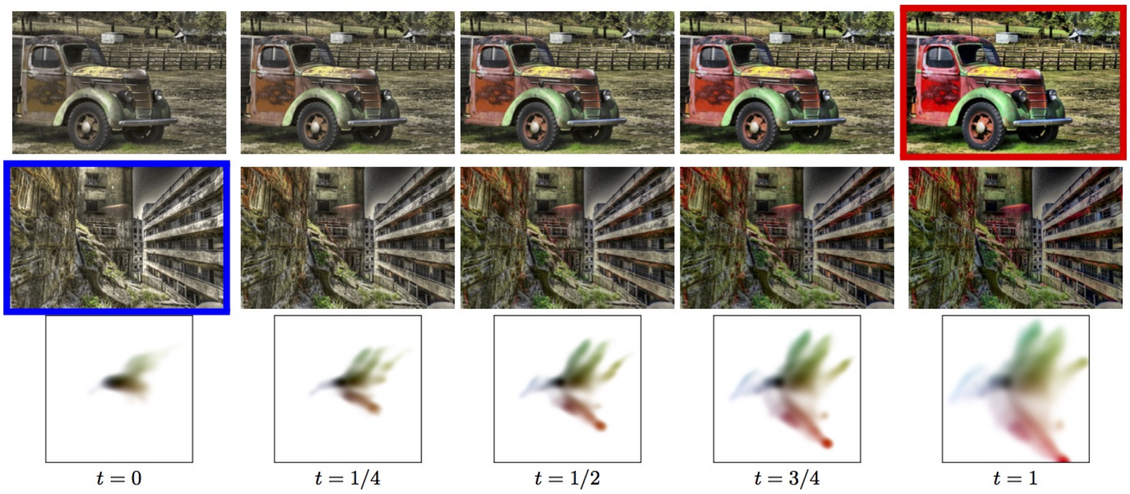

Figure 1. Example of color transfer between two images, computed using the method developed in [39], see also [143]. The image framed in red and blue are the input images. Top and middle row: adjusted image where the color of the transported histogram has been imposed. Bottom row: geodesic (displacement) interpolation between the histogram of the chrominance of the image.

2.2.4. Numerical Applications of Optimal Transportation.

Optimal transport has found many applications, starting from its relation with several physical models such as the semi-geostrophic equations in meteorology [107], [89], [88], [36], [117], mesh adaptation [116], the reconstruction of the early mass distribution of the Universe [99], [58] in Astrophysics, and the numerical optimisation of reflectors following the Optimal Transportation interpretation of Oliker [66] and Wang [151]. Extensions of OT such as multi-marginal transport has potential applications in Density Functional Theory , Generalized solution of Euler equations [55] (DFT) and in statistics and finance [33], [101] .... Recently, there has been a spread of interest in applications of OT methods in imaging sciences [50], statistics [47] and machine learning [90]. This is largely due to the emergence of fast numerical schemes to approximate the transportation distance and its generalizations, see for instance [39]. Figure 1 shows an example of application of OT to color transfer. Figure 9 shows an example of application in computer graphics to interpolate between input shapes.

2.3. Diffeomorphisms and Dynamical Transport

2.3.1. Dynamical transport.

While the optimal transport problem, in its original formulation, is a static problem (no time evolution is considered), it makes sense in many applications to rather consider time evolution. This is relevant for instance in applications to fluid dynamics or in medical images to perform registration of organs and model tumor growth.

In this perspective, the optimal transport in Euclidean space corresponds to an evolution where each particule of mass evolves in straight line. This interpretation corresponds to the *Computational Fluid Dynamic* (CFD) formulation proposed by Brenier and Benamou in [35]. These solutions are time curves in the space of densities and geodesics for the Wasserstein distance. The CFD formulation relaxes the non-linear mass conservation constraint into a time dependent continuity equation, the cost function remains convex but is highly non smooth. A remarkable feature of this dynamical formulation is that it can be re-cast as a convex but non smooth optimization problem. This convex dynamical formulation finds many non-trivial extensions and applications, see for instance [37]. The CFD formulation also appears to be a limit case of *Mean Fields games* (MFGs), a large class of economic models introduced by Lasry and Lions [112] leading to a system coupling an Hamilton-Jacobi with a Fokker-Planck equation. In contrast, the Monge case where the ground cost is the euclidan distance leads to a static system of PDEs [52].

2.3.2. Gradient Flows for the Wasserstein Distance.

Another extension is, instead of considering geodesic for transportation metric (i.e. minimizing the Wasserstein distance to a target measure), to make the density evolve in order to minimize some functional. Computing the steepest descent direction with respect to the Wasserstein distance defines a so-called Wasserstein gradient flow, also known as *JKO gradient flows* after its authors [108]. This is a popular tool to study a large class of non-linear diffusion equations. Two interesting examples are the Keller-Segel system for chemotaxis [109], [82] and a model of congested crowd motion proposed by Maury, Santambrogio and Roudneff-Chupin [123]. From the numerical point of view, these schemes are understood to be the natural analogue of implicit scheme for linear parabolic equations. The resolution is however costly as it involves taking the derivative in the Wasserstein sense of the relevant energy, which in turn requires the resolution of a large scale convex but non-smooth minimization.

2.3.3. Geodesic on infinite dimensional Riemannian spaces.

To tackle more complicated warping problems, such as those encountered in medical image analysis, one unfortunately has to drop the convexity of the functional involved in defining the gradient flow. This gradient flow can either be understood as defining a geodesic on the (infinite dimensional) group of diffeomorphisms [32], or on a (infinite dimensional) space of curves or surfaces [152]. The de-facto standard to define, analyze and compute these geodesics is the “Large Deformation Diffeomorphic Metric Mapping” (LDDMM) framework of Trouvé, Younes, Holm and co-authors [32], [106]. While in the CFD formulation of optimal transport, the metric on infinitesimal deformations is just the L^2 norm (measure according to the density being transported), in LDDMM, one needs to use a stronger regularizing metric, such as Sobolev-like norms or reproducing kernel Hilbert spaces (RKHS). This enables a control over the smoothness of the deformation which is crucial for many applications. The price to pay is the need to solve a non-convex optimization

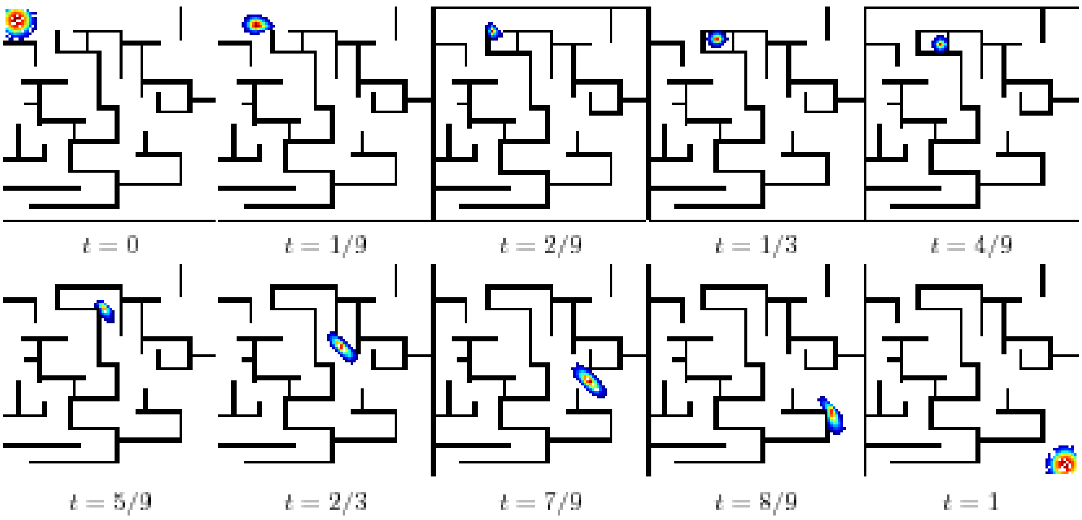

Figure 2. Examples of displacement interpolation (geodesic for optimal transport) according to a non-Euclidean Riemannian metric (the mass is constrained to move inside a maze) between to input Gaussian distributions. Note that the maze is dynamic: its topology change over time, the mass being "trapped" at time $t = 1/3$.

problem through geodesic shooting method [124], which requires to integrate backward and forward the geodesic ODE. The resulting strong Riemannian geodesic structure on spaces of diffeomorphisms or shapes is also pivotal to allow us to perform statistical analysis on the tangent space, to define mean shapes and perform dimensionality reduction when analyzing large collection of input shapes (e.g. to study evolution of a diseases in time or the variation across patients) [67].

2.4. Sparsity in Imaging

2.4.1. Sparse ℓ^1 regularization.

Beside image warping and registration in medical image analysis, a key problem in nearly all imaging applications is the reconstruction of high quality data from low resolution observations. This field, commonly referred to as "inverse problems", is very often concerned with the precise location of features such as point sources (modeled as Dirac masses) or sharp contours of objects (modeled as gradients being Dirac masses along curves). The underlying intuition behind these ideas is the so-called sparsity model (either of the data itself, its gradient, or other more complicated representations such as wavelets, curvelets, bandlets [122] and learned representation [153]).

The huge interest in these ideas started mostly from the introduction of convex methods to serve as proxy for these sparse regularizations. The most well known is the ℓ^1 norm introduced independently in imaging by Donoho and co-workers under the name "Basis Pursuit" [85] and in statistics by Tibshirani [144] under the name "Lasso". A more recent resurgence of this interest dates back to 10 years ago with the introduction of the so-called "compressed sensing" acquisition techniques [69], which make use of randomized forward operators and ℓ^1-type reconstruction.

2.4.2. Regularization over measure spaces.

However, the theoretical analysis of sparse reconstructions involving real-life acquisition operators (such as those found in seismic imaging, neuro-imaging, astro-physical imaging, etc.) is still mostly an open problem. A recent research direction, triggered by a paper of Candès and Fernandez-Granda [71], is to study directly the infinite dimensional problem of reconstruction of sparse measures (i.e. sum of Dirac masses) using the total variation of measures (not to be mistaken for the total variation of 2-D functions). Several works [70], [95], [92] have used this framework to provide theoretical performance guarantees by basically studying how the distance between neighboring spikes impacts noise stability.

Segmentation input output Zooming input output

Figure 3. Two example of application of the total variation regularization of functions. Left: image segmentation into homogeneous color regions. Right: image zooming (increasing the number of pixels while keeping the edges sharp).

2.4.3. Low complexity regularization and partial smoothness.

In image processing, one of the most popular methods is the total variation regularization [139], [62]. It favors low-complexity images that are piecewise constant, see Figure 3 for some examples on how to solve some image processing problems. Beside applications in image processing, sparsity-related ideas also had a deep impact in statistics [144] and machine learning [29]. As a typical example, for applications to recommendation systems, it makes sense to consider sparsity of the singular values of matrices, which can be relaxed using the so-called nuclear norm (a.k.a. trace norm) [30]. The underlying methodology is to make use of low-complexity regularization models, which turns out to be equivalent to the use of partly-smooth regularization functionals [115], [146] enforcing the solution to belong to a low-dimensional manifold.

2.5. Mokaplan unified point of view

The dynamical formulation of optimal transport creates a link between optimal transport and geodesics on diffeomorphisms groups. This formal link has at least two strong implications that MOKAPLAN will elaborate on: (i) the development of novel models that bridge the gap between these two fields ; (ii) the introduction of novel fast numerical solvers based on ideas from both non-smooth optimization techniques and Bregman metrics, as highlighted in Section 3.2.3.

In a similar line of ideas, we believe a unified approach is needed to tackle both sparse regularization in imaging and various generalized OT problems. Both require to solve related non-smooth and large scale optimization problems. Ideas from proximal optimization has proved crucial to address problems in both fields (see for instance [35], [137]). Transportation metrics are also the correct way to compare and regularize variational problems that arise in image processing (see for instance the Radon inversion method proposed in [39]) and machine learning (see [90]). This unity in term of numerical methods is once again at the core of Section 3.2.3.

3. Research Program

3.1. Modeling and Analysis

The first layer of methodological tools developed by our team is a set of theoretical continuous models that aim at formalizing the problems studied in the applications. These theoretical findings will also pave the way to efficient numerical solvers that are detailed in Section 3.2.

3.1.1. Static Optimal Transport and Generalizations

3.1.1.1. Convexity constraint and Principal Agent problem in Economics.

(*Participants:* G. Carlier, J-D. Benamou, V. Duval, Xavier Dupuis (LUISS Guido Carli University, Roma)) The principal agent problem plays a distinguished role in the literature on asymmetric information and contract theory (with important contributions from several Nobel prizes such as Mirrlees, Myerson or Spence) and it has many important applications in optimal taxation, insurance, nonlinear pricing. The typical problem consists in finding a cost minimizing strategy for a monopolist facing a population of agents who have an unobservable characteristic, the principal therefore has to take into account the so-called incentive compatibilty constraint which is very similar to the cyclical monotonicity condition which characterizes optimal transport plans. In a special case, Rochet and Choné [138] reformulated the problem as a variational problem subject to a convexity constraint. For more general models, and using ideas from Optimal Transportation, Carlier [73] considered the more general c-convexity constraint and proved a general existence result. Using the formulation of [73] McCann, Figalli and Kim [96] gave conditions under which the principal agent problem can be written as an infinite dimensional convex variational problem. The important results of [96] are intimately connected to the regularity theory for optimal transport and showed that there is some hope to numerically solve the principal-agent problem for general utility functions.

Our expertise: We have already contributed to the numerical resolution of the Principal Agent problem in the case of the convexity constraint, see [78], [128], [125].

Goals: So far, the mathematical PA model can be numerically solved for simple utility functions. A Bregman approach inspired by [39] is currently being developed [76] for more general functions. It would be extremely useful as a complement to the theoretical analysis. A new semi-Discrete Geometric approach is also investigated where the method reduces to non-convex polynomial optimization.

3.1.1.2. Optimal transport and conditional constraints in statistics and finance.

(*Participants:* G. Carlier, J-D. Benamou, G. Peyré) A challenging branch of emerging generalizations of Optimal Transportation arising in *economics, statistics and finance* concerns Optimal Transportation with *conditional* constraints. The *martingale optimal transport* [33], [101] which appears naturally in mathematical finance aims at computing robust bounds on option prices as the value of an optimal transport problem where not only the marginals are fixed but the coupling should be the law of a martingale, since it represents the prices of the underlying asset under the risk-neutral probability at the different dates. Note that as soon as more than two dates are involved, we are facing a multimarginal problem.

Our expertise: Our team has a deep expertise on the topic of OT and its generalization, including many already existing collaboration between its members, see for instance [39], [44], [37] for some representative recent collaborative publications.

Goals: This is a non trivial extension of Optimal Transportation theory and MOKAPLAN will develop numerical methods (in the spirit of entropic regularization) to address it. A popular problem in statistics is the so-called quantile regression problem, recently Carlier, Chernozhukov and Galichon [74] used an Optimal Transportation approach to extend quantile regression to several dimensions. In this approach again, not only fixed marginals constraints are present but also constraints on conditional means. As in the martingale Optimal Transportation problem, one has to deal with an extra conditional constraint. The duality approach usually breaks down under such constraints and characterization of optimal couplings is a challenging task both from a theoretical and numerical viewpoint.

3.1.1.3. JKO gradient flows.

(*Participants:* G. Carlier, J-D. Benamou, M. Laborde, Q. Mérigot, V. Duval) The connection between the static and dynamic transportation problems (see Section 2.3) opens the door to many extensions, most notably by leveraging the use of gradient flows in metric spaces. The flow with respect to the transportation distance has been introduced by Jordan-Kindelherer-Otto (JKO) [108] and provides a variational formulation of many linear and non-linear diffusion equations. The prototypical example is the Fokker Planck equation. We will explore this formalism to study new variational problems over probability spaces, and also to derive innovative numerical solvers. The JKO scheme has been very successfully used to study evolution equations that have the structure of a gradient flow in the Wasserstein space. Indeed many important PDEs have this structure: the Fokker-Planck equation (as was first considered by [108]), the porous medium equations, the granular media equation, just to give a few examples. It also finds application in image processing [61]. Figure 4 shows examples of gradient flows.

Our expertise: There is an ongoing collaboration between the team members on the theoretical and numerical analysis of gradient flows.

Goals: We apply and extend our research on JKO numerical methods to treat various extensions:

- Wasserstein gradient flows with a non displacement convex energy (as in the parabolic-elliptic Keller-Segel chemotaxis model [80])
- systems of evolution equations which can be written as gradient flows of some energy on a product space (possibly mixing the Wasserstein and L^2 structures) : multi-species models or the parabolic-parabolic Keller-Segel model [48]
- perturbation of gradient flows: multi-species or kinetic models are not gradient flows, but may be viewed as a perturbation of Wasserstein gradient flows, we shall therefore investigate convergence of splitting methods for such equations or systems.

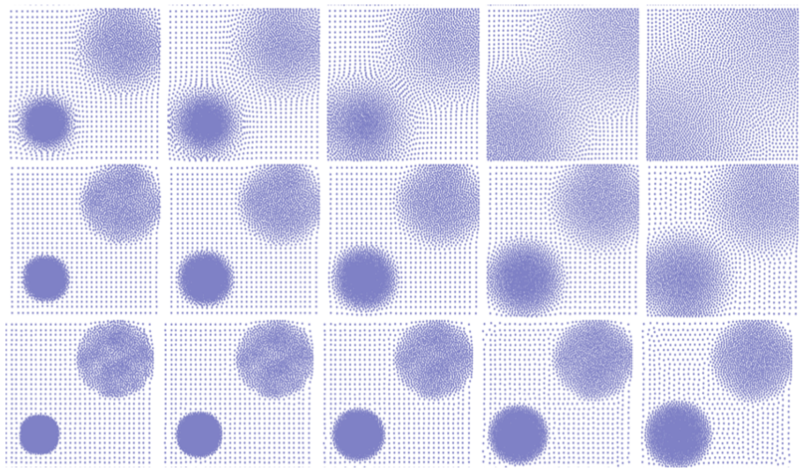

Figure 4. Example of non-linear diffusion equations solved with a JKO flow [40]. The horizontal axis shows the time evolution minimizing the functional $\int \frac{\rho^\alpha}{\alpha-1}$ on the density ρ (discretized here using point clouds, i.e. sum of Diracs' with equal mass). Each row shows a different value of $\alpha = (0.6, 2, 3)$

3.1.1.4. From networks to continuum congestion models.

(*Participants:* G. Carlier, J-D. Benamou, G. Peyré) Congested transport theory in the discrete framework of networks has received a lot of attention since the 50's starting with the seminal work of Wardrop. A few years later, Beckmann proved that equilibria are characterized as solution of a convex minimization problem. However, this minimization problem involves one flow variable per path on the network, its dimension thus quickly becomes too large in practice. An alternative, is to consider continuous in space models of congested optimal transport as was done in [77] which leads to very degenerate PDEs [53].

Our expertise: MOKAPLAN members have contributed a lot to the analysis of congested transport problems and to optimization problems with respect to a metric which can be attacked numerically by fast marching methods [44].

Goals: The case of general networks/anisotropies is still not well understood, general Γ-convergence results will be investigated as well as a detailed analysis of the corresponding PDEs and numerical methods to solve them. Benamou and Carlier already studied numerically some of these PDEs by an augmented Lagrangian method see figure 5. Note that these class of problems share important similarities with metric learning problem in machine learning, detailed below.

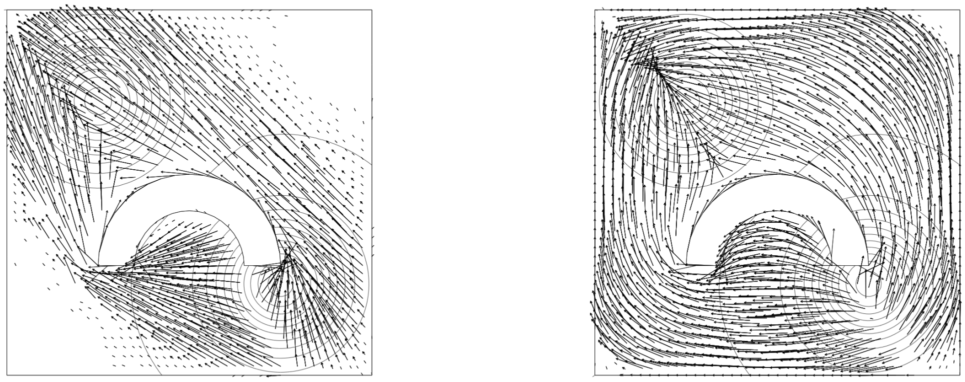

Figure 5. Monge and Wardrop flows of mass around an obstacle [37]. the source/target mass is represented by the level curves. Left : no congestion, Right : congestion.

3.1.2. Diffeomorphisms and Dynamical Transport

3.1.2.1. Growth Models for Dynamical Optimal Transport.

(*Participants:* F-X. Vialard, J-D. Benamou, G. Peyré, L. Chizat) A major issue with the standard dynamical formulation of OT is that it does not allow for variation of mass during the evolution, which is required when tackling medical imaging applications such as tumor growth modeling [64] or tracking elastic organ movements [142]. Previous attempts [119], [135] to introduce a source term in the evolution typically lead to mass teleportation (propagation of mass with infinite speed), which is not always satisfactory.

Our expertise: Our team has already established key contributions both to connect OT to fluid dynamics [35] and to define geodesic metrics on the space of shapes and diffeomorphisms [84].

Goals: Lenaic Chizat's PhD thesis aims at bridging the gap between dynamical OT formulation, and LDDDM diffeomorphisms models (see Section 2.3). This will lead to biologically-plausible evolution models that are both more tractable numerically than LDDM competitors, and benefit from strong theoretical guarantees associated to properties of OT.

3.1.2.2. Mean-field games.

(*Participants:* G. Carlier, J-D. Benamou) The Optimal Transportation Computational Fluid Dynamics (CFD) formulation is a limit case of variational Mean-Field Games (MFGs), a new branch of game theory recently developed by J-M. Lasry and P-L. Lions [112] with an extremely wide range of potential applications [104]. Non-smooth proximal optimization methods used successfully for the Optimal Transportation can be used in the case of deterministic MFGs with singular data and/or potentials [38]. They provide a robust treatment of the positivity constraint on the density of players.
Our expertise: J.-D. Benamou has pioneered with Brenier the CFD approach to Optimal Transportation. Regarding MFGs, on the numerical side, our team has already worked on the use of augmented Lagrangian methods in MFGs [37] and on the analytical side [72] has explored rigorously the optimality system for a singular CFD problem similar to the MFG system.
Goals: We will work on the extension to stochastic MFGs. It leads to non-trivial numerical difficulties already pointed out in [26].

3.1.2.3. Macroscopic Crowd motion, congestion and equilibria.

(*Participants:* G. Carlier, J-D. Benamou, Q. Mérigot, F. Santambrogio (U. Paris-Sud), Y. Achdou (Univ. Paris 7), R. Andreev (Univ. Paris 7)) Many models from PDEs and fluid mechanics have been used to give a description of *people or vehicles moving in a congested environment*. These models have to be classified according to the dimension (1D model are mostly used for cars on traffic networks, while 2-D models are most suitable for pedestrians), to the congestion effects ("soft" congestion standing for the phenomenon where high densities slow down the movement, "hard" congestion for the sudden effects when contacts occur, or a certain threshold is attained), and to the possible rationality of the agents Maury et al [123] recently developed a theory for 2D hard congestion models without rationality, first in a discrete and then in a continuous framework. This model produces a PDE that is difficult to attack with usual PDE methods, but has been successfully studied via Optimal Transportation techniques again related to the JKO gradient flow paradigm. Another possibility to model crowd motion is to use the mean field game approach of Lions and Lasry which limits of Nash equilibria when the number of players is large. This also gives macroscopic models where congestion may appear but this time a global equilibrium strategy is modelled rather than local optimisation by players like in the JKO approach. Numerical methods are starting to be available, see for instance [26], [60].
Our expertise: We have developed numerical methods to tackle both the JKO approach and the MFG approach. The Augmented Lagrangian (proximal) numerical method can actually be applied to both models [37], JKO and deterministic MFGs.
Goals: We want to extend our numerical approach to more realistic congestion model where the speed of agents depends on the density, see Figure 6 for preliminary results. Comparison with different numerical approaches will also be performed inside the ANR ISOTACE. Extension of the Augmented Lagrangian approach to Stochastic MFG will be studied.

3.1.2.4. Diffeomorphic image matching.

(*Participants:* F-X. Vialard, G. Peyré, B. Schmitzer, L. Chizat) Diffeomorphic image registration is widely used in medical image analysis. This class of problems can be seen as the computation of a generalized optimal transport, where the optimal path is a geodesic on a group of diffeomorphisms. The major difference between the two approaches being that optimal transport leads to non smooth optimal maps in general, which is however compulsory in diffeomorphic image matching. In contrast, optimal transport enjoys a convex variational formulation whereas in LDDMM the minimization problem is non convex.
Our expertise: F-X. Vialard is an expert of diffeomorphic image matching (LDDMM) [147], [59], [145]. Our team has already studied flows and geodesics over non-Riemannian shape spaces, which allows for piecewise smooth deformations [84].

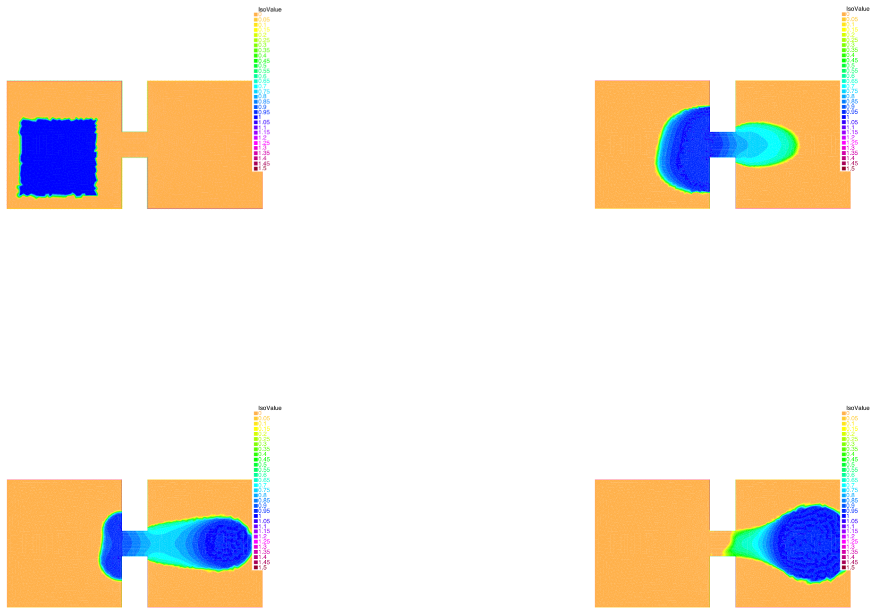

Figure 6. Example of crowd congestion with density dependent speed. The macroscopic density, at 4 different times, of people forced to exit from one room towards a meeting point in a second room.

Goals: Our aim consists in bridging the gap between standard optimal transport and diffeomorphic methods by building new diffeomorphic matching variational formulations that are convex (geometric obstructions might however appear). A related perspective is the development of new registration/transport models in a Lagrangian framework, in the spirit of [141], [142] to obtain more meaningful statistics on longitudinal studies.

Diffeomorphic matching consists in the minimization of a functional that is a sum of a deformation cost and a similarity measure. The choice of the similarity measure is as important as the deformation cost. It is often chosen as a norm on a Hilbert space such as functions, currents or varifolds. From a Bayesian perspective, these similarity measures are related to the noise model on the observed data which is of geometric nature and it is not taken into account when using Hilbert norms. Optimal transport fidelity have been used in the context of signal and image denoising [114], and it is an important question to extends these approach to registration problems. Therefore, we propose to develop similarity measures that are geometric and computationally very efficient using entropic regularization of optimal transport.

Our approach is to use a regularized optimal transport to design new similarity measures on all of those Hilbert spaces. Understanding the precise connections between the evolution of shapes and probability distributions will be investigated to cross-fertilize both fields by developing novel transportation metrics and diffeomorphic shape flows.

The corresponding numerical schemes are however computationally very costly. Leveraging our understanding of the dynamic optimal transport problem and its numerical resolution, we propose to develop new algorithms. These algorithms will use the smoothness of the Riemannian metric to improve both accuracy and speed, using for instance higher order minimization algorithm on (infinite dimensional) manifolds.

3.1.2.5. Metric learning and parallel transport for statistical applications.

(*Participants:* F-X. Vialard, G. Peyré, B. Schmitzer, L. Chizat) The LDDMM framework has been advocated to enable statistics on the space of shapes or images that benefit from the estimation of the deformation. The statistical results of it strongly depend on the choice of the Riemannian metric. A possible direction consists in learning the right invariant Riemannian metric as done in [148] where a correlation matrix (Figure 7) is learnt which represents the covariance matrix of the deformation fields for a given population of shapes. In the same direction, a question of emerging interest in medical imaging is the analysis of time sequence of shapes (called longitudinal analysis) for early diagnosis of disease, for instance [97]. A key question is the inter subject comparison of the organ evolution which is usually done by transport of the time evolution in a common coordinate system via parallel transport or other more basic methods. Once again, the statistical results (Figure 8) strongly depend on the choice of the metric or more generally on the connection that defines parallel transport.
Our expertise: Our team has already studied statistics on longitudinal evolutions in [97], [98].
Goals: Developing higher order numerical schemes for parallel transport (only low order schemes are available at the moment) and developing variational models to learn the metric or the connections for improving statistical results.

3.1.3. Sparsity in Imaging

3.1.3.1. Inverse problems over measures spaces.

(*Participants:* G. Peyré, V. Duval, C. Poon, Q. Denoyelle) As detailed in Section 2.4, popular methods for regularizing inverse problems in imaging make use of variational analysis over infinite-dimensional (typically non-reflexive) Banach spaces, such as Radon measures or bounded variation functions.
Our expertise: We have recently shown in [146] how – in the finite dimensional case – the non-smoothness of the functionals at stake is crucial to enforce the emergence of geometrical structures (edges in images or fractures in physical materials [49]) for discrete (finite dimensional) problems. We extended this result in a simple infinite dimensional setting, namely sparse regularization of Radon measures for deconvolution [92]. A deep understanding of those continuous inverse problems is crucial to analyze the behavior of their discrete counterparts, and in [93] we have taken advantage of this understanding to develop a fine analysis of the artifacts induced by discrete (*i.e.* which involve grids) deconvolution models. These works are also closely

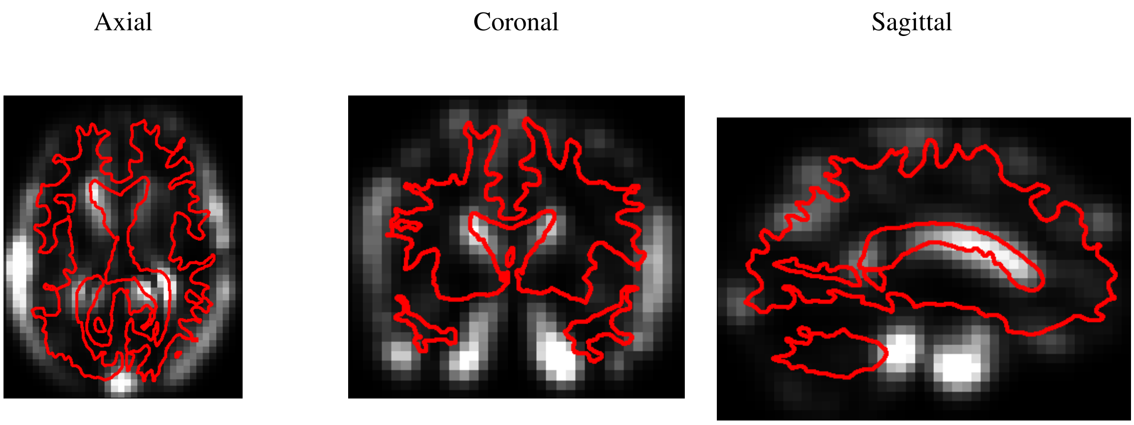

Figure 7. Learning Riemannian metrics in diffeomorphic image matching to capture the brain variability: a diagonal operator that encodes the Riemannian metric is learnt on a template brain out of a collection of brain images. The values of the diagonal operator are shown in greyscale. The red curves represent the boundary between white and grey matter. For more details, we refer the reader to [148], which was a first step towards designing effective and robust metric learning algorithms.

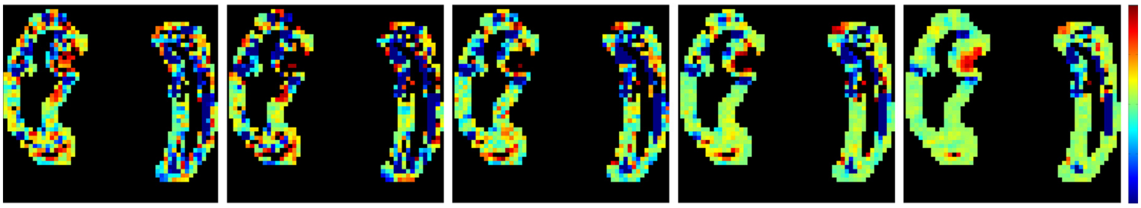

Figure 8. Statistics on initial momenta: In [97], we compared several intersubject transport methodologies to perform statistics on longitudinal evolutions. These longitudinal evolutions are represented by an initial velocity field on the shapes boundaries and these velocity fields are then compared using logistic regression methods that are regularized. The four pictures represent different regularization methods such as L^2, H^1 and regularization including a sparsity prior such as Lasso, Fused Lasso and TV.

related to the problem of limit analysis and yield design in mechanical plasticity, see [75], [49] for an existing collaboration between MOKAPLAN's team members.

Goals: A current major front of research in the mathematical analysis of inverse problems is to extend these results for more complicated infinite dimensional signal and image models, such as for instance the set of piecewise regular functions. The key bottleneck is that, contrary to sparse measures (which are finite sums of Dirac masses), here the objects to recover (smooth edge curves) are not parameterized by a finite number of degrees of freedom. The relevant previous work in this direction are the fundamental results of Chambolle, Caselles and co-workers [34], [28], [81]. They however only deal with the specific case where there is no degradation operator and no noise in the observations. We believe that adapting these approaches using our construction of vanishing derivative pre-certificate [92] could lead to a solution to these theoretical questions.

3.1.3.2. Sub-Riemannian diffusions.

(*Participants:* G. Peyré, J-M. Mirebeau, D. Prandi) Modeling and processing natural images require to take into account their geometry through anisotropic diffusion operators, in order to denoise and enhance directional features such as edges and textures [134], [94]. This requirement is also at the heart of recently proposed models of cortical processing [133]. A mathematical model for these processing is diffusion on sub-Riemanian manifold. These methods assume a fixed, usually linear, mapping from the 2-D image to a lifted function defined on the product of space and orientation (which in turn is equipped with a sub-Riemannian manifold structure).

Our expertise: J-M. Mirebeau is an expert in the discretization of highly anisotropic diffusions through the use of locally adaptive computational stencils [126], [94]. G. Peyré has done several contributions on the definition of geometric wavelets transform and directional texture models, see for instance [134]. Dario Prandi has recently applied methods from sub-Riemannian geometry to image restoration [51].

Goals: A first aspect of this work is to study non-linear, data-adaptive, lifting from the image to the space/orientation domain. This mapping will be implicitly defined as the solution of a convex variational problem. This will open both theoretical questions (existence of a solution and its geometrical properties, when the image to recover is piecewise regular) and numerical ones (how to provide a faithful discretization and fast second order Newton-like solvers). A second aspect of this task is to study the implication of these models for biological vision, in a collaboration with the UNIC Laboratory (directed by Yves Fregnac), located in Gif-sur-Yvette. In particular, the study of the geometry of singular vectors (or "ground states" using the terminology of [45]) of the non-linear sub-Riemannian diffusion operators is highly relevant from a biological modeling point of view.

3.1.3.3. Sparse reconstruction from scanner data.

(*Participants:* G. Peyré, V. Duval, C. Poon) Scanner data acquisition is mathematically modeled as a (sub-sampled) Radon transform [105]. It is a difficult inverse problem because the Radon transform is ill-posed and the set of observations is often aggressively sub-sampled and noisy [140]. Typical approaches [111] try to recover piecewise smooth solutions in order to recover precisely the position of the organ being imaged. There is however a very poor understanding of the actual performance of these methods, and little is known on how to enhance the recovery.

Our expertise: We have obtained a good understanding of the performance of inverse problem regularization on *compact* domains for pointwise sources localization [92].

Goals: We aim at extending the theoretical performance analysis obtained for sparse measures [92] to the set of piecewise regular 2-D and 3-D functions. Some interesting previous work of C. Poon et al [136] (C. Poon is currently a postdoc in MOKAPLAN) have tackled related questions in the field of variable Fourier sampling for compressed sensing application (which is a toy model for fMRI imaging). These approaches are however not directly applicable to Radon sampling, and require some non-trivial adaptations. We also aim at better exploring the connection of these methods with optimal-transport based fidelity terms such as those introduced in [25].

3.1.3.4. Tumor growth modeling in medical image analysis.

(*Participants:* G. Peyré, F-X. Vialard, J-D. Benamou, L. Chizat) Some applications in medical image analysis require to track shapes whose evolution is governed by a growth process. A typical example is tumor growth, where the evolution depends on some typically unknown but meaningful parameters that need to be estimated. There exist well-established mathematical models [64], [132] of non-linear diffusions that take into account recently biologically observed property of tumors. Some related optimal transport models with mass variations have also recently been proposed [121], which are connected to so-called metamorphoses models in the LDDMM framework [46].

Our expertise: Our team has a strong experience on both dynamical optimal transport models and diffeomorphic matching methods (see Section 3.1.2).

Goals: The close connection between tumor growth models [64], [132] and gradient flows for (possibly non-Euclidean) Wasserstein metrics (see Section 3.1.2) makes the application of the numerical methods we develop particularly appealing to tackle large scale forward tumor evolution simulation. A significant departure from the classical OT-based convex models is however required. The final problem we wish to solve is the backward (inverse) problem of estimating tumor parameters from noisy and partial observations. This also requires to set-up a meaningful and robust data fidelity term, which can be for instance a generalized optimal transport metric.

3.2. Numerical Tools

The above continuous models require a careful discretization, so that the fundamental properties of the models are transferred to the discrete setting. Our team aims at developing innovative discretization schemes as well as associated fast numerical solvers, that can deal with the geometric complexity of the variational problems studied in the applications. This will ensure that the discrete solution is correct and converges to the solution of the continuous model within a guaranteed precision. We give below examples for which a careful mathematical analysis of the continuous to discrete model is essential, and where dedicated non-smooth optimization solvers are required.

3.2.1. Geometric Discretization Schemes

3.2.1.1. Discretizing the cone of convex constraints.

(*Participants:* J-D. Benamou, G. Carlier, J-M. Mirebeau, Q. Mérigot) Optimal transportation models as well as continuous models in economics can be formulated as infinite dimensional convex variational problems with the constraint that the solution belongs to the cone of convex functions. Discretizing this constraint is however a tricky problem, and usual finite element discretizations fail to converge.

Our expertise: Our team is currently investigating new discretizations, see in particular the recent proposal [43] for the Monge-Ampère equation and [125] for general non-linear variational problems. Both offer convergence guarantees and are amenable to fast numerical resolution techniques such as Newton solvers. Since [43] explaining how to treat efficiently and in full generality Transport Boundary Conditions for Monge-Ampère, this is a promising fast and new approach to compute Optimal Transportation viscosity solutions. A monotone scheme is needed. One is based on Froese Oberman work [100], a new different and more accurate approach has been proposed by Mirebeau, Benamou and Collino [41]. As shown in [86], discretizing the constraint for a continuous function to be convex is not trivial. Our group has largely contributed to solve this problem with G. Carlier [78], Quentin Mérigot [128] and J-M. Mirebeau [125]. This problem is connected to the construction of monotone schemes for the Monge-Ampère equation.

Goals: The current available methods are 2-D. They need to be optimized and parallelized. A non-trivial extension to 3-D is necessary for many applications. The notion of c-convexity appears in optimal transport for generalized displacement costs. How to construct an adapted discretization with “good” numerical properties is however an open problem.

3.2.1.2. *Numerical JKO gradient flows.*

(*Participants:* J-D. Benamou, G. Carlier, J-M. Mirebeau, G. Peyré, Q. Mérigot) As detailed in Section 2.3, gradient Flows for the Wasserstein metric (aka JKO gradient flows [108]) provides a variational formulation of many non-linear diffusion equations. They also open the way to novel discretization schemes. From a computational point, although the JKO scheme is constructive (it is based on the implicit Euler scheme), it has not been very much used in practice numerically because the Wasserstein term is difficult to handle (except in dimension one).
Our expertise:

Solving one step of a JKO gradient flow is similar to solving an Optimal transport problem. A geometrical a discretization of the Monge-Ampère operator approach has been proposed by Mérigot, Carlier, Oudet and Benamou in [40] see Figure 4. The Gamma convergence of the discretisation (in space) has been proved.
Goals: We are also investigating the application of other numerical approaches to Optimal Transport to JKO gradient flows either based on the CFD formulation or on the entropic regularization of the Monge-Kantorovich problem (see section 3.2.3). An in-depth study and comparison of all these methods will be necessary.

3.2.2. Sparse Discretization and Optimization

3.2.2.1. *From discrete to continuous sparse regularization and transport.*

(*Participants:* V. Duval, G. Peyré, G. Carlier, Jalal Fadili (ENSICaen), Jérôme Malick (CNRS, Univ. Grenoble)) While pervasive in the numerical analysis community, the problem of discretization and Γ-convergence from discrete to continuous is surprisingly over-looked in imaging sciences. To the best of our knowledge, our recent work [92], [93] is the first to give a rigorous answer to the transition from discrete to continuous in the case of the spike deconvolution problem. Similar problems of Γ-convergence are progressively being investigated in the optimal transport community, see in particular [79].
Our expertise: We have provided the first results on the discrete-to-continous convergence in both sparse regularization variational problems [92], [93] and the static formulation of OT and Wasserstein barycenters [79]
Goals: In a collaboration with Jérôme Malick (Inria Grenoble), our first goal is to generalize the result of [92] to generic partly-smooth convex regularizers routinely used in imaging science and machine learning, a prototypal example being the nuclear norm (see [146] for a review of this class of functionals). Our second goal is to extend the results of [79] to the novel class of entropic discretization schemes we have proposed [39], to lay out the theoretical foundation of these ground-breaking numerical schemes.

3.2.2.2. *Polynomial optimization for grid-free regularization.*

(*Participants:* G. Peyré, V. Duval, I. Waldspurger) There has been a recent spark of attention of the imaging community on so-called “grid free” methods, where one tries to directly tackle the infinite dimensional recovery problem over the space of measures, see for instance [71], [92]. The general idea is that if the range of the imaging operator is finite dimensional, the associated dual optimization problem is also finite dimensional (for deconvolution, it corresponds to optimization over the set of trigonometric polynomials).
Our expertise: We have provided in [92] a sharp analysis of the support recovery property of this class of methods for the case of sparse spikes deconvolution.
Goals: A key bottleneck of these approaches is that, while being finite dimensional, the dual problem necessitates to handle a constraint of polynomial positivity, which is notoriously difficult to manipulate (except in the very particular case of 1-D problems, which is the one exposed in [71]). A possible, but very costly, methodology is to ressort to Lasserre’s SDP representation hierarchy [113]. We will make use of these approaches and study how restricting the level of the hierarchy (to obtain fast algorithms) impacts the recovery performances (since this corresponds to only computing approximate solutions). We will pay a particular attention to the recovery of 2-D piecewise constant functions (the so-called total variation of functions regularization [139]), see Figure 3 for some illustrative applications of this method.

3.2.3. First Order Proximal Schemes

3.2.3.1. L^2 proximal methods.

(*Participants:* G. Peyré, J-D. Benamou, G. Carlier, Jalal Fadili (ENSICaen)) Both sparse regularization problems in imaging (see Section 2.4) and dynamical optimal transport (see Section 2.3) are instances of large scale, highly structured, non-smooth convex optimization problems. First order proximal splitting optimization algorithms have recently gained lots of interest for these applications because they are the only ones capable of scaling to giga-pixel discretizations of images and volumes and at the same time handling non-smooth objective functions. They have been successfully applied to optimal transport [35], [129], congested optimal transport [63] and to sparse regularizations (see for instance [137] and the references therein).
Our expertise: The pioneering work of our team has shown how these proximal solvers can be used to tackle the dynamical optimal transport problem [35], see also [129]. We have also recently developed new proximal schemes that can cope with non-smooth composite objectives functions [137].
Goals: We aim at extending these solvers to a wider class of variational problems, most notably optimization under divergence constraints [37]. Another subject we are investigating is the extension of these solvers to both non-smooth and non-convex objective functionals, which are mandatory to handle more general transportation problems and novel imaging regularization penalties.

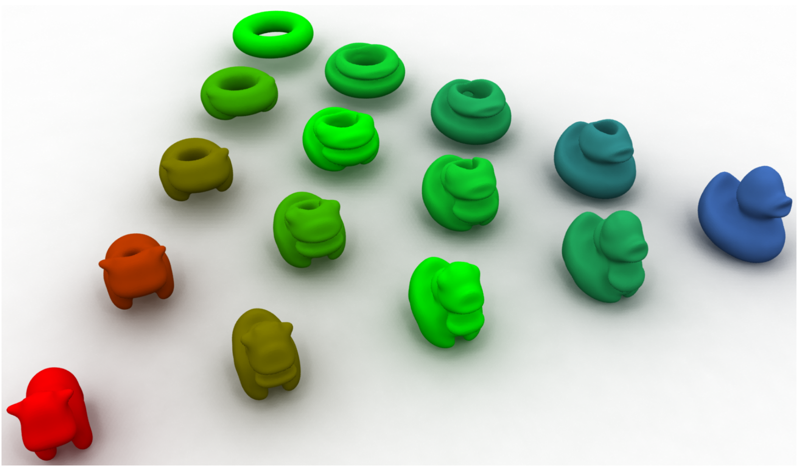

Figure 9. Example of barycenter between shapes computed using optimal transport barycenters of the uniform densities inside the 3 extremal shapes, computed as detailed in [143]. Note that the barycenters are not in general uniform distributions, and we display them as the surface defined by a suitable level-set of the density.

3.2.3.2. Bregman proximal methods.

(*Participants:* G. Peyré G. Carlier, L. Nenna, J-D. Benamou, L. Nenna, Marco Cuturi (Kyoto Univ.)) The entropic regularization of the Kantorovich linear program for OT has been shown to be surprisingly simple and efficient, in particular for applications in machine learning [90]. As shown in [39], this is a special instance of the general method of Bregman iterations, which is also a particular instance of first order proximal schemes according to the Kullback-Leibler divergence.

Our expertise: We have recently [39] shown how Bregman projections [54] and Dykstra algorithm [31] offer a generic optimization framework to solve a variety of generalized OT problems. Carlier and Dupuis [76] have designed a new method based on alternate Dykstra projections and applied it to the *principal-agent problem* in microeconomics. We have applied this method in computer graphics in a paper accepted in SIGGRAPH 2015 [143]. Figure 9 shows the potential of our approach to handle giga-voxel datasets: the input volumetric densities are discretized on a 100^3 computational grid.
Goals: Following some recent works (see in particular [83]) we first aim at studying primal-dual optimization schemes according to Bregman divergences (that would go much beyond gradient descent and iterative projections), in order to offer a versatile and very effective framework to solve variational problems involving OT terms. We then also aim at extending the scope of usage of this method to applications in quantum mechanics (Density Functional Theory, see [87]) and fluid dynamics (Brenier's weak solutions of the incompressible Euler equation, see [55]). The computational challenge is that realistic physical examples are of a huge size not only because of the space discretization of one marginal but also because of the large number of marginals involved (for incompressible Euler the number of marginals equals the number of time steps).

4. New Software and Platforms

4.1. ALG2

FUNCTIONAL DESCRIPTION: ALG2 for Monge Mean-Field Games, Monge problem and Variational problems under divergence constraint. A generalisation of the ALG2 algorithm has been implemented in FreeFem++.

- Contact: Jean-David Benamou
- URL: https://team.inria.fr/mokaplan/augmented-lagrangian-simulations/

4.2. Mokabajour

FUNCTIONAL DESCRIPTION: We design a software resolving the following inverse problem: define the shape of a mirror which reflects the light from a source to a defined target, distribution and support of densities being prescribed. Classical applications include the conception of solar oven, public lightning, car headlights… Mathematical modeling of this problem, related to the optimal transport theory, takes the form of a nonlinear Monge-Ampere type PDE. The numerical resolution of these models remained until recently a largely open problem. MOKABAJOUR project aims to develop, using algorithms invented especially at Inria and LJK, a reflector design software more efficient than geometrical methods used so far. The final step is to realize and physically test prototype reflectors.

- Participants: Boris Thibert, Jean-David Benamou and Quentin Mérigot
- Contact: Jean-David Benamou
- URL: https://project.inria.fr/mokabajour/

5. New Results

5.1. An augmented Lagrangian approach to Wasserstein gradient flows and applications

Benamou, Jean-David and Carlier, Guillaume and Laborde, Maxime. In [2] : Taking advantage of the Benamou-Brenier dynamic formulation of optimal transport, we propose a convex formulation for each step of the JKO scheme for Wasserstein gradient flows which can be attacked by an augmented Lagrangian method which we call the ALG2-JKO scheme. We test the algorithm in particular on the porous medium equation. We also consider a semi implicit variant which enables us to treat nonlocal interactions as well as systems of interacting species. Regarding systems, we can also use the ALG2-JKO scheme for the simulation of crowd motion models with several species.

5.2. An entropy minimization approach to second-order variational mean-field games

Benamou, Jean-David and Carlier, Guillaume and Marino, Simone Di and Nenna, Luca. In [18] : We propose an entropy minimization viewpoint on variational meanfield games with diffusion and quadratic Hamiltonian. We carefully analyze the time-discretization of such problems, establish Γ-convergence results as the time step vanishes and propose an efficient algorithm relying on this entropic interpretation as well as on the Sinkhorn scaling algorithm.

5.3. Minimal convex extensions and finite difference discretization of the quadratic Monge-Kantorovich problem

Benamou, Jean-David and Duval, Vincent. In [3] : We propose an adaptation of the MA-LBR scheme to the Monge-Ampère equation with second boundary value condition, provided the target is a convex set. This yields a fast adaptive method to numerically solve the Optimal Transport problem between two absolutely continuous measures, the second of which has convex support. The proposed numerical method actually captures a specific Brenier solution which is minimal in some sense. We prove the convergence of the method as the grid stepsize vanishes and we show with numerical experiments that it is able to reproduce subtle properties of the Optimal Transport problem

We propose an entropy minimization viewpoint on variational meanfield games with diffusion and quadratic Hamiltonian. We carefully analyze the time-discretization of such problems, establish Γ-convergence results as the time step vanishes and propose an efficient algorithm relying on this entropic interpretation as well as on the Sinkhorn scaling algorithm.

5.4. Sparse analysis methods for Mesoscale Convective Systems (MCS)

Jean-Baptiste Courbot, Vincent Duval and Bernard Legras. In [20] : Mesoscale Convective Systems (MCS) are organized cloud systems which constitute the major part of the high cloud cover in the tropical region, where they can reach sizes of up to 1000 km in diameter. Under some favorable circumstances, they can eventually organize as tropical cyclones. Due to the complexity and the huge amount of available information, there is a need for an automatic tool able to detect and keep track of MCS in time series of these images. In [20] we have proposed a new method for the tracking of those cloud systems in infrared satellite images.

5.5. Pareto optimality with transport costs

Bruno Nazaret, Xavier Bacon, Guillaume Carlier, Thomas Gallouët. Arrived in September 2019, B. Nazaret, with his Phd student X. Bacon, has started a collaboration with G. Carlier on a Pareto optimality model with transport costs. He also has initiated a study with T.O. Gallouët of of a first order proximal scheme for densities on Euclidean half-spaces, with the additional feature that the boundary moves while the mass still remains constant along the time.

5.6. An epigraphical approach to the representer theorem

Duval, Vincent. In [22] : Describing the solutions of inverse problems arising in signal or image processing is an important issue both for theoretical and numerical purposes. In [22], we have proposed a principle which describes the solutions to convex variational problems involving a finite number of measurements. We discuss its optimality on various problems concerning the recovery of Radon measures.

5.7. Approximate Optimal Designs for Multivariate Polynomial Regression

De Castro, Yohann et al. In [8] : We introduce a new approach aiming at computing approximate optimal designs for multivariate polynomial regressions on compact (semi-algebraic) design spaces. We use the moment-sum-of-squares hierarchy of semidefinite programming problems to solve numerically the approximate optimal design problem. The geometry of the design is recovered via semidefinite programming duality theory. This article shows that the hierarchy converges to the approximate optimal design as the order of the hierarchy increases. Furthermore, we provide a dual certificate ensuring finite convergence of the hierarchy and showing that the approximate optimal design can be computed numerically with our method. As a byproduct, we revisit the equivalence theorem of the experimental design theory: it is linked to the Christoffel polynomial and it characterizes finite convergence of the moment-sum-of-square hierarchies. Describing the solutions of inverse problems arising

5.8. Sparse Recovery from Extreme Eigenvalues Deviation Inequalities

De Castro, Yohann et al. In [7] : This article provides a new toolbox to derive sparse recovery guarantees. It is ususally referred to as "stable and robust sparse regression" (SRSR) ' from deviations on extreme singular values or extreme eigenvalues obtained in Random Matrix Theory. This work is based on Restricted Isometry Constants (RICs) which are a pivotal notion in Compressed Sensing and High-Dimensional Statistics as these constants finely assess how a linear operator is conditioned on the set of sparse vectors and hence how it performs in SRSR. While it is an open problem to construct deterministic matrices with apposite RICs, one can prove that such matrices exist using random matrices models. In this paper, we show upper bounds on RICs for Gaussian and Rademacher matrices using state-of-the-art deviation estimates on their extreme eigenvalues. This allows us to derive a lower bound on the probability of getting SRSR. One benefit of this paper is a direct and explicit derivation of upper bounds on RICs and lower bounds on SRSR from deviations on the extreme eigenvalues given by Random Matrix theory.

5.9. Some new Stein operators for product distributions

Mijoule G. et al. In [12] : We provide a general result for finding Stein operators for the product of two independent random variables whose Stein operators satisfy a certain assumption, extending a recent result of Gaunt, R. E., Mijoule, G. and Swan. This framework applies to non-centered normal and non-centered gamma random variables, as well as a general sub-family of the variance-gamma distributions. Curiously, there is an increase in complexity in the Stein operators for products of independent normals as one moves, for example, from centered to non-centered normals. As applications, we give a simple derivation of the characteristic function of the product of independent normals, and provide insight into why the probability density function of this distribution is much more complicated in the non-centered case than the centered case.

5.10. Simulation of multiphase porous media flows with minimizing movement and finite volume schemes

C. Cancès, T. O. Gallouët, M. Laborde, L. Monsaingeon. In [6]: the Wasserstein gradient flow structure of the PDE system governing multiphase flows in porous media was recently highlighted in [68]. The model can thus be approximated by means of the minimizing movement (or JKO) scheme. We solve the JKO scheme using the ALG2-JKO scheme proposed in [2]. The numerical results are compared to a classical upstream mobility Finite Volume scheme, for which strong stability properties can be established.

5.11. An unbalanced optimal transport splitting scheme for general advection-reaction-diffusion problems

T. O. Gallouët, M. Laborde, L. Monsaingeon. In [10] the authors show that unbalanced optimal transport provides a convenient framework to handle reaction and diffusion processes in a unified metric framework. We use a constructive method, alternating minimizing movements for the Wasserstein distance and for the

Fisher-Rao distance, and prove existence of weak solutions for general scalar reaction-diffusion-advection equations. We extend the approach to systems of multiple interacting species, and also consider an application to a very degenerate diffusion problem involving a Gamma-limit. Moreover, some numerical simulations are included.

5.12. An unbalanced optimal transport splitting scheme for general advection-reaction-diffusion problems

C .Cancès, T. O. Gallouët, G. Todeschi. We propose a variational finite volume scheme to approximate the solutions to Wasserstein gradient flows. The time discretization is based on an implicit linearization of the Wasserstein distance expressed thanks to Benamou-Brenier formula, whereas space discretization relies on upstream mobility two-point flux approximation finite volumes. Our scheme is based on a first discretize then optimize approach in order to preserve the variational structure of the continuous model at the discrete level. Our scheme can be applied to a wide range of energies, guarantees non-negativity of the discrete solutions as well as decay of the energy. We show that our scheme admits a unique solution whatever the convex energy involved in the continuous problem , and we prove its convergence in the case of the linear Fokker-Planck equation with positive initial density. Numerical illustrations show that it is first order accurate in both time and space, and robust with respect to both the energy and the initial profile.

5.13. Generalized compressible fluid flows and solutions of the Camassa-Holm variational model

T. O. Gallouët, A. Natale, F-X. Vialard. In [11] : The Camassa-Holm equation on a domain M $\in \mathbb{R}^d$, in one of its possible multi-dimensional generalizations, describes geodesics on the group of diffeomorphisms with respect to the H(div) metric. It has been recently reformulated as a geodesic equation for the L2 metric on a subgroup of the diffeomorphism group of the cone over M. We use such an interpretation to construct an analogue of Brenier's generalized incompressible Euler flows for the Camassa-Holm equation. This involves describing the fluid motion using probability measures on the space of paths on the cone, so that particles are allowed to split and cross. Differently from Brenier's model, however, we are also able to account for compressibility by employing an explicit probabilistic representation of the Jacobian of the flow map. We formulate the boundary value problem associated to the Camassa-Holm equation using such generalized flows. We prove existence of solutions and that, for short times, smooth solutions of the Camassa-Holm equations are the unique solutions of our model. We propose a numerical scheme to construct generalized solutions on the cone and present some numerical results illustrating the relation between the generalized Camassa-Holm and incompressible Euler solutions.

5.14. Metric completion of $\mathrm{Diff}([0,1])$ with the right-invariant H^1 metric

S. di Marino, A. Natale, R. Tahraoui, F-X. Vialard. In [21]: We consider the group of smooth increasing diffeomorphisms Diff on the unit interval endowed with the right-invariant H1 metric. We compute the metric completion of this space which appears to be the space of increasing maps of the unit interval with boundary conditions at 0 and 1. We compute the lower-semicontinuous envelope associated with the length minimizing geodesic variational problem. We discuss the Eulerian and Lagrangian formulation of this relaxation and we show that smooth solutions of the EPDiff equation are length minimizing for short times.

5.15. Embedding Camassa-Holm equations in incompressible Euler

A. Natale, F-X. Vialard. In [13] : In this article, we show how to embed the so-called CH2 equations into the geodesic flow of the $H(\mathrm{div})$ metric in 2D, which, itself, can be embedded in the incompressible Euler equation of a non compact Riemannian manifold. The method consists in embedding the incompressible Euler equation with a potential term coming from classical mechanics into incompressible Euler of a manifold and seeing the CH2 equation as a particular case of such fluid dynamic equation.

5.16. Second order models for optimal transport and cubic splines on the Wasserstein space

J-D. Benamou, T. O. Gallouët, F-X. Vialard. In [4] : On the space of probability densities, we extend the Wasserstein geodesics to the case of higher-order interpolation such as cubic spline interpolation. After presenting the natural extension of cubic splines to the Wasserstein space, we propose a simpler approach based on the relaxation of the variational problem on the path space. We explore two different numerical approaches, one based on multi-marginal optimal transport and entropic regularization and the other based on semi-discrete optimal transport.

6. Partnerships and Cooperations

6.1. National Initiatives

6.1.1. ANR

V. Duval is the PI of the CIPRESSI (ANR JCJC) project. Its aim is to develop novel numerical schemes which respect the continuous nature of the variational problems in image or signal processing.
J-D. Benamou and G. Carlier are members of the ANR MFG (ANR-16-CE40-0015-01). Scientific topics of the project: Mean field analysis Analysis of the MFG systems and of the Master equation Numerical analysis Models and applications
J-D. Benamou G. Carlier F-X. Vialard and T. O. Gallouët are members of ANR MAGA (ANR-13-JS01-0007-01). The Monge-Ampère equation is a fully nonlinear elliptic equation, which plays a central role in geometry and in the theory of optimal transport. However, the singular and non-linear nature of the equation is a serious obstruction to its efficient numerical resolution. The first aim of the MAGA project is to study and to implement discretizations of optimal transport and Monge-Ampère equations which rely on tools from computational geometry (Laguerre diagrams). In a second step, these solvers will be applied to concrete problems from various fields involving optimal transport or Monge-Ampère equations such as computational physics: early universe reconstruction problem, congestion/incompressibility constraints economics: principal agent problems, geometry: variational problems over convex bodies, reflector and refractor design for non-imaging optics
T. O. Gallouët is member of the ANR GEOPOR (JCJC of C. Cancès) Scientific topic: geometrical approach, based on Wasserstein gradient flow, for multiphase flows in porous media. Theory and Numerics.
T. O. Gallouët is member of the ANR MESA (JCJC of M. Fathi) Scientific topic: Stein methods.

6.2. European Initiatives

6.2.1. FP7 & H2020 Projects

J-D. Benamou and Giorgi Rukhaia are members of ROMSOC ITN-EID.

6.3. International Initiatives

6.3.1. Inria International Partners

6.3.1.1. Informal International Partners

The team has strong ties with Technische Universität München, dept. of Math. (Profs. Daniel Matthes, Gero Friesecke, Bernhardt Schmitzer)

6.4. International Research Visitors

6.4.1. Visits of International Scientists

15-28/02 Visit of Prof. Yanir Rubinstein (University of Maryland).

6.4.2. Visits to International Teams

6.4.2.1. Research Stays Abroad

P. Pegon has been invited by Maria Colombo (Chair of Mathematical Analysis, Calculus of Variations and PDEs) at EPFL, Lausanne for 4 months (Feb-June 2019) to work on optimal and branched transport problems.

7. Dissemination

7.1. Promoting Scientific Activities

7.1.1. Scientific Events: Organisation

7.1.1.1. General Chair, Scientific Chair

- School, A Numerical introduction to optimal transport https://team.inria.fr/ecoleceainriaedf/en/
- School, Sparsity for Physics, Signal, Learning https://sparsity4psl.github.io/

7.1.1.2. Member of the Organizing Committees

I. Waldspurger has co-organized a session on phase retrieval at the SamTA conference 2019.

7.1.2. Scientific Events: Selection

7.1.2.1. Reviewer

V. Duval has reviewed several contributions for the GRETSI conference.

7.1.3. Journal

7.1.3.1. Member of the Editorial Boards

G. Carlier is on the Editorial board of J. Ec. Polytechnique, Mathematics and Fin. Econ., Applied Math. and Optim., Journal of Mathematical Analysis and Applications and J. Dyn. Games,

7.1.3.2. Reviewer - Reviewing Activities

I. Waldspurger has performed reviews for several journals (Applied and Computational Harmonic Analysis, Journal of Fourier Analysis and Applications, IEEE Transactions on Signal Processing and IEEE Transactions on Information Theory) and for the ICML conference.
V. Duval has performed reviews for SIAM Journal on Imaging Sciences, Information and Inference: a Journal of the IMA, Bulletin of the London Mathematical Society, Inverse Problems.
T. O. Gallouët has performed reviews for Analysis & PDE, Archive for Rational Mechanics and Analysis ARMA, Mathematics of computation (Math. of Comp.).
A. Natale has performed reviews for Journal of Scientific Computing (JOMP).
P. Pegon has performed reviews for Journal of Mathematical Analysis and Applications and Advances in Calculus of Variations.

7.1.4. Invited Talks

- J-D. Benamou, Seminar Monash University (Aug., Melbourne), Workshop on Monge-Ampère in honor of Prof. J. Urbas (Aug. , Kiama), Workshop on Optimal Transport and Optimal Patterns (Sept. , Edinburgh), Workshop on Risk in Finance (Oct. , Marseille)
- V. Duval: Workshop Variational methods and optimization in Imaging (IHP, Paris), Workshop on Signal and Image Analysis (MSIA'19, Burghausen, Germany), Two workshops at Conference on Applied Inverse Problems (AIP'19, Grenoble), Workshop Computational aspects of Geometry (U. Paul Sabatier, Toulouse), Conference Optimization on Measure Spaces (U. Paul Sabatier, Toulouse)
- G. Carlier Orsay (ANEDP), workshop Transport optimal Toulouse, ANR Shapo (Paris 7), Workshop Optimal Transport and Economics (Fields Institute, Toronto), Optimal transport in analysis and probability (Vienne, E. Schrodinger Institute).
- I. Waldspurger: Workshop on operator theoretic methods in dynamic data analysis and control (UCLA, États-Unis), workshop on imaging and machine learning (IHP, Paris), series of lectures on waves and imaging (ETH Zurich), séminaire parisien d'optimisation (IHP, Paris), workshop on computational aspects of geometry (U. Paul Sabatier, Toulouse)
- A. Natale: MAGA days (Université Paris-Sud, Paris), Journé de rentrée de l'equipe ANEDP (Université Paris-Sud, Paris), Rencontres Inria-LJLL en calcul scientifique (LJLL, Paris), Workshop on Variational Discretization for GFD (Fields Institute, Canada).
- P. Pegon: Séminaire d'analyse (EPFL, Lausanne, Suisse)
- T. O. Gallouët: Séminaire CMAP, École polytechnique.

7.2. Teaching - Supervision - Juries

7.2.1. Teaching

- Licence : J-D. Benamou, Méthodes Numériques , 39 H. équivalent TD, niveau (L2), université Paris Dauphine , FR
- Licence : V. Duval, Probabilités Multidimensionnelles, 39 h équivalent TD, niveau L2, Université Paris-Dauphine, FR
- Master : V. Duval, Optimization for Machine Learning, 9h, niveau M2, Université PSL/ENS, FR
- Licence : I. Waldspurger, Analyse 2, 45 h équivalent TD, niveau L1, Université Paris-Dauphine, FR
- Master : I. Waldspurger, Optimization for Machine Learning, 6h, niveau M2, Université PSL/ENS, FR
- Licence : G. Carlier, algebre 1, L1 78h, Dauphine, FR
- Master : G. Carlier Variational and transport methods in economics, M2 Masef, 27h, Dauphine, FR
- Licence : A. Natale, Python, 44 h équivalent TD, niveau L2, Université Paris-Sud , FR
- Licence : P. Pegon, Analyse 3, 51 H. équivalent TD, TD niveau L2, Université Paris-Dauphine, FR
- Licence : P. Pegon, Intégrale de Lebesgue et probabilités, 44 H. équivalent TD, TD niveau L3, Université Paris-Dauphine, FR
- Master : P. Pegon, Pré-rentrée d'analyse, 16 H. équivalent TD, cours/TD niveau M1, Université Paris-Dauphine, FR
- Licence : T. O. Gallouët, Optimisation, 24h équivalent TD, niveau L3, Université d'Orsay), FR

7.2.2. Supervision

- PhD in progress : Paul Catala, Optimisation polynomiale pour les problèmes inverses en imagerie, 01/10/2016, V. Duval
- PhD in progress : Romain Petit, Méthodes sans grille pour l'imagerie, 01/10/2019, V. Duval

- PhD in progress : J-D. Benamou, Miao Yu, Application of optimal transport theory in seismic full waveform inversion, 1/10/2016. Co-supervised by J.-P. Vilotte (IPGP)
- PhD in progress : J-D. Benamou, Lucas Martinet, Calcul Haute Performance pour le Transport Optimal, application en Astrophysique, 1/10/2017
- PhD in progress : J-D. Benamou, Giorgi Rukhaia , An Optimal Transportation computational approach of inverse free-form optical surfaces design for extended sources, 1/05/2018
- PhD in progress: G. Carlier, Quentin Petit, mean-field games for cities modeling, (co-supervision with Y. Achdou and D. Tonon), 1/09/2018
- PhD in progress: G. Carlier, Katharina Eichinger, Systems of Monge-Ampere equations: a variational approach 1/09/2019
- PhD in progress: T. O. Gallouët, Gabriele Todeschi, Optimal transport and finite volume schemes, 1/09/2018

7.2.3. Juries

J-D. Benamou, PhD Defense of Aude Genevay (13/03).
G. Carlier was in the Ph.D committee of Michael Orieux, Aymeric Baradat and Rui Chen.

7.3. Popularization

- I. Waldspurger, *Quand les films prennent des couleurs*, exposé au cycle SMAI & Musée des arts des métiers

8. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. GENEVAY.*Entropy-Regularized Optimal Transport for Machine Learning*, PSL University, March 2019, https://tel.archives-ouvertes.fr/tel-02319318

Articles in International Peer-Reviewed Journal

[2] J.-D. BENAMOU, G. CARLIER, M. LABORDE.*An augmented Lagrangian approach to Wasserstein gradient flows and applications*, in "ESAIM: Proceedings and Surveys", August 2019, https://hal.archives-ouvertes.fr/hal-01245184

[3] J.-D. BENAMOU, V. DUVAL.*Minimal convex extensions and finite difference discretization of the quadratic Monge-Kantorovich problem*, in "European Journal of Applied Mathematics", 2019, https://arxiv.org/abs/1710.05594 , forthcoming [*DOI :* 10.1017/S0956792518000451], https://hal.inria.fr/hal-01616842

[4] J.-D. BENAMOU, T. GALLOUËT, F.-X. VIALARD.*Second order models for optimal transport and cubic splines on the Wasserstein space*, in "Foundations of Computational Mathematics", October 2019, https://arxiv.org/abs/1801.04144 , https://hal.archives-ouvertes.fr/hal-01682107

[5] C. BOYER, A. CHAMBOLLE, Y. DE CASTRO, V. DUVAL, F. DE GOURNAY, P. WEISS.*On Representer Theorems and Convex Regularization*, in "SIAM Journal on Optimization", May 2019, vol. 29, n° 2, p. 1260–1281, https://arxiv.org/abs/1806.09810 [*DOI :* 10.1137/18M1200750], https://hal.archives-ouvertes.fr/hal-01823135

[6] C. Cancès, T. Gallouët, M. Laborde, L. Monsaingeon.*Simulation of multiphase porous media flows with minimizing movement and finite volume schemes*, in "European Journal of Applied Mathematics", 2019, vol. 30, n^o 6, p. 1123-1152 [*DOI :* 10.1017/S0956792518000633], https://hal.archives-ouvertes.fr/hal-01700952

[7] S. Dallaporta, Y. De Castro.*Sparse Recovery from Extreme Eigenvalues Deviation Inequalities*, in "ESAIM: Probability and Statistics", 2019, https://arxiv.org/abs/1604.01171 - 33 pages, 1 figure [*DOI :* 10.1051/PS/2018024], https://hal.archives-ouvertes.fr/hal-01309439

[8] Y. De Castro, F. Gamboa, D. Henrion, R. Hess, J. B. Lasserre.*Approximate Optimal Designs for Multivariate Polynomial Regression*, in "Annals of Statistics", January 2019, vol. 47, n^o 1, p. 127-155 [*DOI :* 10.1214/18-AOS1683], https://hal.laas.fr/hal-01483490

[9] Q. Denoyelle, V. Duval, G. Peyré, E. Soubies.*The Sliding Frank-Wolfe Algorithm and its Application to Super-Resolution Microscopy*, in "Inverse Problems", 2019, https://arxiv.org/abs/1811.06416 , forthcoming [*DOI :* 10.1088/1361-6420/AB2A29], https://hal.archives-ouvertes.fr/hal-01921604

[10] T. Gallouët, M. Laborde, L. Monsaingeon.*An unbalanced optimal transport splitting scheme for general advection-reaction-diffusion problems*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2019, vol. 25, n^o 8, https://arxiv.org/abs/1704.04541 [*DOI :* 10.1051/COCV/2018001], https://hal.archives-ouvertes.fr/hal-01508911

[11] T. Gallouët, A. Natale, F.-X. Vialard.*Generalized compressible flows and solutions of the H(div) geodesic problem*, in "Archive for Rational Mechanics and Analysis", 2020, https://arxiv.org/abs/1806.10825 , forthcoming, https://hal.archives-ouvertes.fr/hal-01815531

[12] R. E. Gaunt, G. Mijoule, Y. Swan.*Some new Stein operators for product distributions*, in "Brazilian Journal of Probability and Statistics", 2019, https://arxiv.org/abs/1901.11460 - 13 pages, forthcoming, https://hal.archives-ouvertes.fr/hal-02017801

[13] A. Natale, F.-X. Vialard.*Embedding Camassa-Holm equations in incompressible Euler*, in "Journal of Geometric Mechanics", June 2019, https://arxiv.org/abs/1804.11080 , https://hal.archives-ouvertes.fr/hal-01781162

[14] P. Pegon, F. Santambrogio, Q. Xia.*A fractal shape optimization problem in branched transport*, in "Journal de Mathématiques Pures et Appliquées", March 2019, https://arxiv.org/abs/1709.01415 , https://hal.archives-ouvertes.fr/hal-01581675

International Conferences with Proceedings

[15] J.-B. Courbot, E. Monfrini, V. Mazet, C. Collet.*Triplet markov trees for image segmentation*, in "SSP 2018: IEEE Workshop on Statistical Signal Processing", Fribourg-en-Brisgau, Germany, 2018 IEEE Statistical Signal Processing Workshop (SSP), IEEE Computer Society, 2019, p. 233-237 [*DOI :* 10.1109/SSP.2018.8450841], https://hal.archives-ouvertes.fr/hal-01815562

Conferences without Proceedings

[16] J. M. Fadili, G. Garrigos, J. Malick, G. Peyré.*Model Consistency for Learning with Mirror-Stratifiable Regularizers*, in "International Conference on Artificial Intelligence and Statistics (AISTATS)", Naha, Japan, April 2019, https://hal.archives-ouvertes.fr/hal-01988309

Other Publications

[17] J.-M. AZAÏS, Y. DE CASTRO.*Multiple Testing and Variable Selection along Least Angle Regression's path*, July 2019, https://arxiv.org/abs/1906.12072 - 39 pages, 7 figures, https://hal.archives-ouvertes.fr/hal-02170476

[18] J.-D. BENAMOU, G. CARLIER, S. D. MARINO, L. NENNA.*An entropy minimization approach to second-order variational mean-field games*, September 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01848370

[19] C. CANCÈS, T. GALLOUËT, G. TODESCHI.*A variational finite volume scheme for Wasserstein gradient flows*, July 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02189050

[20] J.-B. COURBOT, V. DUVAL, B. LEGRAS.*Sparse analysis for mesoscale convective systems tracking*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02010436

[21] S. DI MARINO, A. NATALE, R. TAHRAOUI, F.-X. VIALARD.*Metric completion of $Diff([0,1])$ with the $H1$ right-invariant metric*, June 2019, https://arxiv.org/abs/1906.09139 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02161686

[22] V. DUVAL.*An Epigraphical Approach to the Representer Theorem*, December 2019, https://arxiv.org/abs/1912.13224 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02424908

[23] G. MIJOULE, G. REINERT, Y. SWAN.*Stein operators, kernels and discrepancies for multivariate continuous distributions*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02420874

[24] A. NATALE, G. TODESCHI.*TPFA Finite Volume Approximation of Wasserstein Gradient Flows*, January 2020, https://arxiv.org/abs/2001.07005 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02444833

References in notes

[25] I. ABRAHAM, R. ABRAHAM, M. BERGOUNIOUX, G. CARLIER.*Tomographic reconstruction from a few views: a multi-marginal optimal transport approach*, in "Preprint Hal-01065981", 2014

[26] Y. ACHDOU, V. PEREZ.*Iterative strategies for solving linearized discrete mean field games systems*, in "Netw. Heterog. Media", 2012, vol. 7, n^{o} 2, p. 197–217, http://dx.doi.org/10.3934/nhm.2012.7.197

[27] M. AGUEH, G. CARLIER.*Barycenters in the Wasserstein space*, in "SIAM J. Math. Anal.", 2011, vol. 43, n^{o} 2, p. 904–924

[28] F. ALTER, V. CASELLES, A. CHAMBOLLE.*Evolution of Convex Sets in the Plane by Minimizing the Total Variation Flow*, in "Interfaces and Free Boundaries", 2005, vol. 332, p. 329–366

[29] F. R. BACH.*Consistency of the Group Lasso and Multiple Kernel Learning*, in "J. Mach. Learn. Res.", June 2008, vol. 9, p. 1179–1225, http://dl.acm.org/citation.cfm?id=1390681.1390721

[30] F. R. BACH.*Consistency of Trace Norm Minimization*, in "J. Mach. Learn. Res.", June 2008, vol. 9, p. 1019–1048, http://dl.acm.org/citation.cfm?id=1390681.1390716

[31] H. H. BAUSCHKE, P. L. COMBETTES.*A Dykstra-like algorithm for two monotone operators*, in "Pacific Journal of Optimization", 2008, vol. 4, n^{o} 3, p. 383–391

[32] M. F. BEG, M. I. MILLER, A. TROUVÉ, L. YOUNES.*Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms*, in "International Journal of Computer Vision", February 2005, vol. 61, n^{o} 2, p. 139–157, http://dx.doi.org/10.1023/B:VISI.0000043755.93987.aa

[33] M. BEIGLBOCK, P. HENRY-LABORDÈRE, F. PENKNER.*Model-independent bounds for option prices mass transport approach*, in "Finance and Stochastics", 2013, vol. 17, n^{o} 3, p. 477-501, http://dx.doi.org/10.1007/s00780-013-0205-8

[34] G. BELLETTINI, V. CASELLES, M. NOVAGA.*The Total Variation Flow in R^N*, in "J. Differential Equations", 2002, vol. 184, n^{o} 2, p. 475–525

[35] J.-D. BENAMOU, Y. BRENIER.*A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, in "Numer. Math.", 2000, vol. 84, n^{o} 3, p. 375–393, http://dx.doi.org/10.1007/s002110050002

[36] J.-D. BENAMOU, Y. BRENIER.*Weak existence for the semigeostrophic equations formulated as a coupled Monge-Ampère/transport problem*, in "SIAM J. Appl. Math.", 1998, vol. 58, n^{o} 5, p. 1450–1461

[37] J.-D. BENAMOU, G. CARLIER.*Augmented Lagrangian algorithms for variational problems with divergence constraints*, in "JOTA", 2015

[38] J.-D. BENAMOU, G. CARLIER, N. BONNE.*An Augmented Lagrangian Numerical approach to solving Mean-Fields Games*, Inria, December 2013, 30, http://hal.inria.fr/hal-00922349

[39] J.-D. BENAMOU, G. CARLIER, M. CUTURI, L. NENNA, G. PEYRÉ.*Iterative Bregman Projections for Regularized Transportation Problems*, in "SIAM J. Sci. Comp.", 2015, to appear

[40] J.-D. BENAMOU, G. CARLIER, Q. MÉRIGOT, É. OUDET.*Discretization of functionals involving the Monge-Ampère operator*, HAL, July 2014, https://hal.archives-ouvertes.fr/hal-01056452

[41] J.-D. BENAMOU, F. COLLINO, J.-M. MIREBEAU.*Monotone and Consistent discretization of the Monge-Ampère operator*, in "arXiv preprint arXiv:1409.6694", 2014, to appear in Math of Comp

[42] J.-D. BENAMOU, B. D. FROESE, A. OBERMAN.*Two numerical methods for the elliptic Monge-Ampère equation*, in "M2AN Math. Model. Numer. Anal.", 2010, vol. 44, n^{o} 4, p. 737–758

[43] J.-D. BENAMOU, B. D. FROESE, A. OBERMAN.*Numerical solution of the optimal transportation problem using the Monge–Ampere equation*, in "Journal of Computational Physics", 2014, vol. 260, p. 107–126

[44] F. BENMANSOUR, G. CARLIER, G. PEYRÉ, F. SANTAMBROGIO.*Numerical approximation of continuous traffic congestion equilibria*, in "Netw. Heterog. Media", 2009, vol. 4, n^{o} 3, p. 605–623

[45] M. BENNING, M. BURGER.*Ground states and singular vectors of convex variational regularization methods*, in "Meth. Appl. Analysis", 2013, vol. 20, p. 295–334

[46] B. BERKELS, A. EFFLAND, M. RUMPF.*Time discrete geodesic paths in the space of images*, in "Arxiv preprint", 2014

[47] J. BIGOT, T. KLEIN.*Consistent estimation of a population barycenter in the Wasserstein space*, in "Preprint arXiv:1212.2562", 2012

[48] A. BLANCHET, P. LAURENÇOT.*The parabolic-parabolic Keller-Segel system with critical diffusion as a gradient flow in R^d, $d \geq 3$*, in "Comm. Partial Differential Equations", 2013, vol. 38, n$^{\text{o}}$ 4, p. 658–686, http://dx.doi.org/10.1080/03605302.2012.757705

[49] J. BLEYER, G. CARLIER, V. DUVAL, J.-M. MIREBEAU, G. PEYRÉ.*A Γ Convergence Result for the Upper Bound Limit Analysis of Plates*, in "ESAIM: Mathematical Modelling and Numerical Analysis", January 2016, vol. 50, n$^{\text{o}}$ 1, p. 215–235 [*DOI :* 10.1051/M2AN/2015040], https://www.esaim-m2an.org/articles/m2an/abs/2016/01/m2an141087/m2an141087.html

[50] N. BONNEEL, J. RABIN, G. PEYRÉ, H. PFISTER.*Sliced and Radon Wasserstein Barycenters of Measures*, in "Journal of Mathematical Imaging and Vision", 2015, vol. 51, n$^{\text{o}}$ 1, p. 22–45, http://hal.archives-ouvertes.fr/hal-00881872/

[51] U. BOSCAIN, R. CHERTOVSKIH, J.-P. GAUTHIER, D. PRANDI, A. REMIZOV.*Highly corrupted image inpainting through hypoelliptic diffusion*, Preprint CMAP, 2014, http://hal.archives-ouvertes.fr/hal-00842603/

[52] G. BOUCHITTÉ, G. BUTTAZZO.*Characterization of optimal shapes and masses through Monge-Kantorovich equation*, in "J. Eur. Math. Soc. (JEMS)", 2001, vol. 3, n$^{\text{o}}$ 2, p. 139–168, http://dx.doi.org/10.1007/s100970000027

[53] L. BRASCO, G. CARLIER, F. SANTAMBROGIO.*Congested traffic dynamics, weak flows and very degenerate elliptic equations*, in "J. Math. Pures Appl. (9)", 2010, vol. 93, n$^{\text{o}}$ 6, p. 652–671

[54] L. M. BREGMAN.*The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming*, in "USSR computational mathematics and mathematical physics", 1967, vol. 7, n$^{\text{o}}$ 3, p. 200–217

[55] Y. BRENIER.*Generalized solutions and hydrostatic approximation of the Euler equations*, in "Phys. D", 2008, vol. 237, n$^{\text{o}}$ 14-17, p. 1982–1988, http://dx.doi.org/10.1016/j.physd.2008.02.026

[56] Y. BRENIER.*Décomposition polaire et réarrangement monotone des champs de vecteurs*, in "C. R. Acad. Sci. Paris Sér. I Math.", 1987, vol. 305, n$^{\text{o}}$ 19, p. 805–808

[57] Y. BRENIER.*Polar factorization and monotone rearrangement of vector-valued functions*, in "Comm. Pure Appl. Math.", 1991, vol. 44, n$^{\text{o}}$ 4, p. 375–417, http://dx.doi.org/10.1002/cpa.3160440402

[58] Y. BRENIER, U. FRISCH, M. HENON, G. LOEPER, S. MATARRESE, R. MOHAYAEE, A. SOBOLEVSKI.*Reconstruction of the early universe as a convex optimization problem*, in "Mon. Not. Roy. Astron. Soc.", 2003, vol. 346, p. 501–524, http://arxiv.org/pdf/astro-ph/0304214.pdf

[59] M. BRUVERIS, L. RISSER, F.-X. VIALARD.*Mixture of Kernels and Iterated Semidirect Product of Diffeomorphisms Groups*, in "Multiscale Modeling & Simulation", 2012, vol. 10, n$^{\text{o}}$ 4, p. 1344-1368

[60] M. BURGER, M. DIFRANCESCO, P. MARKOWICH, M. T. WOLFRAM.*Mean field games with nonlinear mobilities in pedestrian dynamics*, in "DCDS B", 2014, vol. 19

[61] M. BURGER, M. FRANEK, C. SCHONLIEB.*Regularized regression and density estimation based on optimal transport*, in "Appl. Math. Res. Expr.", 2012, vol. 2, p. 209–253

[62] M. BURGER, S. OSHER.*A guide to the TV zoo*, in "Level-Set and PDE-based Reconstruction Methods, Springer", 2013

[63] G. BUTTAZZO, C. JIMENEZ, É. OUDET.*An optimization problem for mass transportation with congested dynamics*, in "SIAM J. Control Optim.", 2009, vol. 48, n^{o} 3, p. 1961–1976

[64] H. BYRNE, D. DRASDO.*Individual-based and continuum models of growing cell populations: a comparison*, in "Journal of Mathematical Biology", 2009, vol. 58, n^{o} 4-5, p. 657-687

[65] L. A. CAFFARELLI.*The regularity of mappings with a convex potential*, in "J. Amer. Math. Soc.", 1992, vol. 5, n^{o} 1, p. 99–104, http://dx.doi.org/10.2307/2152752

[66] L. A. CAFFARELLI, S. A. KOCHENGIN, V. OLIKER.*On the numerical solution of the problem of reflector design with given far-field scattering data*, in "Monge Ampère equation: applications to geometry and optimization (Deerfield Beach, FL, 1997)", Providence, RI, Contemp. Math., Amer. Math. Soc., 1999, vol. 226, p. 13–32, http://dx.doi.org/10.1090/conm/226/03233

[67] C. CANCERITOGLU.*Computational Analysis of LDDMM for Brain Mapping*, in "Frontiers in Neuroscience", 2013, vol. 7

[68] C. CANCÈS, T. GALLOUËT, L. MONSAINGEON.*Incompressible immiscible multiphase flows in porous media: a variational approach*, in "Analysis & PDE", 2017, vol. 10, n^{o} 8, p. 1845–1876 [*DOI :* 10.2140/APDE.2017.10.1845], https://hal.archives-ouvertes.fr/hal-01345438

[69] E. CANDES, M. WAKIN.*An Introduction to Compressive Sensing*, in "IEEE Signal Processing Magazine", 2008, vol. 25, n^{o} 2, p. 21–30

[70] E. J. CANDÈS, C. FERNANDEZ-GRANDA.*Super-Resolution from Noisy Data*, in "Journal of Fourier Analysis and Applications", 2013, vol. 19, n^{o} 6, p. 1229–1254

[71] E. J. CANDÈS, C. FERNANDEZ-GRANDA.*Towards a Mathematical Theory of Super-Resolution*, in "Communications on Pure and Applied Mathematics", 2014, vol. 67, n^{o} 6, p. 906–956

[72] P. CARDALIAGUET, G. CARLIER, B. NAZARET.*Geodesics for a class of distances in the space of probability measures*, in "Calc. Var. Partial Differential Equations", 2013, vol. 48, n^{o} 3-4, p. 395–420

[73] G. CARLIER.*A general existence result for the principal-agent problem with adverse selection*, in "J. Math. Econom.", 2001, vol. 35, n^{o} 1, p. 129–150

[74] G. CARLIER, V. CHERNOZHUKOV, A. GALICHON.*Vector Quantile Regression*, Arxiv 1406.4643, 2014

[75] G. CARLIER, M. COMTE, I. IONESCU, G. PEYRÉ.*A Projection Approach to the Numerical Analysis of Limit Load Problems*, in "Mathematical Models and Methods in Applied Sciences", 2011, vol. 21, n^o 6, p. 1291–1316 [*DOI :* DOI:10.1142/S0218202511005325], http://hal.archives-ouvertes.fr/hal-00450000/

[76] G. CARLIER, X. DUPUIS.*An iterated projection approach to variational problems under generalized convexity constraints and applications*, In preparation, 2015

[77] G. CARLIER, C. JIMENEZ, F. SANTAMBROGIO.*Optimal Transportation with Traffic Congestion and Wardrop Equilibria*, in "SIAM Journal on Control and Optimization", 2008, vol. 47, n^o 3, p. 1330-1350

[78] G. CARLIER, T. LACHAND-ROBERT, B. MAURY.*A numerical approach to variational problems subject to convexity constraint*, in "Numer. Math.", 2001, vol. 88, n^o 2, p. 299–318, http://dx.doi.org/10.1007/PL00005446

[79] G. CARLIER, A. OBERMAN, É. OUDET.*Numerical methods for matching for teams and Wasserstein barycenters*, in "M2AN", 2015, to appear

[80] J. A. CARRILLO, S. LISINI, E. MAININI.*Uniqueness for Keller-Segel-type chemotaxis models*, in "Discrete Contin. Dyn. Syst.", 2014, vol. 34, n^o 4, p. 1319–1338, http://dx.doi.org/10.3934/dcds.2014.34.1319

[81] V. CASELLES, A. CHAMBOLLE, M. NOVAGA.*The discontinuity set of solutions of the TV denoising problem and some extensions*, in "Multiscale Modeling and Simulation", 2007, vol. 6, n^o 3, p. 879–894

[82] F. A. C. C. CHALUB, P. A. MARKOWICH, B. PERTHAME, C. SCHMEISER.*Kinetic models for chemotaxis and their drift-diffusion limits*, in "Monatsh. Math.", 2004, vol. 142, n^o 1-2, p. 123–141, http://dx.doi.org/10.1007/s00605-004-0234-7

[83] A. CHAMBOLLE, T. POCK.*On the ergodic convergence rates of a first-order primal-dual algorithm*, in "Preprint OO/2014/09/4532", 2014

[84] G. CHARPIAT, G. NARDI, G. PEYRÉ, F.-X. VIALARD.*Finsler Steepest Descent with Applications to Piecewise-regular Curve Evolution*, Preprint hal-00849885, 2013, http://hal.archives-ouvertes.fr/hal-00849885/

[85] S. S. CHEN, D. L. DONOHO, M. A. SAUNDERS.*Atomic decomposition by basis pursuit*, in "SIAM journal on scientific computing", 1999, vol. 20, n^o 1, p. 33–61

[86] P. CHONÉ, H. V. J. LE MEUR.*Non-convergence result for conformal approximation of variational problems subject to a convexity constraint*, in "Numer. Funct. Anal. Optim.", 2001, vol. 22, n^o 5-6, p. 529–547, http://dx.doi.org/10.1081/NFA-100105306

[87] C. COTAR, G. FRIESECKE, C. KLUPPELBERG.*Density Functional Theory and Optimal Transportation with Coulomb Cost*, in "Communications on Pure and Applied Mathematics", 2013, vol. 66, n^o 4, p. 548–599, http://dx.doi.org/10.1002/cpa.21437

[88] M. J. P. CULLEN, W. GANGBO, G. PISANTE.*The semigeostrophic equations discretized in reference and dual variables*, in "Arch. Ration. Mech. Anal.", 2007, vol. 185, n^o 2, p. 341–363, http://dx.doi.org/10.1007/s00205-006-0040-6

[89] M. J. P. CULLEN, J. NORBURY, R. J. PURSER. *Generalised Lagrangian solutions for atmospheric and oceanic flows*, in "SIAM J. Appl. Math.", 1991, vol. 51, n^{o} 1, p. 20–31

[90] M. CUTURI. *Sinkhorn Distances: Lightspeed Computation of Optimal Transport*, in "Proc. NIPS", C. J. C. BURGES, L. BOTTOU, Z. GHAHRAMANI, K. Q. WEINBERGER (editors), 2013, p. 2292–2300

[91] E. J. DEAN, R. GLOWINSKI. *Numerical methods for fully nonlinear elliptic equations of the Monge-Ampère type*, in "Comput. Methods Appl. Mech. Engrg.", 2006, vol. 195, n^{o} 13-16, p. 1344–1386

[92] V. DUVAL, G. PEYRÉ. *Exact Support Recovery for Sparse Spikes Deconvolution*, in "Foundations of Computational Mathematics", 2014, p. 1-41, http://dx.doi.org/10.1007/s10208-014-9228-6

[93] V. DUVAL, G. PEYRÉ. *Sparse regularization on thin grids I: the L asso*, in "Inverse Problems", 2017, vol. 33, n^{o} 5, 055008 [*DOI :* 10.1088/1361-6420/AA5E12], http://stacks.iop.org/0266-5611/33/i=5/a=055008

[94] J. FEHRENBACH, J.-M. MIREBEAU. *Sparse Non-negative Stencils for Anisotropic Diffusion*, in "Journal of Mathematical Imaging and Vision", 2014, vol. 49, n^{o} 1, p. 123-147, http://dx.doi.org/10.1007/s10851-013-0446-3

[95] C. FERNANDEZ-GRANDA. *Support detection in super-resolution*, in "Proc. Proceedings of the 10th International Conference on Sampling Theory and Applications", 2013, p. 145–148

[96] A. FIGALLI, R. MCCANN, Y. KIM. *When is multi-dimensional screening a convex program?*, in "Journal of Economic Theory", 2011

[97] J.-B. FIOT, H. RAGUET, L. RISSER, L. D. COHEN, J. FRIPP, F.-X. VIALARD. *Longitudinal deformation models, spatial regularizations and learning strategies to quantify Alzheimer's disease progression*, in "NeuroImage: Clinical", 2014, vol. 4, n^{o} 0, p. 718 - 729 [*DOI :* 10.1016/J.NICL.2014.02.002], http://www.sciencedirect.com/science/article/pii/S2213158214000205

[98] J.-B. FIOT, L. RISSER, L. D. COHEN, J. FRIPP, F.-X. VIALARD. *Local vs Global Descriptors of Hippocampus Shape Evolution for Alzheimer's Longitudinal Population Analysis*, in "Spatio-temporal Image Analysis for Longitudinal and Time-Series Image Data", Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2012, vol. 7570, p. 13-24, http://dx.doi.org/10.1007/978-3-642-33555-6_2

[99] U. FRISCH, S. MATARRESE, R. MOHAYAEE, A. SOBOLEVSKI. *Monge-Ampère-Kantorovitch (MAK) reconstruction of the eary universe*, in "Nature", 2002, vol. 417, n^{o} 260

[100] B. D. FROESE, A. OBERMAN. *Convergent filtered schemes for the Monge-Ampère partial differential equation*, in "SIAM J. Numer. Anal.", 2013, vol. 51, n^{o} 1, p. 423–444

[101] A. GALICHON, P. HENRY-LABORDÈRE, N. TOUZI. *A stochastic control approach to No-Arbitrage bounds given marginals, with an application to Loopback options*, in "submitted to Annals of Applied Probability", 2011

[102] W. GANGBO, R. MCCANN. *The geometry of optimal transportation*, in "Acta Math.", 1996, vol. 177, n^{o} 2, p. 113–161, http://dx.doi.org/10.1007/BF02392620

[103] E. GHYS.*Gaspard Monge, Le mémoire sur les déblais et les remblais*, in "Image des mathématiques, CNRS", 2012, http://images.math.cnrs.fr/Gaspard-Monge,1094.html

[104] O. GUÉANT, J.-M. LASRY, P.-L. LIONS.*Mean field games and applications*, in "Paris-Princeton Lectures on Mathematical Finance 2010", Berlin, Lecture Notes in Math., Springer, 2011, vol. 2003, p. 205–266, http://dx.doi.org/10.1007/978-3-642-14660-2_3

[105] G. HERMAN.*Image reconstruction from projections: the fundamentals of computerized tomography*, Academic Press, 1980

[106] D. D. HOLM, J. T. RATNANATHER, A. TROUVÉ, L. YOUNES.*Soliton dynamics in computational anatomy*, in "NeuroImage", 2004, vol. 23, p. S170–S178

[107] B. J. HOSKINS.*The mathematical theory of frontogenesis*, in "Annual review of fluid mechanics, Vol. 14", Palo Alto, CA, Annual Reviews, 1982, p. 131–151

[108] R. JORDAN, D. KINDERLEHRER, F. OTTO.*The variational formulation of the Fokker-Planck equation*, in "SIAM J. Math. Anal.", 1998, vol. 29, n^{o} 1, p. 1–17

[109] W. JÄGER, S. LUCKHAUS.*On explosions of solutions to a system of partial differential equations modelling chemotaxis*, in "Trans. Amer. Math. Soc.", 1992, vol. 329, n^{o} 2, p. 819–824, http://dx.doi.org/10.2307/2153966

[110] L. KANTOROVITCH.*On the translocation of masses*, in "C. R. (Doklady) Acad. Sci. URSS (N.S.)", 1942, vol. 37, p. 199–201

[111] E. KLANN.*A Mumford-Shah-Like Method for Limited Data Tomography with an Application to Electron Tomography*, in "SIAM J. Imaging Sciences", 2011, vol. 4, n^{o} 4, p. 1029–1048

[112] J.-M. LASRY, P.-L. LIONS.*Mean field games*, in "Jpn. J. Math.", 2007, vol. 2, n^{o} 1, p. 229–260, http://dx.doi.org/10.1007/s11537-007-0657-8

[113] J. LASSERRE.*Global Optimization with Polynomials and the Problem of Moments*, in "SIAM Journal on Optimization", 2001, vol. 11, n^{o} 3, p. 796-817

[114] J. LELLMANN, D. A. LORENZ, C. SCHÖNLIEB, T. VALKONEN.*Imaging with Kantorovich-Rubinstein Discrepancy*, in "SIAM J. Imaging Sciences", 2014, vol. 7, n^{o} 4, p. 2833–2859

[115] A. S. LEWIS.*Active sets, nonsmoothness, and sensitivity*, in "SIAM Journal on Optimization", 2003, vol. 13, n^{o} 3, p. 702–725

[116] B. LI, F. HABBAL, M. ORTIZ.*Optimal transportation meshfree approximation schemes for Fluid and plastic Flows*, in "Int. J. Numer. Meth. Engng 83:1541–579", 2010, vol. 83, p. 1541–1579

[117] G. LOEPER.*A fully nonlinear version of the incompressible Euler equations: the semigeostrophic system*, in "SIAM J. Math. Anal.", 2006, vol. 38, n^{o} 3, p. 795–823 (electronic)

[118] G. LOEPER, F. RAPETTI.*Numerical solution of the Monge-Ampére equation by a Newton's algorithm*, in "C. R. Math. Acad. Sci. Paris", 2005, vol. 340, n^o 4, p. 319–324

[119] D. LOMBARDI, E. MAITRE.*Eulerian models and algorithms for unbalanced optimal transport*, in "Preprint hal-00976501", 2013

[120] C. LÉONARD.*A survey of the Schrödinger problem and some of its connections with optimal transport*, in "Discrete Contin. Dyn. Syst.", 2014, vol. 34, n^o 4, p. 1533–1574, http://dx.doi.org/10.3934/dcds.2014.34.1533

[121] J. MAAS, M. RUMPF, C. SCHONLIEB, S. SIMON.*A generalized model for optimal transport of images including dissipation and density modulation*, in "Arxiv preprint", 2014

[122] S. G. MALLAT.*A wavelet tour of signal processing*, Third, Elsevier/Academic Press, Amsterdam, 2009

[123] B. MAURY, A. ROUDNEFF-CHUPIN, F. SANTAMBROGIO.*A macroscopic crowd motion model of gradient flow type*, in "Math. Models Methods Appl. Sci.", 2010, vol. 20, n^o 10, p. 1787–1821, http://dx.doi.org/10.1142/S0218202510004799

[124] M. I. MILLER, A. TROUVÉ, L. YOUNES.*Geodesic Shooting for Computational Anatomy*, in "Journal of Mathematical Imaging and Vision", March 2006, vol. 24, n^o 2, p. 209–228, http://dx.doi.org/10.1007/s10851-005-3624-0

[125] J.-M. MIREBEAU.*Adaptive, Anisotropic and Hierarchical cones of Discrete Convex functions*, in "Preprint", 2014

[126] J.-M. MIREBEAU.*Anisotropic Fast-Marching on Cartesian Grids Using Lattice Basis Reduction*, in "SIAM Journal on Numerical Analysis", 2014, vol. 52, n^o 4, p. 1573-1599

[127] Q. MÉRIGOT.*A multiscale approach to optimal transport*, in "Computer Graphics Forum", 2011, vol. 30, n^o 5, p. 1583–1592

[128] Q. MÉRIGOT, É. OUDET.*Handling Convexity-Like Constraints in Variational Problems*, in "SIAM J. Numer. Anal.", 2014, vol. 52, n^o 5, p. 2466–2487

[129] N. PAPADAKIS, G. PEYRÉ, É. OUDET.*Optimal Transport with Proximal Splitting*, in "SIAM Journal on Imaging Sciences", 2014, vol. 7, n^o 1, p. 212–238 [*DOI :* 10.1137/130920058], http://hal.archives-ouvertes.fr/hal-00816211/

[130] B. PASS, N. GHOUSSOUB.*Optimal transport: From moving soil to same-sex marriage*, in "CMS Notes", 2013, vol. 45, p. 14–15

[131] B. PASS.*Uniqueness and Monge Solutions in the Multimarginal Optimal Transportation Problem*, in "SIAM Journal on Mathematical Analysis", 2011, vol. 43, n^o 6, p. 2758-2775

[132] B. PERTHAME, F. QUIROS, J. L. VAZQUEZ.*The Hele-Shaw Asymptotics for Mechanical Models of Tumor Growth*, in "Archive for Rational Mechanics and Analysis", 2014, vol. 212, n^o 1, p. 93-127, http://dx.doi.org/10.1007/s00205-013-0704-y

[133] J. PETITOT.*The neurogeometry of pinwheels as a sub-riemannian contact structure*, in "Journal of Physiology-Paris", 2003, vol. 97, n° 23, p. 265–309

[134] G. PEYRÉ.*Texture Synthesis with Grouplets*, in "Pattern Analysis and Machine Intelligence, IEEE Transactions on", April 2010, vol. 32, n° 4, p. 733–746

[135] B. PICCOLI, F. ROSSI.*Generalized Wasserstein distance and its application to transport equations with source*, in "Archive for Rational Mechanics and Analysis", 2014, vol. 211, n° 1, p. 335–358

[136] C. POON.*Structure dependent sampling in compressed sensing: theoretical guarantees for tight frames*, in "Applied and Computational Harmonic Analysis", 2015

[137] H. RAGUET, J. FADILI, G. PEYRÉ.*A Generalized Forward-Backward Splitting*, in "SIAM Journal on Imaging Sciences", 2013, vol. 6, n° 3, p. 1199–1226 [*DOI :* 10.1137/120872802], http://hal.archives-ouvertes.fr/hal-00613637/

[138] J.-C. ROCHET, P. CHONÉ.*Ironing, Sweeping and multi-dimensional screening*, in "Econometrica", 1998

[139] L. RUDIN, S. OSHER, E. FATEMI.*Nonlinear total variation based noise removal algorithms*, in "Physica D: Nonlinear Phenomena", 1992, vol. 60, n° 1, p. 259–268, http://dx.doi.org/10.1016/0167-2789(92)90242-F

[140] O. SCHERZER, M. GRASMAIR, H. GROSSAUER, M. HALTMEIER, F. LENZEN.*Variational Methods in Imaging*, Springer, 2008

[141] T. SCHMAH, L. RISSER, F.-X. VIALARD.*Left-Invariant Metrics for Diffeomorphic Image Registration with Spatially-Varying Regularisation*, in "MICCAI (1)", 2013, p. 203-210

[142] T. SCHMAH, L. RISSER, F.-X. VIALARD.*Diffeomorphic image matching with left-invariant metrics*, in "Fields Institute Communications series, special volume in memory of Jerrold E. Marsden", January 2014

[143] J. SOLOMON, F. DE GOES, G. PEYRÉ, M. CUTURI, A. BUTSCHER, A. NGUYEN, T. DU, L. GUIBAS.*Convolutional Wasserstein Distances: Efficient Optimal Transportation on Geometric Domains*, in "ACM Transaction on Graphics, Proc. SIGGRAPH'15", 2015, to appear

[144] R. TIBSHIRANI.*Regression shrinkage and selection via the Lasso*, in "Journal of the Royal Statistical Society. Series B. Methodological", 1996, vol. 58, n° 1, p. 267–288

[145] A. TROUVÉ, F.-X. VIALARD.*Shape splines and stochastic shape evolutions: A second order point of view*, in "Quarterly of Applied Mathematics", 2012

[146] S. VAITER, M. GOLBABAEE, J. FADILI, G. PEYRÉ.*Model Selection with Piecewise Regular Gauges*, in "Information and Inference", 2015, to appear, http://hal.archives-ouvertes.fr/hal-00842603/

[147] F.-X. VIALARD, L. RISSER, D. RUECKERT, C. COTTER.*Diffeomorphic 3D Image Registration via Geodesic Shooting Using an Efficient Adjoint Calculation*, in "International Journal of Computer Vision", 2012, vol. 97, n° 2, p. 229-241, http://dx.doi.org/10.1007/s11263-011-0481-8

[148] F.-X. VIALARD, L. RISSER.*Spatially-Varying Metric Learning for Diffeomorphic Image Registration: A Variational Framework*, in "Medical Image Computing and Computer-Assisted Intervention MICCAI 2014", Lecture Notes in Computer Science, Springer International Publishing, 2014, vol. 8673, p. 227-234

[149] C. VILLANI.*Topics in optimal transportation*, Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003, vol. 58, xvi+370

[150] C. VILLANI.*Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 2009, vol. 338, xxii+973, Old and new, http://dx.doi.org/10.1007/978-3-540-71050-9

[151] X.-J. WANG.*On the design of a reflector antenna. II*, in "Calc. Var. Partial Differential Equations", 2004, vol. 20, n° 3, p. 329–341, http://dx.doi.org/10.1007/s00526-003-0239-4

[152] B. WIRTH, L. BAR, M. RUMPF, G. SAPIRO.*A continuum mechanical approach to geodesics in shape space*, in "International Journal of Computer Vision", 2011, vol. 93, n° 3, p. 293–318

[153] J. WRIGHT, Y. MA, J. MAIRAL, G. SAPIRO, T. S. HUANG, S. YAN.*Sparse representation for computer vision and pattern recognition*, in "Proceedings of the IEEE", 2010, vol. 98, n° 6, p. 1031–1044

Project-Team OURAGAN

Tools for resolutions in algebra, geometry and their applications

IN COLLABORATION WITH: Institut de Mathématiques de Jussieu

IN PARTNERSHIP WITH:
CNRS
Sorbonne Université (UPMC)
Université Denis Diderot (Paris 7)

RESEARCH CENTER
Paris

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team OURAGAN

Creation of the Team: 2012 January 01, updated into Project-Team: 2019 May 01

Keywords:

Computer Science and Digital Science:

A4.3. - Cryptography
A4.3.1. - Public key cryptography
A4.3.2. - Secret key cryptography
A4.3.3. - Cryptographic protocols
A4.3.4. - Quantum Cryptography
A7.1. - Algorithms
A7.1.4. - Quantum algorithms
A8.1. - Discrete mathematics, combinatorics
A8.3. - Geometry, Topology
A8.4. - Computer Algebra
A8.5. - Number theory
A8.10. - Computer arithmetic

Other Research Topics and Application Domains:

B5.6. - Robotic systems
B9.5.1. - Computer science
B9.5.2. - Mathematics

1. Team, Visitors, External Collaborators

Research Scientists

Fabrice Rouillier [Team leader, Inria, Senior Researcher, HDR]
Mohamed Yacine Bouzidi [Inria, Starting Research Position]
Alban Quadrat [Inria, Senior Researcher, HDR]
Robin Timsit [Univ Pierre et Marie Curie, Researcher, until Sep 2019]
Elias Tsigaridas [Inria, Researcher, from Sep 2019]

Faculty Members

Jean Bajard [Univ Pierre et Marie Curie, Professor, from Sep 2019]
Elisha Falbel [Univ Pierre et Marie Curie, Professor]
Antonin Guilloux [Univ Pierre et Marie Curie, Associate Professor, HDR]
Antoine Joux [Univ Pierre et Marie Curie, Associate Professor, HDR]
Pierre-Vincent Koseleff [Univ Pierre et Marie Curie, Associate Professor, HDR]
Pascal Molin [Univ Denis Diderot, Associate Professor]

External Collaborator

Razvan Barbulescu [CNRS, until Aug 2019]

PhD Students

Christina Katsamaki [Inria, PhD Student, from Oct 2019]
Mahya Mehrabdollahei [Inria, PhD Student]
Sudarshan Shinde [Univ Pierre et Marie Curie, PhD Student, until Sep 2019]
Grace Younes [Inria, PhD Student]

Visiting Scientist

Thomas Espitau [Univ de Nanterre, Sep 2019]

Administrative Assistants

Laurence Bourcier [Inria, Administrative Assistant]

Maryse Desnous [Inria, Administrative Assistant, until Apr 2019]

Julien Guieu [Inria, Administrative Assistant, from Apr 2019]

2. Overall Objectives

2.1. Overall Objectives

OURAGAN proposes to focus on the transfer of computational algebraic methods to some related fields (computational geometry, topology, number theory, etc.) and some carefully chosen application domains (robotics, control theory, evaluation of the security of cryptographic systems, etc.), which implies working equally on the use (modeling, know - how) and on the development of new algorithms. The latest breakthrough developments and applications where algebraic methods are currently decisive remain few and very targeted. We wish to contribute to increase the impact of these methods but also the number of domains where the use of computational algebraic methods represent a significant added value. This transfer-oriented positioning does not imply to stop working on the algorithms, it simply sets the priorities.

An original aspect of the OURAGAN proposal is to blend into an environment of fundamental mathematics, at the Institut de Mathématiques de Jussieu – Paris Rive Gauche (IMJ-PRG CNRS 7586), and to be cross-functional to several teams (Algebraic Analysis, Complex Analysis and Geometry, Number Theory to name only the main ones), which will be our first source of transfer of computational know-how. The success of this coupling allows to maintain a strong theoretical basis and to measure objectively our transfer activity in the direction of mathematicians (in geometry, topology, number theory, algebraic analysis, etc.) and to consolidate the presence of Inria in scientific areas among the most theoretical.

We propose three general directions with five particular targets:

- Number Theory
  - Algorithmic Number Theory
  - Rigorous Numerical Computations
- Topology in small dimension
  - Character varieties
  - Knot theory
  - Computational geometry
- Algebraic analysis of functional systems

These actions come, of course, in addition to the study and development of a common set of core elements of

- Basic theory and algorithms in algebra and geometry [Transverse activity].

This core activity is the invention and study of fundamental algebraic algorithms and objects that can be grouped into 2 categories: algorithms designed to operate on finite fields and algorithms running on fields of characteristic 0; with 2 types of computational strategies: the exactness and the use of approximate arithmetic (but with exact results). This mix also installs joint studies between the various axes and is an originality of the project-team. For example many kinds of arithmetic tools around algebraic numbers have to face to similar theoretical problems such as finding a good representation for a number field; almost all problems related to the resolution of algebraic systems will reduce to the study of varieties in small dimension and in particular, most of the time, to the effective computation of the topology of curves and surfaces, or the certified drawing of non algebraic function over an algebraic variety.

The tools and objects developed for research on algorithmic number theory as well as in computational geometry apply quite directly on some selected connected challenging subjects:

- Security of cryptographic systems
- Control theory
- Robotics
- Signal processing

These applications will serve for the evaluation of the general tools we develop when used in a different context, in particular their capability to tackle state of the art problems.

2.2. Scientific ground

2.2.1. Basic computable objects and algorithms

The basic computable objects and algorithms we study, use, optimize or develop are among the most classical ones in computer algebra and are studied by many people around the world: they mainly focus on basic computer arithmetic, linear algebra, lattices, and both polynomial system and differential system solving.

In the context of OURAGAN, it is important to avoid reinventing the wheel and to re-use wherever possible existing objects and algorithms, not necessarily developed in our team so that the main effort is focused on finding good formulations/modelisations for an efficient use. Also, our approach for the development of basic computable objects and algorithms is *application driven* and follows a simple strategy : use the existing tools in priority, develop missing tools when required and then optimize the critical operations. First, for some selected problems, we do propose and develop general key algorithms (isolation of real roots of univariate polynomials, parametrisations of solutions of zero-dimensional polynomial systems, solutions of parametric equations, equidimensional decompositions, etc.) in order to complement the existing set computable objects developed and studied around the world (Gröbner bases, resultants [64], subresultants [85], critical point methods [41], etc.) which are also deeply used in our developments. Second, for a selection of well-known problems, we propose different computational strategies (for example the use of approximate arithmetic to speed up LLL algorithm or root isolators, still certifying the final result). Last, we propose specialized variants of known algorithms optimized for a given problem (for example, dedicated solvers for degenerated bivariate polynomials to be used in the computation of the topology of plane curves).

In the activity of OURAGAN, many key objects or algorithms around the resolution of algebraic systems are developed or optimized within the team, such as the resolution of polynomials in one variable with real coefficients [105] [14], rational parameterizations of solutions of zero-dimensional systems with rational coefficients [49] [13] or discriminant varieties for solving systems depending on parameters [11], but we are also power users of existing software (mainly Sage [0], Maple [0], Pari-GP [0],Snappea [0]) and libraries (mainly gmp [0], mpfr [0], flint [0], arb [0], etc.) to which we contribute when it makes sense.

For our studies in number theory and applications to the security of cryptographic systems, our team works on three categories of basic algorithms: discrete logarithm computations [99] (for example to make progress on the computation of class groups in number fields [90]), network reductions by means of LLL variants [75] and, obviously, various computations in linear algebra, for example dedicated to *almost sparse* matrices [100].

Finally, for the algorithmic approach to algebraic analysis of functional equations [45] [103] [104], we developed the effective study of both module theory and homological algebra [136] over certain noncommutative polynomial rings of functional operators [4], of Stafford's famous theorems on the Weyl algebras [127], of the equidimensional decomposition of functional systems [122], etc.

[0] http://www.sagemath.org/
[0] https://maplesoft.com
[0] https://pari.math.u-bordeaux.fr
[0] http://www.geometrygames.org/SnapPea/
[0] https://gmplib.org/
[0] https://www.mpfr.org/
[0] http://www.flintlib.org/
[0] http://arblib.org/

2.2.2. Computational Number Theory

Many frontiers between computable objects, algorithms (above section), computational number theory and applications, especially in cryptography are porous. However, one can classify our work in computational number theory into two classes of studies : computational algebraic number theory and (rigorous) numerical computations in number theory.

Our work on rigorous numerical computations is somehow a transverse activity in Ouragan : floating point arithmetic is used in many basic algorithms we develop (root isolation, LLL) and is thus present in almost all our research directions. However there are specific developments that could be labelized *Number Theory*, in particular contributions to numerical evaluations of L-functions which are deeply used in many problems in number theory (for example the Riemann Zeta function). We participate, for example to the *L-functions and Modular Forms Database*[0] a world wide collaborative project.

Our work in computational algebraic number theory is driven by the algorithmic improvement to solve presumably hard problems relevant to cryptography. The use of number-theoretic hard problems in cryptography dates back to the invention of public-key cryptography by Diffie and Hellman [71], where they proposed a first instantiation of their paradigm based on the discrete logarithm problem in prime fields. The invention of RSA [134], based on the hardness of factoring came as a second example. The introduction of discrete logarithms on elliptic curves [106] [116] only confirmed this trend.

These crypto-systems attracted a lot of interest on the problems of factoring and discrete log. Their study led to the invention of fascinating new algorithms that can solve the problems much faster than initially expected :

- the elliptic curve method (ECM) [101]
- the quadratic field for factoring [120] and its variant for discrete log called the Gaussian integers method [114]
- the number field sieve (NFS) [39]

Since the invention of NFS in the 90's, many optimizations of this algorithm have been performed. However, an algorithm with better complexity hasn't been found for factoring and discrete logarithms in large characteristic.

While factorization and discrete logarithm problems have a long history in cryptography, the recent post-quantum cryptosystems introduce a new variety of presumably hard problems/objects/algorithms with cryptographic relevance: the shortest vector problem (SVP), the closest vector problem (CVP) or the computation of isogenies between elliptic curves, especially in the supersingular case.

Members of OURAGAN started working on the topic of discrete logarithms around 1998, with several computation records that were announced on the *NMBRTHRY* mailing list. In large characteristic, especially for the case of prime fields, the best current method is the number field sieve (NFS) algorithm. In particular, they published the first NFS based record computation[10]. Despite huge practical improvements, the prime field case algorithm hasn't really changed since that first record. Around the same time, we also presented small characteristic computation record based on simplifications of the Function Field Sieve (FFS) algorithm [98].

In 2006, important changes occurred concerning the FFS and NFS algorithms, indeed, while the algorithms only covered the extreme case of constant characteristic and constant extension degree, two papers extended their ranges of applicability to all finite fields. At the same time, this permitted a big simplification of the FFS, removing the need for function fields.

Starting from 2012, new results appeared in small characteristic. Initially based on a simplification of the 2006 result, they quickly blossomed into the Frobenial representation methods, with quasi-polynomial time complexity [99], [91].

An interesting side-effect of this research was the need to revisit the key sizes of pairing-based cryptography. This type of cryptography is also a topic of interest for OURAGAN. In particular, it was introduced in 2000 [9].

[0] http://www.lmfdb.org

The computations of *class groups in number fields* has strong links with the computations of discrete logarithms or factorizations using the NFS (number field sieve) strategy which as the name suggests is based on the use of number fields. Roughly speaking, the NFS algorithm uses two number fields and the strategy consists in choosing number fields with small sized coefficients in their definition polynomials. On the contrary, in class group computations, there is a single number field, which is clearly a simplification, but this field is given as input by some fixed definition polynomial. Obviously, the degree of this polynomial as well as the size of its coefficients are both influencing the complexity of the computations so that finding other polynomials representing the same class group but with a better characterization (degree or coefficient's sizes) is a mathematical problem with direct practical consequences. We proposed a method to address the problem [90], but many issues remain open.

Computing generators of principal ideals of cyclotomic fields is also strongly related to the computation of class groups in number fields. Ideals in cyclotomic fields are used in a number of recent public-key cryptosystems. Among the difficult problems that ensure the safety of these systems, there is one that consists in finding a small generator, if it exists, of an ideal. The case of cyclotomic fields is considered [44].

2.2.3. Topology in small dimension

2.2.3.1. Character varieties

There is a tradition of using computations and software to study and understand the topology of small dimensional manifolds, going back at least to Thurston's works (and before him, Riley's pioneering work). The underlying philosophy of these tools is to build combinatorial models of manifolds (for example, the torus is often described as a square with an identification of the sides). For dimensions 2, 3 and 4, this approach is relevant and effective. In the team OURAGAN, we focus on the dimension 3, where the manifolds are modelized by a finite number of tetrahedra with identification of the faces. The software SnapPy [0] implements this strategy [139] and is regularly used as a starting point in our work. Along the same philosophy of implementation, we can also cite Regina [0]. A specific trait of SnapPy is that it focuses on hyperbolic structures on the 3-dimensional manifolds. This setting is the object of a huge amount of theoretical work that were used to speed up computations. For example, some Newton methods were implemented without certification for solving a system of equations, but the theoretical knowledge of the uniqueness of the solution made this implementation efficient enough for the target applications. In recent years, in part under the influence of our team [0], more attention has been given to certified computations (at least with an error control) and now this is implemented in SnapPy.

This philosophy (modelization of manifolds by quite simple combinatoric models to compute such complicated objects as representations of the fundamental group) was applied in a pioneering work of Falbel [8] when he begins to look for another type of geometry on 3-dimensional manifolds (called CR-spherical geometry). From a computational point of view, this change of objectives was a jump in the unknown: the theoretical justification for the computations were missing, and the number of variables of the systems were multiplied by four. So instead of a relatively small system that could be tackled by Newton methods and numerical approximations, we had to deal with/study (were in front of) relatively big systems (the smallest example being 8 variables of degree 6) with no a priori description of the solutions.

Still, the computable objects that appear from the theoretical study are very often outside the reach of automated computations and are to be handled case by case. A few experts around the world have been tackling this kind of computations (Dunfield, Goerner, Heusener, Porti, Tillman, Zickert) and the main current achievement is the *Ptolemy module*[0] for SnapPy.

From these early computational needs, topology in small dimension has historically been the source of collaboration with the IMJ-PRG laboratory. At the beginning, the goal was essentially to provide computational tools for finding geometric structures in triangulated 3-dimensional varieties. Triangulated varieties can be

[0] https://www.math.uic.edu/t3m/SnapPy/
[0] https://regina-normal.github.io
[0] as part of the CURVE project
[0] https://www.math.uic.edu/t3m/SnapPy/ptolemy.html

topologically encoded by a collection of tetrahedra with gluing constraints (this can be called a triangulation or mesh, but it is not an approximation of the variety by simple structures, rather a combinatorial model). Imposing a geometric structure on this combinatorial object defines a number of constraints that we can translate into an algebraic system that we then have to solve to study geometric structures of the initial variety, for example in relying on solutions to study representations of the fundamental group of the variety. For these studies, a large part of the computable objects or algorithms we develop are required, from the algorithms for univariate polynomials to systems depending on parameters. It should be noted that most of the computational work lies in the modeling of problems [43][7] that have strictly no chance to be solved by blindly running the most powerful black boxes: we usually deal here with systems that have 24 to 64 variables, depend on 4 to 8 parameters and with degrees exceeding 10 in each variable. With an ANR [0] funding on the subject, the progress that we did [79] were (much) more significant than expected. In particular, we have introduced new computable objects with an immediate theoretical meaning (let us say rather with a theoretical link established with the usual objects of the domain), namely, the so-called *deformation variety*.

2.2.3.2. Knot theory

Knot theory is a wide area of mathematics. We are interested in polynomial representations of long knots, that is to say polynomial embeddings $\mathbf{R} \to \mathbf{R}^3 \subset \mathbf{S}^3$. Every knot admits a polynomial representation and a natural question is to determine explicit parameterizations, minimal degree parameterizations. On the other hand we are interested to determine what is the knot of a given polynomial smooth embedding $\mathbf{R} \to \mathbf{R}^3$. These questions involve real algebraic curves. This subject was first considered by Vassiliev in the 90's [138].

A Chebyshev knot [108], is a polynomial knot parameterized by a Chebyshev curve $(T_a(t), T_b(t), T_c(t + \varphi))$ where $T_n(t) = \cos(n \arccos t)$ is the n-th Chebyshev polynomial of the first kind. Chebyshev knots are polynomial analogues of Lissajous knots that have been studied by Jones, Hoste, Lamm... It was first established that any knot can be parameterized by Chebyshev polynomials, then we have studied the properties of harmonic nodes [110] which then opened the way to effective computations.

Our activity in Knot theory is a bridge between our work in computational geometry (topology and drawing of real space curves) and our work on topology in small dimensions (varieties defined as a knot complement).

Two-bridge knots (or rational knots) are particularly studied because they are much easier to study. The first 26 knots (except 8_5) are two-bridge knots. We were able to give an exhaustive, minimal and certified list of Chebyshev parameterizations of the first rational two-bridge knots, using blind computations [111]. On the other hand, we propose the identification of Chebyshev knot diagrams [112] by developing new certified algorithms for computing trigonometric expressions [113]. These works share many tools with our action in visualization and computational geometry.

We made use of Chebyshev polynomials so as Fibonacci polynomials which are families of orthogonal polynomials. Considering the Alexander-Conway polynomials as continuant polynomials in the Fibonacci basis, we were able to give a partial answer to Hoste's conjecture on the roots of Alexander polynomials of alternating knots ([109]).

We study the lexicographic degree of the two-bridge knots, that is to say the minimal (multi)degree of a polynomial representation of a N-crossing two-bridge knot. We show that this degree is $(3, b, c)$ with $b + c = 3N$. We have determined the lexicographic degree of the first 362 first two-bridge knots with 12 crossings or fewer [58] [0]. These results make use of the braid theoretical approach developped by Y. Orevkov to study real plane curves and the use of real pseudoholomorphic curves [56], the slide isotopies on trigonal diagrams, namely those that never increase the number of crossings [57].

2.2.3.3. Visualization and Computational Geometry

The drawing of algebraic curves and surfaces is a critical action in OURAGAN since it is a key ingredient in numerous developments. For example, a certified plot of a discriminant variety could be the only admissible answer that can be proposed for engineering problems that need the resolution of parametric algebraic systems:

[0] ANR project Structures Géométriques et Triangulations
[0] Minimal degrees are listed in https://webusers.imj-prg.fr/~pierre-vincent.koseleff/knots/2bk-lexdeg.html

this variety (and the connected components of its counter part) defines a partition of the parameter's space in regions above which the solutions are numerically stable and topologically simple. Several directions have been explored since the last century, ranging from pure numerical computations to infallible exact ones, depending on the needs (global topology, local topology, simple drawing, etc.). For plane real algebraic curves, one can mention the cylindrical algebraic decomposition [63], grids methods (for ex. the marching square algorithm), subdivision methods, etc.

As mentioned above, we focus on curves and surfaces coming from the study of parametric systems. They mostly come from some elimination process, they highly (numerically) unstable (a small deformation of the coefficients might change a lot the topology of the curve) and we are mostly interested in getting qualitative information about their counter part in the parameter's space.

For this work, we are associated with the GAMBLE EPI (Inria Nancy Grand Est) with the aim of developing computational techniques for the study, plotting and topology. In this collaboration, Ouragan focuses on CAD-Like methods while Gamble develops numerical strategies (that could also apply on non algebraic curves). Ouragan's work involves the development of effective methods for the resolution of algebraic systems with 2 or 3 variables [49], [105], [50], [51] which are basic engines for computing the topology [118], [70] and / or plotting.

2.2.4. Algebraic analysis of functional systems

Systems of functional equations or simply functional systems are systems whose unknowns are functions, such as systems of ordinary or partial differential equations, of differential time-delay equations, of difference equations, of integro-differential equations, etc.

Numerical aspects of functional systems, especially differential systems, have been widely studied in applied mathematics due to the importance of numerical simulation issues.

Complementary approaches, based on algebraic methods, are usually upstream or help the numerical simulation of systems of functional systems. These methods also tackle a different range of questions and problems such as algebraic preconditioning, elimination and simplification, completion to formal integrability or involution, computation of integrability conditions and compatibility conditions, index reduction, reduction of variables, choice of adapted coordinate systems based on symmetries, computation of first integrals of motion, conservation laws and Lax pairs, Liouville integrability, study of the (asymptotic) behavior of solutions at a singularity, etc. Although not yet very popular in applied mathematics, these theories have lengthy been studied in fundamental mathematics and were developed by Lie, Cartan, Janet, Ritt, Kolchin, Spencer, etc. [94] [103] [104] [107] [133] [121].

Over the past years, certain of these algebraic approaches to functional systems have been investigated within an algorithmic viewpoint, mostly driven by applications to engineering sciences such as mathematical systems theory and control theory. We have played a role towards these effective developments, especially in the direction of an algorithmic approach to the so-called *algebraic analysis* [103], [104], [45], a mathematical theory developed by the Japanese school of Sato, which studies linear differential systems by means of both algebraic and analytic methods. To develop an effective approach to algebraic analysis, we first have to make algorithmic standard results on rings of functional operators, module theory, homological algebra, algebraic geometry, sheaf theory, category theory, etc., and to implement them in computer algebra systems. Based on elimination theory (Gröbner or Janet bases [94], [62], [135], differential algebra [47] [76], Spencer's theory [121], etc.), in [4], [5], we have initiated such a computational algebraic analysis approach for general classes of functional systems (and not only for holonomic systems as done in the literature of computer algebra [62]). Based on the effective aspects to algebraic analysis approach, the parametrizability problem [4], the reduction and (Serre) decomposition problems [5], the equidimensional decomposition [122], Stafford's famous theorems for the Weyl algebras [127], etc., have been studied and solutions have been implemented in `Maple`, `Mathematica`, and `GAP` [61][5]. But these results are only the first steps towards computational algebraic analysis, its implementation in computer algebra systems, and its applications to mathematical systems, control theory, signal processing, mathematical physics, etc.

2.2.5. Synergies

Outside applications which can clearly be seen as transversal acitivies, our development directions are linked at several levels : shared computable objects, computational strategies and transversal research directions.
Sharing basic algebraic objects As seen above, is the well-known fact that the elimination theory for functional systems is deeply intertwined with the one for polynomial systems so that, topology in small dimension, applications in control theory, signal theory and robotics share naturally a large set of computable objects developped in our project team.

Performing efficient basic arithmetic operations in number fields is also a key ingredient to most of our algorithms, in Number theory as well as in topology in small dimension or , more generally in the use of roots of polynomials systems. In particular, finding good representations of number fields, lead to the same computational problems as working with roots of polynomial systems by means of triangular systems (towers of number fields) or rational parameterizations (unique number field). Making any progress in one direction will probably have direct consequences for almost all the problems we want to tackle.
Symbolic-numeric strategies. Several general low-level tools are also shared such as the use of approximate arithmetic to speed up certified computations. Sometimes these can also lead to improvement for a different purpose (for example computations over the rationals, deeply used in geometry can often be performed in parallel combining computations in finite fields together with fast Chinese remaindering and modular evaluations).

As simple example of this sharing of tools and strategies, the use of approximate arithmetic is common to the work on LLL (used in the evaluation of the security of cryptographic systems), resolutions of real-world algebraic systems (used in our applications in robotics, control theory, and signal theory), computations of signs of trigonometric expressions used in knot theory or to certified evaluations of dilogarithm functions on an algebraic variety for the computation of volumes of representations in our work in topology, numerical integration and computations of L-functions.
Transversal research directions. The study of the topology of complex algebraic curves is central in the computation of periods of algebraic curves (number theory) but also in the study of character varieties (topology in small dimension) as well as in control theory (stability criteria). Very few computational tools exists for that purpose and they mostly convert the problem to the one of variety over the reals (we can then recycle our work in computational geometry).

As for real algebraic curves, finding a way to describe the topology (an equivalent to the graph obtained in the real case) or computing certified drawings (in the case of a complex plane curve, a useful drawing is the so called associated amoeba) are central subjects for Ouragan.

As mentioned in the section 3.3.1 the computation of the Mahler measure of an algebraic implicit curve is either a challenging problem in number theory and a new direction in topology. The basic formula requires the study of points of moduli 1 , as for stability problems in Control Theory (stability problems), and certified numerical evaluations of non algebraic functions at algebraic points as for many computations for L-Functions.

3. Research Program

3.1. Basic computable objects and algorithms

The development of basic computable objects is somehow *on demand* and depends on all the other directions. However, some critical computations are already known to be bottlenecks and are sources of constant efforts.

Computations with algebraic numbers appear in almost all our activities: when working with number fields in our work in algorithmic number theory as well as in all the computations that involve the use of solutions of zero-dimensional systems of polynomial equations. Among the identified problems: finding good representations for single number fields (optimizing the size and degree of the defining polynomials), finding good representations for towers or products of number fields (typically working with a tower or finding a unique good extension), efficiently computing in practice with number fields (using certified approximation vs working with the formal description based on polynomial arithmetics). Strong efforts are currently done in the understanding of the various strategies by means of tight theoretical complexity studies [70], [115], [50] and many other efforts will be required to find the right representation for the right problem in practice. For example, for isolating critical points of plane algebraic curves, it is still unclear (at least the theoretical complexity cannot help) that an intermediate formal parameterization is more efficient than a triangular decomposition of the system and it is still unclear that these intermediate computations could be dominated in time by the certified final approximation of the roots.

3.2. Algorithmic Number Theory

Concerning algorithmic number theory, the main problems we will be considering in the coming years are the following:

- *Number fields.* We will continue working on the problems of class groups and generators. In particular, the existence and accessibility of *good* defining polynomials for a fixed number field remain very largely open. The impact of better polynomials on the algorithmic performance is a very important parameter, which makes this problem essential.
- *Lattice reduction.* Despite a great amount of work in the past 35 years on the LLL algorithm and its successors, many open problems remain. We will continue the study of the use of interval arithmetic in this field and the analysis of variants of LLL along the lines of the *Potential*-LLL which provides improved reduction comparable to BKZ with a small block size but has better performance.
- *Elliptic curves and Drinfeld modules.* The study of elliptic curves is a very fruitful area of number theory with many applications in crypto and algorithms. Drinfeld modules are “cousins” of elliptic curves which have been less explored in the algorithm context. However, some recent advances [74] have used them to provide some fast sophisticated factoring algorithms. As a consequence, it is natural to include these objects in our research directions.

3.2.1. Rigorous numerical computations

Some studies in this area will be driven by some other directions, for example, the rigorous evaluation of non algebraic functions on algebraic varieties might become central for some of our work on topology in small dimension (volumes of varieties, drawing of amoeba) or control theory (approximations of discriminant varieties) are our two main current sources of interesting problems. In the same spirit, the work on L-functions computations (extending the computation range, algorithmic tools for computing algebraic data from the L function) will naturally follow.

On the other hand, another objective is to extend existing results on periods of algebraic curves to general curves and higher dimensional varieties is a general promising direction. This project aims at providing tools for integration on higher homology groups of algebraic curves, ie computing Gauss-Manin connections. It requires good understanding of their topology, and more algorithmic tools on differential equations.

3.3. Topology in small dimension

3.3.1. Character varieties

The brute force approach to computable objects from topology of small dimension will not allow any significant progress. As explained above, the systems that arise from these problems are simply outside the range of doable computations. We still continue the work in this direction by a four-fold approach, with all three directions deeply inter-related. First, we focus on a couple of especially meaningful (for the applications) cases, in particular the 3-dimensional manifold called Whitehead link complement. At this point, we are able to make steps in the computation and describe part of the solutions [79], [89]; we hope to be able to complete the computation using every piece of information to simplify the system. Second, we continue the theoretical work to understand more properties of these systems [77]. These properties may prove how useful for the mathematical understanding is the resolution of such systems - or at least the extraction of meaningful information. This approach is for example carried on by Falbel and his work on configuration of flags [80], [82]. Third, we position ourselves as experts in the know-how of this kind of computations and natural interlocutors for colleagues coming up with a question on such a computable object (see [87] and [89]). This also allows us to push forward the kind of computation we actually do and make progress in the direction of the second point. We are credible interlocutors because our team has the blend of theoretical knowledge and computational capabilities that grants effective resolutions of the problems we are presented. And last, we use the knowledge already acquired to pursue our theoretical study of the CR-spherical geometry [69], [81], [78].

Another direction of work is the help to the community in experimental mathematics on new objects. It involves downsizing the system we are looking at (for example by going back to systems coming from hyperbolic geometry and not CR-spherical geometry) and get the most out of what we can compute, by studying new objects. An example of this research direction is the work of Guilloux around the volume function on deformation varieties. This is a real-analytic function defined on the varieties we specialized in computing. Being able to do effective computations with this function led first to a conjecture [86]. Then, theoretical discussions around this conjecture led to a paper on a new approach to the Mahler measure of some 2-variables polynomials [88]. In turn, this last paper gave a formula for the Mahler measure in terms of a function akin to the volume function applied at points in an algebraic variety whose moduli of coordinates are 1. The OURAGAN team has the expertise to compute all the objects appearing in this formula, opening the way to another area of application. This area is deeply linked with number theory as well as topology of small dimension. It requires all the tools at disposition within OURAGAN.

3.3.2. Knot theory

We will carry on the exhaustive search for the lexicographic degrees for the rational knots. They correspond to trigonal space curves: computations in the braid group B_3, explicit parametrization of trigonal curves corresponding to "dessins d'enfants", etc. The problem seems much more harder when looking for more general knots.

On the other hand, a natural direction would be: given an explicit polynomial space curve, determine the under/over nature of the crossings when projecting, draw it and determine the known knot [0] it is isotopic to.

3.3.3. Vizualisation and Computational Geometry

As mentioned above, the drawing of algebraic curves and surfaces is a critical action in OURAGAN since it is a key ingredient in numerous developments. In some cases, one will need a fully certified study of the variety for deciding existence of solutions (for example a region in a robot's parameter's space with solutions to the DKP above or deciding if some variety crosses the unit polydisk for some stability problems in control-theory), in some other cases just a partial but certified approximation of a surface (path planning in robotics, evaluation of non algebraic functions over an algebraic variety for volumes of knot complements in the study of character varieties).

On the one hand, we will contribute to general tools like ISOTOP [0] under the supervision of the GAMBLE project-team and, on the other hand, we will propose ad-hoc solutions by gluing some of our basic tools (problems of high degrees in robust control theory). The priority is to provide a first software that implements methods that fit as most as possible the very last complexity results we got on several (theoretical) algorithms for the computation of the topology of plane curves.

[0] for example the first rational knots are listed at https://team.Inria.fr/ouragan/knots
[0] https://isotop.gamble.loria.fr

A particular effort will be devoted to the resolution of overconstraint bivariate systems which are useful for the studies of singular points and to polynomials systems in 3 variables in the same spirit : avoid the use of Gröbner basis and propose a new algorithm with a state-of-the-art complexity and with a good practical behavior.

In parallel, one will have to carefully study the drawing of graphs of non algebraic functions over algebraic complex surfaces for providing several tools which are useful for mathematicians working on topology in small dimension (a well known example is the drawing of amoebia, a way of representing a complex curve on a sheet of paper).

3.4. Algebraic analysis of functional systems

We want to further develop our expertise in the computational aspects of algebraic analysis by continuing to develop effective versions of results of module theory, homological algebra, category theory and sheaf theory [136] which play important roles in algebraic analysis [45], [103], [104] and in the algorithmic study of linear functional systems. In particular, we shall focus on linear systems of integro-differential-constant/varying/distributed delay equations [124], [126] which play an important role in mathematical systems theory, control theory, and signal processing [124], [131], [125], [128].

The rings of integro-differential operators are highly more complicated than the purely differential case (i.e. Weyl algebras) [12], due to the existence of zero-divisors, or the fact of having a coherent ring instead of a noetherian ring [42]. Therefore, we want to develop an algorithmic study of these rings. Following the direction initiated in [126] for the computation of zero divisors (based on the polynomial null spaces of certain operators), we first want to develop algorithms for the computation of left/right kernels and left/right/generalized inverses of matrices with entries in such rings, and to use these results in module theory (e.g. computation of syzygy modules, (shorter/shortest) free resolutions, split short/long exact sequences). Moreover, Stafford's results [137], algorithmically developed in [12] for rings of partial differential operators (i.e. the Weyl algebras), are known to still hold for rings of integro-differential operators. We shall study their algorithmic extensions. Our corresponding implementation will be extended accordingly.

Finally, within a computer algebra viewpoint, we shall continue to algorithmically study issues on rings of integro-differential-delay operators [124], [125] and their applications to the study of equivalences of differential constant/varying/distributed delay systems (e.g. Artstein's reduction, Fiagbedzi-Pearson's transformation) which play an important role in control theory.

4. Application Domains

4.1. Security of cryptographic systems

The study of the security of asymmetric cryptographic systems comes as an application of the work carried out in algorithmic number theory and revolves around the development and the use of a small number of general purpose algorithms (lattice reduction, class groups in number fields, discrete logarithms in finite fields, ...). For example, the computation of generators of principal ideals of cyclotomic fields can be seen as one of these applications since these are used in a number of recent public key cryptosystems.

The cryptographic community is currently very actively assessing the threat coming for the development of quantum computers. Indeed, such computers would permit tremendous progress on many number theoretic problems such as factoring or discrete logarithm computations and would put the security of current cryptosystem under a major risk. For this reason, there is a large global research effort dedicated to finding alternative methods of securing data. For example, the US standardization agency called NIST has recently launched a standardization process around this issue. In this context, OURAGAN is part of the competition and has submitted a candidate (which has not been selected) [40]. This method is based on number-theoretic ideas involving a new presumably difficult problem concerning the Hamming distance of integers modulo large numbers of Mersenne.

4.2. Robotics

Algebraic computations have tremendously been used in Robotics, especially in kinematics, since the last quarter of the 20th century [93]. For example, one can find algebraic proofs for the 40 possible solutions to the direct kinematics problem [117] for steward platforms and companion experiments based on Gröbner basis computations [83]. On the one hand, hard general kinematics problems involve too many variables for pure algebraic methods to be used in place of existing numerical or semi-numerical methods everywhere and everytime, and on the other hand, global algebraic studies allow to propose exhaustive classifications that cannot be reached by other methods,for some quite large classes.

Robotics is a long-standing collaborative work with LS2N (Laboratory of Numerical Sciences of Nantes). Work has recently focused on the offline study of mechanisms, mostly parallel, their singularities or at least some types of singularities (cuspidals robots [140]).

For most parallel or serial manipulators, pose variables and joints variables are linked by algebraic equations and thus lie an algebraic variety. The two-kinematics problems (the direct kinematics problem - DKP- and the inverse kinematics problem - IKP) consist in studying the preimage of the projection of this algebraic variety onto a subset of unknowns. Solving the DKP remains to computing the possible positions for a given set of joint variables values while solving the IKP remains to computing the possible joints variables values for a given position. Algebraic methods have been deeply used in several situations for studying parallel and serial mechanisms, but finally their use stays quite confidential in the design process. Cylindrical Algebraic Decomposition coupled with variable's eliminations by means of Gröbner based computations can be used to model the workspace, the joint space and the computation of singularities. On the one hand, such methods suffer immediately when increasing the number of parameters or when working with imprecise data. On the other hand, when the problem can be handled, they might provide full and exhaustive classifications. The tools we use in that context [60], [59], [95], [97], [96] depend mainly on the resolution of parameter-based systems and therefore of study-dependent curves or flat algebraic surfaces (2 or 3 parameters), thus joining our thematic *Computational Geometry*.

4.3. Control theory

Certain problems studied in mathematical systems theory and control theory can be better understood and finely studied by means of algebraic structures and methods. Hence, the rich interplay between algebra, computer algebra, and control theory has a long history. For instance, the first main paper on Gröbner bases written by their creators, Buchberger, was published in Bose's book [46] on control theory of multidimensional systems. Moreover, the differential algebra approach to nonlinear control theory (see [72], [73] and the references therein) was a major motivation for the algorithmic study of differential algebra [47], [76]. Finally, the behaviour approach to linear systems theory [141], [119] advocates for an algorithmic study of algebraic analysis (see Section 2.2.4). More generally, control theory is porous to computer algebra since one finds algebraic criteria of all kinds in the literature even if the control theory community has a very few knowledge in computer algebra.

OURAGAN has a strong interest in the computer algebra aspects of mathematical systems theory and control theory related to both functional and polynomial systems, particularly in the direction of robust stability analysis and robust stabilization problems for multidimensional systems [46], [119] and infinite-dimensional systems [66] (such as, e.g., differential time-delay systems).

Let us shortly state a few points of our recent interests in this direction.

In control theory, stability analysis of linear time-invariant control systems is based on the famous Routh-Hurwitz criterion (late 19th century) and its relation with Sturm sequences and Cauchy index. Thus, stability tests were only involving tools for univariate polynomials [102]. While extending those tests to multidimensional systems or differential time-delay systems, one had to tackle multivariate problems recursively with respect to the variables [46]. Recent works use a mix of symbolic/numeric strategies, Linear Matrix Inequalities (LMI), sums of squares, etc. But still very few practical experiments are currently involving certified algebraic computations based on general solvers for polynomial equations. We have recently started to study certified stability tests for multidimensional systems or differential time-delay systems with an important observation: with a correct modelization, some recent algebraic methods – derived from our work in algorithmic geometry and shared with applications in robotics – can now handle previously impossible computations and lead to a better understanding of the problems to be solved [52], [54], [55]. The previous approaches seem to be blocked on a recursive use of one-variable methods, whereas our approach involves the direct processing of the problem for a larger number of variables.

The structural stability of n-D discrete linear systems (with $n \geq 2$) is a good source of problems of several kinds ranging from solving univariate polynomials to studying algebraic systems depending on parameters. For instance, we show [53], [54], [55] that the standard characterization of the structural stability of a multivariate rational transfer function (namely, the denominator of the transfer function does not have solutions in the unit polydisc of $\mathbb{C}^n$) is equivalent to deciding whether or not a certain system of polynomial equations has real solutions. The use state-of-the-art computer algebra algorithms to check this last condition, and thus the structural stability of multidimensional systems has been validated in several situations from toy examples with parameters to state-of-the-art examples involving, e.g., the resolution of bivariate systems [51], [50].

The rich interplay between control theory, algebra, and computer algebra is also well illustrated with our recent work on robust stabilization problems for multidimensional and finite/infinite-dimensional systems [48], [123], [129], [132], [130], [131].

4.4. Signal processing

Due to numerous applications (e.g. sensor network, mobile robots), sources and sensors localization has intensively been studied in the literature of signal processing. The *anchor position self calibration problem* is a well-known problem which consists in estimating the positions of both the moving sources and a set of fixed sensors (anchors) when only the distance information between the points from the different sets is available. The position self-calibration problem is a particular case of the *Multidimensional Unfolding* (MDU) problem for the Euclidean space of dimension 3. In the signal processing literature, this problem is attacked by means of optimization problems (see [65] and the references therein). Based on computer algebra methods for polynomial systems, we have recently developed a new approach for the MDU problem which yields closed-form solutions and a very efficient algorithm for the estimation of the positions [68] based only on linear algebra techniques. This first result, done in collaboration with Dagher (Inria Chile) and Zheng (DEFROST, Inria Lille), yielded a recent patent [67]. This result advocates for the study of other localization problems based on the computational polynomial techniques developed in OURAGAN.

In collaboration with *Safran Tech* (Barau, Hubert) and Dagher (Inria Chile), a symbolic-numeric study of the new *multi-carrier demodulation method* [92] has recently been initiated. *Gear fault diagnosis* is an important issue in aeronautics industry since a damage in a gearbox, which is not detected in time, can have dramatic effects on the safety of a plane. Since the vibrations of a spur gear can be modeled as a product of two periodic functions related to the gearbox kinematic, it is proposed to recover each function from the global signal by means of an optimal reconstruction problem which, based on Fourier analysis, can be rewritten as $\mathrm{argmin}_{u \in \mathbb{C}^n, v_1, v_2 \in \mathbb{C}^m} \| M - u\, v_1^{\star} - D\, u\, v_2^{\star} \|_F$, where $M \in \mathbb{C}^{n \times m}$ (resp. $D \in \mathbb{C}^{n \times n}$) is a given matrix with a special shape (resp. diagonal matrix), $\| \cdot \|_F$ is the Frobenius norm, and $v^{\star}$ is the Hermitian transpose of v. We have recently obtained closed-form solutions for the exact problem, i.e., $M = u\, v_1^{\star} + D\, u\, v_2^{\star}$, which is a polynomial system with parameters. This first result gives interesting new insides for the study of the non-exact case, i.e. for the above optimization problem.

Our expertise on *algebraic parameter estimation problem*, developed in the former NON-A project-team (Inria Lille), will be further developed. Following this work [84], the problem consists in estimating a set θ of parameters of a signal $x(\theta, t)$– which satisfies a certain dynamics – when the signal $y(t) = x(\theta, t) + \gamma(t) + \varpi(t)$ is observed, where γ denotes a structured perturbation and ϖ a noise. It has been shown that θ can sometimes be explicitly determined by means of closed-form expressions using iterated integrals of y. These integrals are used to filter the noise ϖ. Based on a combination of algebraic analysis techniques (rings of differential operators), differential elimination theory (Gröbner basis techniques for Weyl algebras), and operational calculus (Laplace transform, convolution), an algorithmic approach to algebraic parameter estimation problem has been initiated in [125] for a particular type of structured perturbations (i.e. bias) and was implemented in the `Maple` prototype `NonA`. The case of a general structured perturbation is still lacking.

5. Highlights of the Year

5.1. Highlights of the Year

Two new projects have started this year

- the MACAO associated team in collaboration with the University of Wollongong (Australia) - see 9.3.1
- a collaboration with Safran Tech - see 8.1

6. New Software and Platforms

6.1. ISOTOP

Topology and geometry of planar algebraic curves

KEYWORDS: Topology - Curve plotting - Geometric computing

FUNCTIONAL DESCRIPTION: Isotop is a Maple software for computing the topology of an algebraic plane curve, that is, for computing an arrangement of polylines isotopic to the input curve. This problem is a necessary key step for computing arrangements of algebraic curves and has also applications for curve plotting. This software has been developed since 2007 in collaboration with F. Rouillier from Inria Paris - Rocquencourt.

- Participants: Luis Penaranda, Marc Pouget and Sylvain Lazard
- Contact: Marc Pouget
- Publications: Rational Univariate Representations of Bivariate Systems and Applications - Separating Linear Forms for Bivariate Systems - On The Topology of Planar Algebraic Curves - New bivariate system solver and topology of algebraic curves - Improved algorithm for computing separating linear forms for bivariate systems - Solving bivariate systems using Rational Univariate Representations - On the topology of planar algebraic curves - On the topology of real algebraic plane curves - Bivariate triangular decompositions in the presence of asymptotes - Separating linear forms and Rational Univariate Representations of bivariate systems
- URL: https://isotop.gamble.loria.fr/

6.2. RS

FUNCTIONAL DESCRIPTION: Real Roots isolation for algebraic systems with rational coefficients with a finite number of Complex Roots

- Participant: Fabrice Rouillier
- Contact: Fabrice Rouillier
- URL: https://team.inria.fr/ouragan/software/

6.3. A NewDsc

A New Descartes

KEYWORD: Scientific computing

FUNCTIONAL DESCRIPTION: Computations of the real roots of univariate polynomials with rational coefficients.

- Authors: Fabrice Rouillier, Alexander Kobel and Michael Sagraloff
- Partner: Max Planck Institute for Software Systems
- Contact: Fabrice Rouillier
- URL: https://anewdsc.mpi-inf.mpg.de

6.4. SIROPA

KEYWORDS: Robotics - Kinematics

FUNCTIONAL DESCRIPTION: Library of functions for certified computations of the properties of articulated mechanisms, particularly the study of their singularities

- Authors: Damien Chablat, Fabrice Rouillier, Guillaume Moroz and Philippe Wenger
- Partner: LS2N
- Contact: Guillaume Moroz
- URL: http://siropa.gforge.inria.fr/

6.5. MPFI

KEYWORD: Arithmetic

FUNCTIONAL DESCRIPTION: MPFI is a C library based on MPFR and GMP for multi precision floating point arithmetic.

- Contact: Fabrice Rouillier
- URL: http://mpfi.gforge.inria.fr

7. New Results

7.1. Certified non-conservative tests for the structural stability of discrete multidimensional systems

In [18], we present new computer algebra based methods for testing the structural stability of n-D discrete linear systems (with n at least 2). More precisely, we show that the standard characterization of the structural stability of a multivariate rational transfer function (namely, the denominator of the transfer function does not have solutions in the unit polydisc of $\mathbb{C}^n$) is equivalent to the fact that a certain system of polynomials does not have real solutions. We then use state-of-the-art computer algebra algorithms to check this last condition, and thus the structural stability of multidimensional systems.

7.2. Computing period matrices and the Abel-Jacobi map of superelliptic curves

In [24], we present an algorithm for the computation of period matrices and the Abel-Jacobi map of complex superelliptic curves given by an equation y m = f (x). It relies on rigorous numerical integration of differentials between Weierstrass points, which is done using Gauss method if the curve is hyperelliptic (m = 2) or the Double-Exponential method. The algorithm is implemented and makes it possible to reach thousands of digits accuracy even on large genus curves.

7.3. Voronoi diagram of orthogonal polyhedra in two and three dimensions

Voronoi diagrams are a fundamental geometric data structure for obtaining proximity relations. In [28], we consider collections of axis-aligned orthogonal polyhedra in two and three-dimensional space under the max-norm, which is a particularly useful scenario in certain application domains. We construct the exact Voronoi diagram inside an orthogonal polyhedron with holes defined by such polyhedra. Our approach avoids creating full-dimensional elements on the Voronoi diagram and yields a skeletal representation of the input object. We introduce a complete algorithm in 2D and 3D that follows the subdivision paradigm relying on a bounding-volume hierarchy; this is an original approach to the problem. The complexity is adaptive and comparable to that of previous methods. Under a mild assumption it is $O(n/\Delta)$ in 2D or $O(n.\alpha^2/\Delta^2)$ in 3D, where n is the number of sites, namely edges or facets resp., Δ is the maximum cell size for the subdivision to stop, and α bounds vertex cardinality per facet. We also provide a numerically stable, open-source implementation in Julia, illustrating the practical nature of our algorithm.

7.4. A symbolic computation approach towards the asymptotic stability analysis of differential systems with commensurate delays

In [30], the work aims at studying the asymptotic stability of retarded type linear differential systems with commensurate delays. Within the frequency-domain approach, it is well-known that the asymptotic stability of such a system is ensured by the condition that all the roots of the corresponding quasipolynomial have negative real parts. A classical approach for checking this condition consists in computing the set of critical zeros of the quasipolynomial, i.e., the roots (and the corresponding delays) of the quasipolynomial that lie on the imaginary axis, and then analyzing the variation of these roots with respect to the variation of the delay. Following this approach, based on solving algebraic systems techniques, we propose a certified and efficient symbolic-numeric algorithm for computing the set of critical roots of a quasipolynomial. Moreover, using recent algorithmic results developed by the computer algebra community, we present an efficient algorithm for the computation of Puiseux series at a critical zero which allows us to finely analyze the stability of the system with respect to the variation of the delay. Explicit examples are given to illustrate our algorithms.

7.5. On the computation of stabilizing controllers of multidimensional systems

In [25], we consider the open problem consisting in the computation of stabilizing controllers of an internally stabilizable MIMO multidimensional system. Based on homological algebra and the so-called *Polydisk Nullstellensatz*, we propose a general method towards the explicit computation of stabilizing controllers. We show how the homological algebra methods over the ring of structurally stable SISO multidimensional transfer functions can be made algorithmic based on standard Gröbner basis techniques over polynomial rings. The problem of computing stabilizing controllers is then reduced to the problem of obtaining an effective version of the Polydisk Nullstellensatz which, apart from a few cases, stays open and will be studied in forthcoming publications.

7.6. Algebraic aspects of the exact signal demodulation problem

In [29], we introduce a general class of problems originating from gearbox vibration analysis. Based on a previous work where demodulation was formulated as a matrix approximation problem, we study the specific case applicable to amplitude and phase demodulation. This problem can be rewritten as a polynomial system. Based on algebraic methods such as linear algebra and homological algebra, we focus on the characterization of the problem and solve it in the noise-free case.

7.7. General closed-form solutions of the position self-calibration problem

The work in [36] investigates the anchors and sources position self-calibration problem in the 3D space based on range measurements and without any prior restriction on the network configuration. Using a well known low-rank property of Euclidean distance matrices, we first reduce the problem to finding 12 unknowns ascribed in a 3×3 transformation matrix and a 3×1 translation vector. In order to estimate them, we then introduce a polynomial parametrization with 9 unknowns that are estimated by solving a linear system. Afterwards, we identify an intrinsic matrix polynomial system that encodes the solution set of the problem and provide a direct method for solving it. The resulting procedure is simple and straightforward to implement using standard numerical tools. We also show that closed-form solutions can always be obtained when the reference frame is fixed. This is illustrated by adopting reference frames from the literature and by introducing a triangular reference frame whose constraints are imposed only on one position set (anchor or source). Experimental results on synthetic and real sound data show that the proposed closed-form solutions efficiently solve the position self-calibration problem.

7.8. Certified lattice reduction

Quadratic form reduction and lattice reduction are fundamental tools in computational number theory and in computer science, especially in cryptography. The celebrated Lenstra–Lenstra–Lovász reduction algorithm (so-called LLL) has been improved in many ways through the past decades and remains one of the central methods used for reducing integral lattice basis. In particular, its floating-point variants—where the rational arithmetic required by Gram–Schmidt orthogonalization is replaced by floating-point arithmetic—are now the fastest known. However, the systematic study of the reduction theory of real quadratic forms or, more generally, of real lattices is not widely represented in the literature. When the problem arises, the lattice is usually replaced by an integral approximation of (a multiple of) the original lattice, which is then reduced. While practically useful and proven in some special cases, this method doesn't offer any guarantee of success in general. In [22], we present an adaptive-precision version of a generalized LLL algorithm that covers this case in all generality. In particular, we replace floating-point arithmetic by Interval Arithmetic to certify the behavior of the algorithm. We conclude by giving a typical application of the result in algebraic number theory for the reduction of ideal lattices in number fields.

7.9. Using Maple to analyse parallel robots

In [27], we present the SIROPA Maple Library which has been designed to study serial and parallel manipulators at the conception level. We show how modern algorithms in Computer Algebra can be used to study the workspace, the joint space but also the existence of some physical capabilities w.r.t. to some design parameters left as degree of freedom for the designer of the robot.

7.10. On the effective computation of stabilizing controllers of 2D systems

In [26], we show how stabilizing controllers for 2D systems can effectively be computed based on computer algebra methods dedicated to polynomial systems, module theory and homological algebra. The complete chain of algorithms for the computation of stabilizing controllers, implemented in Maple, is illustrated with an explicit example.

7.11. Updating key size estimations for pairings

Recent progress on NFS imposed a new estimation of the security of pairings. In [15], we study the best attacks against some of the most popular pairings. It allows us to propose new pairing-friendly curves of 128 bits and 192 bits of security.

7.12. Matrix formulae for Resultants and Discriminants of Bivariate Tensor-product Polynomials

The construction of optimal resultant formulae for polynomial systems is one of the main areas of research in computational algebraic geometry. However, most of the constructions are restricted to formulae for unmixed polynomial systems, that is, systems of polynomials which all have the same support. Such a condition is restrictive, since mixed systems of equations arise frequently in many problems. Nevertheless, resultant formulae for mixed polynomial systems is a very challenging problem. In [19], we present a square, Koszul-type, matrix, the determinant of which is the resultant of an arbitrary (mixed) bivariate tensor-product polynomial system. The formula generalizes the classical Sylvester matrix of two univariate polynomials, since it expresses a map of degree one, that is, the elements of the corresponding matrix are up to sign the coefficients of the input polynomials. Interestingly, the matrix expresses a primal-dual multiplication map, that is, the tensor product of a univariate multiplication map with a map expressing derivation in a dual space. In addition we prove an impossibility result which states that for tensor-product systems with more than two (affine) variables there are no universal degree-one formulae, unless the system is unmixed. Last but not least, we present applications of the new construction in the efficient computation of discriminants and mixed discriminants.

7.13. Separation bounds for polynomial systems

In [21], we rely on aggregate separation bounds for univariate polynomials to introduce novel worst-case separation bounds for the isolated roots of zero-dimensional, positive-dimensional, and overde- termined polynomial systems. We exploit the structure of the given system, as well as bounds on the height of the sparse (or toric) resultant, by means of mixed volume, thus establishing adaptive bounds. Our bounds improve upon Canny's Gap theorem [9]. Moreover, they exploit sparseness and they apply without any assumptions on the input polynomial system. To evaluate the quality of the bounds, we present polynomial systems whose root separation is asymptotically not far from our bounds. We apply our bounds to three problems. First, we use them to estimate the bitsize of the eigenvalues and eigenvectors of an integer matrix; thus we provide a new proof that the problem has polynomial bit complexity. Second, we bound the value of a positive polynomial over the simplex: we improve by at least one order of magnitude upon all existing bounds. Finally, we asymptotically bound the number of steps of any purely subdivision-based algorithm that isolates all real roots of a polynomial system.

7.14. On the maximal number of real embeddings of minimally rigid graphs in $\mathbb{R}^2$, $\mathbb{R}^3$ and S^2

Rigidity theory studies the properties of graphs that can have rigid embeddings in a euclidean space $\mathbb{R}^d$ or on a sphere and other manifolds which in addition satisfy certain edge length constraints. One of the major open problems in this field is to determine lower and upper bounds on the number of realizations with respect to a given number of vertices. This problem is closely related to the classification of rigid graphs according to their maximal number of real embeddings. In [17], we are interested in finding edge lengths that can maximize the number of real embeddings of minimally rigid graphs in the plane, space, and on the sphere. We use algebraic formulations to provide upper bounds. To find values of the parameters that lead to graphs with a large number of real realizations, possibly attaining the (algebraic) upper bounds, we use some standard heuristics and we also develop a new method inspired by coupler curves. We apply this new method to obtain embeddings in $\mathbb{R}^3$. One of its main novelties is that it allows us to sample efficiently from a larger number of parameters by selecting only a subset of them at each iteration. Our results include a full classification of the 7-vertex graphs according to their maximal numbers of real embeddings in the cases of the embeddings in $\mathbb{R}^2$ and $\mathbb{R}^3$, while in the case of S^2 we achieve this classification for all 6-vertex graphs. Additionally, by increasing the number of embeddings of selected graphs, we improve the previously known asymptotic lower bound on the maximum number of realizations.

7.15. Multilinear Polynomial Systems: Root Isolation and Bit Complexity

In [20], we exploit structure in polynomial system solving by considering polyno-mials that are linear in subsets of the variables. We focus on algorithms and their Boolean complexity for computing isolating hyperboxes for all the isolated complex roots of well-constrained, unmixed systems of multilinear polynomials based on resultant methods. We enumerate all expressions of the multihomogeneous (or multigraded) resultant of such systems as a determinant of Sylvester-like matrices, aka generalized Sylvester matrices. We construct these matrices by means of Weyman homological complexes, which generalize the Cayley-Koszul complex. The computation of the determinant of the resultant matrix is the bottleneck for the overall complexity. We exploit the quasi-Toeplitz structure to reduce the problem to efficient matrix-vector multiplication, which corresponds to multivariate polynomial multiplication, by extending the seminal work on Macaulay matrices of Canny, Kaltofen, and Yagati [9] to the multi-homogeneous case. We compute a rational univariate representation of the roots, based on the primitive element method. In the case of 0-dimensional systems we present a Monte Carlo algorithm with probability of success $1 - 1/2^{\eta}$, for a given $\eta \geq 1$, and bit complexity $O_B(n^2 D^{4+\epsilon}(n^{N+1} + \tau) + nD^{2+\epsilon}\eta(D + \eta))$ for any $\epsilon > 0$, where n is the number of variables, D equals the multilinear Bézout bound, N is the number of variable subsets, and τ is the maximum coefficient bitsize. We present an algorithmic variant to compute the isolated roots of overdetermined and positive-dimensional systems. Thus our algorithms and complexity analysis apply in general with no assumptions on the input.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- The objective of our Agrement with WATERLOO MAPLE INC. is to promote software developments to which we actively contribute.

  On the one hand, WMI provides man power, software licenses, technical support (development, documentation and testing) for an inclusion of our developments in their commercial products. On the other hand, OURAGAN offers perpetual licenses for the use of the concerned source code.

  As past results of this agreement one can cite our C-Library *RS* for the computations of the real solutions zero-dimensional systems or also our collaborative development around the Maple package *DV* for solving parametric systems of equations.

  For this term, the agreement covers algorithms developed in areas including but not limited to: 1) solving of systems of polynomial equations, 2) validated numerical polynomial root finding, 3) computational geometry, 4) curves and surfaces topology, 5) parametric algebraic systems, 6) cylindrical algebraic decompositions, 7) robotics applications.

  In particular, it covers our collaborative work with some of our partners, especially the Gamble Project-Team - Inria Nancy Grand Est.
- In 2019, a contract was signed with the company *Safran Tech*. Its goal is to bring our scientific expertise on mathematical and algorithmic aspects on certain problems studied in gearbox vibration analysis. Gear fault diagnosis is an important issue in aeronautics industry since a damage in a gearbox, which is not detected in time, can have dramatic effects on the safety of a plane. Since the vibrations of a spur gear can be modeled as a product of two periodic functions related to the gearbox kinematic, [92] has proposed to recover each function from the global signal by means of an optimal reconstruction problem which, by means of Fourier analysis, yields a Frobenius norm minimization problem for structured matrices. The goal of the collaboration is to use symbolic-numeric to study this problem.

9. Partnerships and Cooperations

9.1. National Initiatives

- FMJH Program, PGMO grant

  ALMA (Algebraic methods in games and optimization).

  Duration: 2018 – 2020. (2 years project)

  Coordinator: Elias Tsigaridas, with Stéphane Gaubert and Xavier Allamigeon (CMAP, École Polytechnique)

9.1.1. ANR

- ANR JCJC GALOP (Games through the lens of ALgebra and OPptimization)

  Coordinator: Elias Tsigaridas

  Duration: 2018 – 2022

  GALOP is a Young Researchers (JCJC) project with the purpose of extending the limits of the state-of-the-art algebraic tools in computer science, especially in stochastic games. It brings original and innovative algebraic tools, based on symbolic-numeric computing, that exploit the geometry and the structure and complement the state-of-the-art. We support our theoretical tools with a highly efficient open-source software for solving polynomials. Using our algebraic tools we study the geometry of the central curve of (semi-definite) optimization problems. The algebraic tools and our results from the geometry of optimization pave the way to introduce algorithms and precise bounds for stochastic games.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

Program: H2020-EU.1.1. - EXCELLENT SCIENCE - European Research Council (ERC)

Project acronym: Almacrypt

Project title: Algorithmic and Mathematical Cryptology

Duration: 01/2016 - 12/2010

Coordinator: Antoine Joux

Abstract: Cryptology is a foundation of information security in the digital world. Today's internet is protected by a form of cryptography based on complexity theoretic hardness assumptions. Ideally, they should be strong to ensure security and versatile to offer a wide range of functionalities and allow efficient implementations. However, these assumptions are largely untested and internet security could be built on sand. The main ambition of Almacrypt is to remedy this issue by challenging the assumptions through an advanced algorithmic analysis. In particular, this proposal questions the two pillars of public-key encryption: factoring and discrete logarithms. Recently, the PI contributed to show that in some cases, the discrete logarithm problem is considerably weaker than previously assumed. A main objective is to ponder the security of other cases of the discrete logarithm problem, including elliptic curves, and of factoring. We will study the generalization of the recent techniques and search for new algorithmic options with comparable or better efficiency. We will also study hardness assumptions based on codes and subset-sum, two candidates for post-quantum cryptography. We will consider the applicability of recent algorithmic and mathematical techniques to the resolution of the corresponding putative hard problems, refine the analysis of the algorithms and design new algorithm tools. Cryptology is not limited to the above assumptions: other hard problems have been proposed to aim at post-quantum security and/or to offer extra functionalities. Should the security of these other assumptions become critical, they would be added to Almacrypt's scope. They could also serve to demonstrate other applications of our algorithmic progress. In addition to its scientific goal, Almacrypt also aims at seeding a strengthened research community dedicated to algorithmic and mathematical cryptology.

9.3. International Initiatives

- Partenariat Hubert Curien franco-turc (PHC Bosphore) with Gebze Technical University, Turkey.

  Title: "Gröbner bases, ResultAnts and Polyhedral gEometry" (GRAPE)

  Duration: 2019 – 2020 (2 years project)

  Coordinator: Elias Tsigaridas

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

9.3.1.1. MACAO

Title: Mathematics and Algorithms for Cryptographic Advanced Objects

International Partner (Institution - Laboratory - Researcher):

University of Wollongong (Australia) - Thomas Plantard

Start year: 2019

See also: https://ssl.informatics.uow.edu.au/MACAO/

Since quantum computers have the ability to break the two main problems on which current public cryptography relies, i.e., the factoring and discrete logarithm problem, every step towards the practical realization of these computers raises fears about potential attacks on cryptographic systems. By scrutinizing the techniques proposed to build post-quantum cryptography, we can identify a few candidate hard problems which underly the proposals. One objective of this international project is to precisely assess the security of these cryptographic algorithms. First, by analyzing in a systematic manner the existing resolution algorithms and by assessing their complexity as a function of security parameters. Then, we will consider new algorithmic techniques to solve these candidate hard Post-Quantum problems, both on classical computers and quantum machines aiming at the discovery of new and better algorithms to solve them.

9.3.2. Inria International Partners

9.3.2.1. Declared Inria International Partners

- University of Wollongong (Australia)

9.3.2.2. Informal International Partners

- CQT Singapour (UMI CNRS Majulab)
- UFPA - Para -Brésil (José Miguel Veloso)
- Institut Joseph Fourier - Université Grenoble Alpes (Martin Deraux, V. Vitse et Pierre Will)
- Max-Planck-Institut für Informatik - Saarbrücken - Germany (Alex. Kobel)
- Holon Institute of Technology, Israel (Jeremy Kaminsky)
- Department of Informatics, National Kapodistrian University of Athens, Greece (Ioannis Emiris)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- Elias Tsigaridas was the program chair of Mathematical Aspects of Computer and Information Sciences (MACIS 2019), which was held at Istanbul, 13-15 November 2019.

10.1.1.2. Member of the Organizing Committees

A. Quadrat was a member of the organization committee of the *Journées Nationales de Calcul Formel* (JNCF), Luminy, France, 04-08/02/2019. He is now a member of the scientific committee of JNCF.

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

Elias Tsigaridas was a program committee member of Computer Algebra in Scientific Computing (CASC 2019).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Elisha Falbel is a member of the editorial board of *São Paulo Journal of Mathematical Sciences - Springer*
- Antoine Joux is a member of the editorial board of *Designs, Codes and Cryptography*
- Alban Quadrat is associate editor of *Multidimensional Systems and Signal Processing*, Springer
- Fabrice Rouillier is a member of the editorial board of *Journal of Symbolic Computation*

10.1.4. Leadership within the Scientific Community

Alban Quadrat co-organized the invited sessions *Algebraic and symbolic methods for mathematical systems theory* and *New trends in computational methods for time-delay systems* at the Join *Joint IFAC Conference 7th Symposium on Systems Structure and Control and 15th Workshop on Time Delay Systems*, 9–11 September 2019, Sinaia, Romania

Elias Tsigaridas organized a session on *Algebraic and geometric tools for optimisation and statistics* during the PGMO days 2019, at the EDF'Lab Palaiseau, December 2019.

10.1.5. Research Administration

- Alban Quadrat is a member of the Conseil d'Administration of the Société Mathématique de France (SMF).
- Fabrice Rouillier is a member of the scientific commitee of the Indo French Centre for Applied Mathematics.
- Elisha Falbel is director of the "École Doctorale Sciences Mathématiques de Paris Centre - ED 386".

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Fabrice Rouillier: Course in Algebraic Computations, M1, 24h, Sorbonne Université
- Fabrice Rouillier: Course in "Agrégation Option - C", M2, 31 heures, Sorbonne Université
- Antonin Guilloux: Courses and administration of the general math courses in L1 at Sorbonne Université
- Elisha Falbel: First year 'mathématiques approfondie' and master class on Riemann surfaces.
- Pierre-Vincent Koseleff : head of the Master 2 "pr éparation à l'agrégation de mathématiques" at Sorbonne Université.
- Pierre-Vincent Koseleff : Course in M2 "Agrégation de Mathématiques", Computer Algebra, Option C, 100h. Sorbonne Université.
- Pierre-Vincent Koseleff : Course in L3, "algèbre appliquée".

10.2.2. Supervision

PhD in progress: Thomas Espitau, 09/2016, directed by Antoine Joux

PhD in progress: Natalia Kharchenko, 09/2016, directed by Antoine Joux

PhD in progress: Mahya Mehrabdollahei, 09/2018, directed by Antonin Guilloux and Fabrice Rouillier

PhD in progress: Grace Younes, 09/2018, directed by Alban Quadrat and Fabrice Rouillier

PhD in progress: Christina Katsamaki, 09/2019, directed by Elias Tsigaridas and Fabrice Rouillier

PhD in progress: Sudarshan Shinde, 09/2016, directed by Razvan Barbulescu and Pierre-Vincent Koseleff

PhD in progress: Raphael Alexandre, 09/2019, directed by Elisha Falbel.

10.2.3. Juries

Alban Quadrat was a member of the PhD defense committee of Elisa Hubert, *Surveillance vibratoire d'une transmission de puissance aéronautique*, University of Lyon & Safran Tech, 28/06/2019.

Elisha Falbel was a member of the Habilitation defense committee of Junyan Cao, *Positivité des images directees, extension d'Ohsawa-Takegoshi et applications*, Sorbonne Université , 25/10/2019.

10.3. Popularization

- Alban Quadrat co-edited the book *Le Jeu de rôle sur table : un laboratoire de l'imaginaire*, Carrefour des lettres modernes, n. 7, Classiques Garnier, 2019.
- Fabrice Rouillier is a member of the editorial board of *Interstices*
- Fabrice Rouillier is *chargé de mission médiation* at Inria Paris
- Fabrice Rouillier is a member of the *comité de pilotage de l'année des mathématiques*
- Fabrice Rouillier is a member of the *comité de pilotage de la semaine des mathématiques*
- Fabrice Rouillier is the president of the association *Animath*[0]
- Fabrice Rouillier is member of the scientific board of the SMF [0]conference *Mathématiques étonnantes*[0]

10.3.1. Internal or external Inria responsibilities

- Alban Quadrat is a member of the technical committee *Linear Systems* of the *International Federation of Automatic Control* (IFAC)
- Elias Tsigaridas is a member of the Evaluation committee (Commission d'Évaluation) of Inria.

10.3.2. Interventions

- Fabrice Rouillier participated to the *semaine des mathématiques* for the Paris Inria Paris research center
- Fabrice Rouillier participated to the welcoming of schoolchildren for their *semaine d'observation* at Inria Paris research center.

11. Bibliography

Major publications by the team in recent years

[1] Y. Bouzidi, S. Lazard, G. Moroz, M. Pouget, F. Rouillier, M. Sagraloff.*Solving bivariate systems using Rational Univariate Representations*, in "Journal of Complexity", 2016, vol. 37, p. 34–75 [*DOI :* 10.1016/j.jco.2016.07.002], https://hal.inria.fr/hal-01342211

[2] E. Brugallé, P.-V. Koseleff, D. Pecker.*On the lexicographic degree of two-bridge knots*, in "Journal Of Knot Theory And Its Ramifications (JKTR)", June 2016, vol. 25, n^o 7, 14p., 21 figs [*DOI :* 10.1142/S0218216516500449], https://hal.archives-ouvertes.fr/hal-01084472

[3] E. Brugallé, P.-V. Koseleff, D. Pecker.*Untangling trigonal diagrams*, in "Journal Of Knot Theory And Its Ramifications (JKTR)", June 2016, vol. 25, n^o 7, 10p., 24 figs [*DOI :* 10.1142/S0218216516500437], https://hal.archives-ouvertes.fr/hal-01084463

[4] F. Chyzak, A. Quadrat, D. Robertz.*Effective algorithms for parametrizing linear control systems over Ore algebras*, in "Applicable Algebra in Engineering, Communications and Computing", 2005, vol. 16, p. 319–376

[0] http://www.animath.fr
[0] Société Mathématique de France
[0] https://smf.emath.fr/actualites-smf/mathematiques-etonnantes-programme

[5] T. Cluzeau, A. Quadrat. *Factoring and decomposing a class of linear functional systems*, in "Linear Algebra and Its Applications", 2008, vol. 428, p. 324–381

[6] E. Falbel, A. Guilloux. *Dimension of character varieties for 3-manifolds*, in "Proceedings of the American Mathematical Society", 2016 [*DOI :* 10.1090/proc/13394], https://hal.archives-ouvertes.fr/hal-01370284

[7] E. Falbel, A. Guilloux, P.-V. Koseleff, F. Rouillier, M. Thistlethwaite. *Character Varieties For SL(3,C): The Figure Eight Knot*, in "Experimental Mathematics", 2016, vol. 25, n^o 2, 17 [*DOI :* 10.1080/10586458.2015.1068249], https://hal.inria.fr/hal-01362208

[8] E. Falbel, J. Wang. *Branched spherical CR structures on the complement of the figure-eight knot*, in "Michigan Mathematical Journal", 2014, vol. 63, p. 635-667, https://hal.archives-ouvertes.fr/hal-01374789

[9] A. Joux. *A one round protocol for tripartite Diffie-Hellman*, in "J. Cryptology", 2004, vol. 17, n^o 4, p. 263–276

[10] A. Joux, R. Lercier. *Improvements to the general number field sieve for discrete logarithms in prime fields. A comparison with the gaussian integer method*, in "Math. Comput.", 2003, vol. 72, n^o 242, p. 953-967

[11] D. Lazard, F. Rouillier. *Solving Parametric Polynomial Systems*, in "Journal of Symbolic Computation", June 2007, vol. 42, p. 636-667

[12] A. Quadrat, D. Robertz. *Computation of bases of free modules over the Weyl algebras*, in "Journal of Symbolic Computation", 2007, vol. 42, p. 1113–1141

[13] F. Rouillier. *Solving zero-dimensional systems through the rational univariate representation*, in "Journal of Applicable Algebra in Engineering, Communication and Computing", 1999, vol. 9, n^o 5, p. 433–461

[14] F. Rouillier, P. Zimmermann. *Efficient Isolation of Polynomial Real Roots*, in "Journal of Computational and Applied Mathematics", 2003, vol. 162, n^o 1, p. 33–50

Publications of the year

Articles in International Peer-Reviewed Journal

[15] R. Barbulescu, S. Duquesne. *Updating key size estimations for pairings*, in "Journal of Cryptology", 2019, vol. 32, n^o 4, p. 1298–1336 [*DOI :* 10.1007/s00145-018-9280-5], https://hal.archives-ouvertes.fr/hal-01534101

[16] R. Barbulescu, J. Ray. *Numerical verification of the Cohen-Lenstra-Martinet heuristics and of Greenberg's p-rationality conjecture*, in "Journal de Théorie des Nombres de Bordeaux", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-01534050

[17] E. Bartzos, I. Z. Emiris, J. Legerský, E. Tsigaridas. *On the maximal number of real embeddings of minimally rigid graphs in* $\mathbb{R}^2$, $\mathbb{R}^3$ *and* S^2, in "Journal of Symbolic Computation", 2019, https://arxiv.org/abs/1811.12800 , forthcoming [*DOI :* 10.1016/j.jsc.2019.10.015], https://hal.archives-ouvertes.fr/hal-02271782

[18] Y. M. BOUZIDI, A. QUADRAT, F. ROUILLIER.*Certified Non-conservative Tests for the Structural Stability of Discrete Multidimensional Systems*, in "Multidimensional Systems and Signal Processing", July 2019, vol. 30, n^{o} 3, 31 [*DOI :* 10.1007/S11045-018-0596-Y], https://hal.inria.fr/hal-01951765

[19] L. BUSÉ, A. MANTZAFLARIS, E. TSIGARIDAS.*Matrix formulae for Resultants and Discriminants of Bivariate Tensor-product Polynomials*, in "Journal of Symbolic Computation", June 2020, vol. 98, p. 65-83 [*DOI :* 10.1016/J.JSC.2019.07.007], https://hal.inria.fr/hal-01654263

[20] I. Z. EMIRIS, A. MANTZAFLARIS, E. TSIGARIDAS.*Multilinear Polynomial Systems: Root Isolation and Bit Complexity*, in "Journal of Symbolic Computation", 2019, Special Issue of the Journal of Symbolic Computation on Milestones in Computer Algebra (MICA 2016), forthcoming, https://hal.inria.fr/hal-02099556

[21] I. Z. EMIRIS, B. MOURRAIN, E. TSIGARIDAS.*Separation bounds for polynomial systems*, in "Journal of Symbolic Computation", 2019 [*DOI :* 10.1016/J.JSC.2019.07.001], https://hal.inria.fr/hal-01105276

[22] T. ESPITAU, A. JOUX.*Certified lattice reduction*, in "Advances in Mathematics of Communications", February 2020, vol. 14, n^{o} 1, p. 137-159 [*DOI :* 10.3934/AMC.2020011], https://hal.archives-ouvertes.fr/hal-02383752

[23] F. GÖLOĞLU, A. JOUX.*A simplified approach to rigorous degree 2 elimination in discrete logarithm algorithms*, in "Mathematics of Computation", 2019, 1, https://hal.archives-ouvertes.fr/hal-01960765

[24] P. MOLIN, C. NEUROHR.*Computing period matrices and the Abel-Jacobi map of superelliptic curves*, in "Mathematics of Computation", 2019, https://hal.inria.fr/hal-02416012

International Conferences with Proceedings

[25] Y. BOUZIDI, T. CLUZEAU, A. QUADRAT.*On the computation of stabilizing controllers of multidimensional systems*, in "SSSC 2019 - 7th IFAC Symposium on Systems Structure and Control", Sinai, Romania, September 2019 [*DOI :* 10.1016/J.IFACOL.2019.11.032], https://hal.inria.fr/hal-02419696

[26] Y. BOUZIDI, T. CLUZEAU, A. QUADRAT, F. ROUILLIER.*On the effective computation of stabilizing controllers of 2D systems*, in "Maple Conference", Waterloo, Canada, October 2019, https://hal.inria.fr/hal-02419719

[27] D. CHABLAT, G. MOROZ, F. ROUILLIER, P. WENGER.*Using Maple to analyse parallel robots*, in "Maple Conference 2019", Waterloo, Canada, October 2019, https://hal.inria.fr/hal-02406703

[28] I. Z. EMIRIS, C. KATSAMAKI.*Voronoi diagram of orthogonal polyhedra in two and three dimensions*, in "SEA 2019 - Symposium on Experimental Algorithms", Kalamata, Greece, I. KOTSIREAS, P. PARDALOS, K. E. PARSOPOULOS, D. SOURAVLIAS, A. TSOKAS (editors), LNCS - Lecture Notes in Computer Science, Springer, June 2019, vol. 11544 [*DOI :* 10.1007/978-3-030-34029-2_1], https://hal.inria.fr/hal-02398736

[29] E. HUBERT, Y. BOUZIDI, R. DAGHER, A. BARRAU, A. QUADRAT.*Algebraic aspects of the exact signal demodulation problem*, in "SSSC 2019 - 7th IFAC Symposium on Systems Structure and Control", Sinaia, Romania, September 2019 [*DOI :* 10.1016/J.IFACOL.2019.11.031], https://hal.inria.fr/hal-02419824

Scientific Books (or Scientific Book chapters)

[30] Y. BOUZIDI, A. POTEAUX, A. QUADRAT.*A symbolic computation approach to the asymptotic stability analysis of differential systems with commensurate delays*, in "Delays and Interconnections: Methodology, Algorithms and Applications, Advances in Delays and Dynamics", G. VALMORBIDA, A. SEURET, I. BOUSSAADA, R. SIPAHI (editors), Springer, October 2019, vol. 10, 16 [*DOI :* 10.1007/978-3-030-11554-8_11], https://hal.inria.fr/hal-01485536

[31] T. CLUZEAU, C. KOUTSCHAN, A. QUADRAT, M. TÕNSO.*Effective algebraic analysis approach to linear systems over Ore algebras*, in "Algebraic and Symbolic Computation Methods in Dynamical Systems", Advances in Delays and Dynamics, Springer, 2020, https://hal.archives-ouvertes.fr/hal-02436985

[32] T. CLUZEAU, A. QUADRAT.*Equivalence of Linear Functional Systems*, in "Algebraic and Symbolic Computation Methods in Dynamical Systems", Advances in Delays and Dynamics, Springer, 2020, https://hal.archives-ouvertes.fr/hal-02436978

Other Publications

[33] J. G. ALCÁZAR, J. CARAVANTES, G. M. DIAZ-TOCA, E. TSIGARIDAS.*Computing the topology of a planar or space hyperelliptic curve*, January 2019, working paper or preprint, https://hal.inria.fr/hal-01968776

[34] R. BARBULESCU, N. EL MRABET, L. GHAMMAM.*A taxonomy of pairings, their security, their complexity*, May 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02129868

[35] R. BARBULESCU, S. SHINDE.*A classification of ECM-friendly families using modular curves*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01822144

[36] R. DAGHER, G. ZHENG, A. QUADRAT.*General closed-form solutions of the position self-calibration problem*, October 2019, Paper under submission, https://hal.inria.fr/hal-02419854

[37] G. IVANYOS, A. JOUX, M. SANTHA.*Discrete logarithm and Diffie-Hellman problems in identity black-box groups*, November 2019, https://arxiv.org/abs/1911.01662 - working paper or preprint, https://hal.sorbonne-universite.fr/hal-02350271

[38] A. JOUX, C. PIERROT.*Algorithmic aspects of elliptic bases in finite field discrete logarithm algorithms*, July 2019, https://arxiv.org/abs/1907.02689 - working paper or preprint, https://hal.sorbonne-universite.fr/hal-02173688

References in notes

[39] A. K. LENSTRA, H. W. LENSTRA (editors). *The development of the number field sieve*, Lecture Notes in Mathematics, Springer-Verlag, 1993, vol. 1554

[40] D. AGGARWAL, A. JOUX, A. PRAKASH, M. SANTHA.*A New Public-Key Cryptosystem via Mersenne Numbers*, in "Advances in Cryptology - CRYPTO 2018 - 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2018, Proceedings, Part III", 2018, p. 459–482, https://doi.org/10.1007/978-3-319-96878-0_16

[41] S. BASU, R. POLLACK, M.-F. ROY.*Algorithms in Real Algebraic Geometry (Algorithms and Computation in Mathematics)*, Springer-Verlag, Berlin, Heidelberg, 2006

[42] V. BAVULA.*The algebra of integro-differential operators on an affine line and its modules*, in "J. Pure Appl. Algebra", 2013, vol. 217, p. 495–529

[43] N. BERGERON, E. FALBEL, A. GUILLOUX.*Tetrahedra of flags, volume and homology of SL(3)*, in "Geometry & Topology Monographs", 2014, vol. 18 [*DOI :* 10.2140/GT.2014.18.1911], https://hal.archives-ouvertes.fr/hal-01370258

[44] J.-F. BIASSE, T. ESPITAU, P.-A. FOUQUE, A. GÉLIN, P. KIRCHNER.*Computing generator in cyclotomic integer rings*, in "36th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT 2017)", Paris, France, Lecture Notes in Computer Science, April 2017, vol. 10210, p. 60-88 [*DOI :* 10.1007/978-3-319-56620-7_3], https://hal.archives-ouvertes.fr/hal-01518438

[45] A. BOREL.*Algebraic D-modules*, Perspectives in mathematics, Academic Press, 1987

[46] N. BOSE.*Multidimensional Systems Theory: Progress, Directions and Open Problems in Multidimensional Systems*, Mathematics and Its Applications, Springer Netherlands, 2001

[47] F. BOULIER, D. LAZARD, F. OLLIVIER, M. PETITOT.*Computing representations for radicals of finitely generated differential ideals*, in "Applicable Algebra in Engineering, Communication and Computing", 2009, vol. 20, p. 73–121

[48] Y. BOUZIDI, T. CLUZEAU, G. MOROZ, A. QUADRAT.*Computing effectively stabilizing controllers for a class of nD systems*, in "The 20th World Congress of the International Federation of Automatic Control", Toulouse, France, July 2017, vol. 50, n^{o} 1, p. 1847 – 1852 [*DOI :* 10.1016/J.IFACOL.2017.08.200], https://hal.archives-ouvertes.fr/hal-01667161

[49] Y. BOUZIDI, S. LAZARD, G. MOROZ, M. POUGET, F. ROUILLIER.*Improved algorithm for computing separating linear forms for bivariate systems*, in "ISSAC - 39th International Symposium on Symbolic and Algebraic Computation", Kobe, Japan, July 2014, https://hal.inria.fr/hal-00992634

[50] Y. BOUZIDI, S. LAZARD, G. MOROZ, M. POUGET, F. ROUILLIER, M. SAGRALOFF.*Solving bivariate systems using Rational Univariate Representations*, in "Journal of Complexity", 2016, vol. 37, p. 34–75 [*DOI :* 10.1016/J.JCO.2016.07.002], https://hal.inria.fr/hal-01342211

[51] Y. BOUZIDI, S. LAZARD, M. POUGET, F. ROUILLIER.*Separating linear forms and Rational Univariate Representations of bivariate systems*, in "Journal of Symbolic Computation", May 2015, vol. 68, n^{o} 0, p. 84-119 [*DOI :* 10.1016/J.JSC.2014.08.009], https://hal.inria.fr/hal-00977671

[52] Y. BOUZIDI, A. POTEAUX, A. QUADRAT.*A symbolic computation approach to the asymptotic stability analysis of differential systems with commensurate delays*, in "Delays and Interconnections: Methodology, Algorithms and Applications", Advances on Delays and Dynamics at Springer, Springer Verlag, March 2017, https://hal.inria.fr/hal-01485536

[53] Y. BOUZIDI, A. QUADRAT, F. ROUILLIER.*Computer algebra methods for testing the structural stability of multidimensional systems*, in " IEEE 9th International Workshop on Multidimensional (nD) Systems (IEEE nDS 2015)", Vila Real, Portugal, Proceedings of the IEEE 9th International Workshop on Multidimensional (nD) Systems (IEEE nDS 2015), September 2015, https://hal-centralesupelec.archives-ouvertes.fr/hal-01259968

[54] Y. BOUZIDI, A. QUADRAT, F. ROUILLIER.*Certified Non-conservative Tests for the Structural Stability of Multidimensional Systems*, August 2017, 31, To appear in Multidimensional Systems and Signal Processing, https://link.springer.com/article/10.1007/s11045-018-0596-y, https://hal.inria.fr/hal-01571230

[55] Y. BOUZIDI, F. ROUILLIER.*Certified Algorithms for proving the structural stability of two dimensional systems possibly with parameters*, in "MNTS 2016 - 22nd International Symposium on Mathematical Theory of Networks and Systems", Minneapolis, United States, Proceedings of the 22nd International Symposium on Mathematical Theory of Networks and Systems, July 2016, https://hal.inria.fr/hal-01366202

[56] E. BRUGALLÉ, P.-V. KOSELEFF, D. PECKER.*On the lexicographic degree of two-bridge knots*, in "Journal Of Knot Theory And Its Ramifications (JKTR)", June 2016, vol. 25, n^{o} 7, 14p., 21 figs [*DOI :* 10.1142/S0218216516500449], https://hal.archives-ouvertes.fr/hal-01084472

[57] E. BRUGALLÉ, P.-V. KOSELEFF, D. PECKER.*Untangling trigonal diagrams*, in "Journal Of Knot Theory And Its Ramifications (JKTR)", June 2016, vol. 25, n^{o} 7, 10p., 24 figs [*DOI :* 10.1142/S0218216516500437], https://hal.archives-ouvertes.fr/hal-01084463

[58] E. BRUGALLÉ, P.-V. KOSELEFF, D. PECKER.*The lexicographic degree of the first two-bridge knots*, September 2018, 30 p., 58 fig., 6 tables, submitted, https://hal.archives-ouvertes.fr/hal-01108678

[59] D. CHABLAT, R. JHA, F. ROUILLIER, G. MOROZ.*Non-singular assembly mode changing trajectories in the workspace for the 3-RPS parallel robot*, in "14th International Symposium on Advances in Robot Kinematics", Ljubljana, Slovenia, June 2014, p. 149 – 159, https://hal.archives-ouvertes.fr/hal-00956325

[60] D. CHABLAT, R. JHA, F. ROUILLIER, G. MOROZ.*Workspace and joint space analysis of the 3-RPS parallel robot*, in "ASME 2013 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference", Buffalo, United States, August 2014, vol. Volume 5A, p. 1-10, https://hal.archives-ouvertes.fr/hal-01006614

[61] F. CHYZAK, A. QUADRAT, D. ROBERTZ.*Effective algorithms for parametrizing linear control systems over Ore algebras*, in "Applicable Algebra in Engineering, Communications and Computing", 2005, vol. 16, p. 319–376

[62] F. CHYZAK, B. SALVY.*Non-commutative elimination in Ore algebras proves multivariate identities*, in "Journal of Symbolic Computation", 1998, vol. 26, n^{o} 2, p. 187–227

[63] G. E. COLLINS.*Quantifier elimination for real closed fields by cylindrical algebraic decompostion*, in "Automata Theory and Formal Languages 2nd GI Conference Kaiserslautern, May 20–23, 1975", Berlin, Heidelberg, H. BRAKHAGE (editor), Springer Berlin Heidelberg, 1975, p. 134–183

[64] D. A. COX, J. LITTLE, D. O'SHEA.*Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3/e (Undergraduate Texts in Mathematics)*, Springer-Verlag, Berlin, Heidelberg, 2007

[65] M. CROCCO, A. DEL BUE, V. MURINO.*A bilinear approach to the position self-calibration of multiple sensors*, in "IEEE Transactions on Signal Processing", 2012, vol. 60, n^{o} 2, p. 660–673

[66] R. CURTAIN, H. ZWART.*An Introduction to Infinite-Dimensional Linear Systems Theory*, Texts in Applied Mathematics, Springer New York, 2012

[67] R. DAGHER, A. QUADRAT, G. ZHENG.*Auto-localisation par mesure de distances*, in "Pattern n. FR1853553", 2018

[68] R. DAGHER, A. QUADRAT, G. ZHENG.*Algebraic solutions to the metric multidimensional unfolding. Application to the position self-calibration problem*, in "in preparation", 2019

[69] M. DERAUX, E. FALBEL.*Complex hyperbolic geometry of the figure eight knot*, in "Geometry and Topology", February 2015, vol. 19, p. 237–293 [*DOI :* 10.2140/GT.2015.19.237], https://hal.archives-ouvertes.fr/hal-00805427

[70] D. N. DIATTA, S. DIATTA, F. ROUILLIER, M.-F. ROY, M. SAGRALOFF.*Bounds for polynomials on algebraic numbers and application to curve topology*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01891417

[71] W. DIFFIE, M. E. HELLMAN.*New directions in cryptography*, in "IEEE Transactions on Information Theory", 1976, vol. 22, n^{o} 6, p. 644-654

[72] S. DIOP.*Elimination in control theory*, in "Math. Control Signals Systems", 1991, vol. 4, p. 17–32

[73] S. DIOP.*Differential-algebraic decision methods and some applications to system theory*, in "Theoret. Comput. Sci.", 1992, vol. 98, p. 137–161

[74] J. DOLISKANI, A. K. NARAYANAN, É. SCHOST.*Drinfeld Modules with Complex Multiplication, Hasse Invariants and Factoring Polynomials over Finite Fields*, in "CoRR", 2017, vol. abs/1712.00669, http://arxiv.org/abs/1712.00669

[75] T. ESPITAU, A. JOUX.*Adaptive precision LLL and Potential-LLL reductions with Interval arithmetic*, in "IACR Cryptology ePrint Archive", 2016, vol. 2016, 528, http://eprint.iacr.org/2016/528

[76] H. EVELYNE.*Notes on Triangular Sets and Triangulation-Decomposition Algorithms II: Differential Systems*, in "Symbolic and Numerical Scientific Computation", F. WINKLER, U. LANGER (editors), Lecture Notes in Computer Science 2630, Springer, 2003, p. 40-87

[77] E. FALBEL, A. GUILLOUX.*Dimension of character varieties for 3-manifolds*, in "Proceedings of the American Mathematical Society", 2016 [*DOI :* 10.1090/PROC/13394], https://hal.archives-ouvertes.fr/hal-01370284

[78] E. FALBEL, A. GUILLOUX, P. WILL.*Hilbert metric, beyond convexity*, 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01768400

[79] E. FALBEL, P.-V. KOSELEFF, F. ROUILLIER.*Representations of fundamental groups of 3-manifolds into PGL(3,C): Exact computations in low complexity*, in "Geometriae Dedicata", August 2015, vol. 177, n^{o} 1, 52 [*DOI :* 10.1007/S10711-014-9987-X], https://hal.inria.fr/hal-00908843

[80] E. FALBEL, M. MACULAN, G. SARFATTI.*Configurations of flags in orbits of real forms*, April 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01779459

[81] E. FALBEL, R. SANTOS THEBALDI.*A Flag structure on a cusped hyperbolic 3-manifold with unipotent holonomy*, in "Pacific Journal of Mathematics", 2015, vol. 278, n^{o} 1, p. 51-78, https://hal.archives-ouvertes.fr/hal-00958255

[82] E. FALBEL, J. VELOSO.*Flag structures on real 3-manifolds*, April 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01778582

[83] J. FAUGÈRE, D. LAZARD.*Combinatorial classes of parallel manipulators*, in "Mechanism and Machine Theory", 1995, vol. 30, n^{o} 6, p. 765 – 776 [*DOI :* 10.1016/0094-114X(94)00069-W], http://www.sciencedirect.com/science/article/pii/0094114X9400069W

[84] M. FLIESS, H. SIRA-RAMIREZ.*An algebraic framework for linear identification*, in "ESAIM Control Optim. Calc. Variat.", 2003, vol. 9, p. 151—168

[85] J. V. Z. GATHEN, J. GERHARD.*Modern Computer Algebra*, 3rd, Cambridge University Press, New York, NY, USA, 2013

[86] A. GUILLOUX.*Volume of representations and birationality of peripheral holonomy*, in "Experimental Mathematics", May 2017, https://hal.archives-ouvertes.fr/hal-01370287

[87] A. GUILLOUX, I. KIM.*Deformation space of discrete groups of SU(2,1) in quaternionic hyperbolic plane*, March 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01736953

[88] A. GUILLOUX, J. MARCHÉ.*Volume function and Mahler measure of exact polynomials*, April 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01758986

[89] A. GUILLOUX, P. WILL.*On SL(3,C)-representations of the Whitehead link group*, 2018, To appear in Geom. Ded, https://hal.archives-ouvertes.fr/hal-01370289

[90] A. GÉLIN, A. JOUX.*Reducing number field defining polynomials: an application to class group computations*, in "Algorithmic Number Theory Symposium XII", Kaiserslautern, Germany, LMS Journal of Computation and Mathematics, August 2016, vol. 19, n^{o} A, p. 315–331 [*DOI :* 10.1112/S1461157016000255], https://hal.archives-ouvertes.fr/hal-01362144

[91] F. GÖLOĞLU, A. JOUX.*A Simplified Approach to Rigorous Degree 2 Elimination in Discrete Logarithm Algorithms*, in "IACR Cryptology ePrint Archive", 2018, vol. 2018, 430, https://eprint.iacr.org/2018/430

[92] E. HUBERT, A. BARRAU, M. EL BADAOUI.*New Multi-Carrier Demodulation Method Applied to Gearbox Vibration Analysis*, 04 2018, p. 2141-2145 [*DOI :* 10.1109/ICASSP.2018.8461924]

[93] M. L. HUSTY, H.-P. SCHRÖCKER.*Algebraic Geometry and Kinematics*, in "Nonlinear Computational Geometry", New York, NY, I. Z. EMIRIS, F. SOTTILE, T. THEOBALD (editors), Springer New York, 2010, p. 85–107

[94] M. JANET.*Leçons sur les systèmes d'équations aux dérivées partielles*, Gauthier-Villars, 1929

[95] R. JHA, D. CHABLAT, L. BARON, F. ROUILLIER, G. MOROZ.*Workspace, Joint space and Singularities of a family of Delta-Like Robot*, in "Mechanism and Machine Theory", September 2018, vol. 127, p. 73-95 [*DOI :* 10.1016/J.MECHMACHTHEORY.2018.05.004], https://hal.archives-ouvertes.fr/hal-01796066

[96] R. JHA, D. CHABLAT, F. ROUILLIER, G. MOROZ.*An algebraic method to check the singularity-free paths for parallel robots*, in "International Design Engineering Technical Conferences & Computers and Information in Engineering Conference", Boston, United States, ASME, August 2015, https://hal.archives-ouvertes.fr/hal-01142989

[97] R. JHA, D. CHABLAT, F. ROUILLIER, G. MOROZ.*Workspace and Singularity analysis of a Delta like family robot*, in "4th IFTOMM International Symposium on Robotics and Mechatronics", Poitiers, France, June 2015, https://hal.archives-ouvertes.fr/hal-01142465

[98] A. JOUX, R. LERCIER.*The function field sieve is quite special*, in "Algorithmic Number Theory-ANTS V", Lecture Notes in Computer Science, Springer, 2002, vol. 2369, p. 431-445

[99] A. JOUX, C. PIERROT.*Improving the Polynomial time Precomputation of Frobenius Representation Discrete Logarithm Algorithms - Simplified Setting for Small Characteristic Finite Fields*, in "20th International Conference on the Theory and Application of Cryptology and Information Security", Kaoshiung, Taiwan, Lecture Notes in Computer Science, Springer Berlin Heidelberg, December 2014, vol. 8873, p. 378-397 [*DOI :* 10.1007/978-3-662-45611-8_20], https://hal.archives-ouvertes.fr/hal-01213649

[100] A. JOUX, C. PIERROT.*Nearly Sparse Linear Algebra and application to Discrete Logarithms Computations*, in "Contemporary Developments in Finite Fields and Applications ", WorldScientific, 2016 [*DOI :* 10.1142/9789814719261_0008], https://hal.inria.fr/hal-01154879

[101] H. L. JR..*Factoring integers with elliptic curves*, in "Annals of Mathematics", 1987, vol. 126, n^o 2, p. 649–673

[102] T. KAILATH.*Linear Systems*, Prentice-Hall, 1980

[103] M. KASHIWARA.*Algebraic study of systems of partial differential equations*, Mémoires de la S. M. F., 1995, vol. 63, Master's thesis 1970 (English translation)

[104] M. KASHIWARA, T. KAWAI, T. KIMURA.*Foundations of Algebraic Analysis*, Princeton University Press, 1986, vol. 37

[105] A. KOBEL, F. ROUILLIER, M. SAGRALOFF.*Computing Real Roots of Real Polynomials ... and now For Real!*, in " ISSAC '16 Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation", Waterloo, Canada, ISSAC '16 Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation, July 2016, 7 [*DOI :* 10.1145/2930889.2930937], https://hal.inria.fr/hal-01363955

[106] N. KOBLITZ.*Elliptic curve cryptosystems*, in "Mathematics of Computation", January 1987, vol. 48, n^o 177, p. 203–209

[107] E. KOLCHIN.*Differential Algebra & Algebraic Groups*, Pure and Applied Mathematics, Elsevier Science, 1973

[108] P.-V. KOSELEFF, D. PECKER.*Chebyshev Knots*, in "Journal of Knot Theory and Its Ramifications", April 2011, vol. 20, n^o 4, p. 575-593 [*DOI :* 10.1142/S0218216511009364], https://hal.archives-ouvertes.fr/hal-00344501

[109] P.-V. KOSELEFF, D. PECKER.*On Alexander–Conway polynomials of two-bridge links*, in "Journal of Symbolic Computation", May 2015, vol. Volume 68, n^o 2, p. 215-229, 15p [*DOI :* 10.1016/J.JSC.2014.09.011], https://hal.archives-ouvertes.fr/hal-00538729

[110] P.-V. KOSELEFF, D. PECKER.*Harmonic Knots*, in "Journal Of Knot Theory And Its Ramifications (JKTR)", 2016, vol. 25, n^o 13, 18, 18 p., 30 fig. [*DOI :* 10.1142/S0218216516500747], https://hal.archives-ouvertes.fr/hal-00680746

[111] P.-V. KOSELEFF, D. PECKER, F. ROUILLIER.*The first rational Chebyshev knots*, in "Journal of Symbolic Computation", December 2010, vol. 45, n^o 12, p. 1341-1358 [*DOI :* 10.1016/J.JSC.2010.06.014], https://hal.archives-ouvertes.fr/hal-00429510

[112] P.-V. KOSELEFF, D. PECKER, F. ROUILLIER, C. TRAN.*Computing Chebyshev knot diagrams*, in "Journal of Symbolic Computation", 2018, vol. 86, 21 [*DOI :* 10.1016/J.JSC.2017.04.001], https://hal.inria.fr/hal-01232181

[113] P.-V. KOSELEFF, F. ROUILLIER, C. TRAN.*On the sign of a trigonometric expression*, in "ISSAC ' 15", Bath, United Kingdom, Proceedings of the 2015 ACM on International Symposium on Symbolic and Algebraic Computation, July 2015 [*DOI :* 10.1145/2755996.2756664], https://hal.inria.fr/hal-01200820

[114] B. A. LAMACCHIA, A. M. ODLYZKO.*Computation of discrete logarithms in prime fields*, in "Designs, Codes and Cryptography", 1991, vol. 1, p. 47–62

[115] S. LAZARD, M. POUGET, F. ROUILLIER.*Bivariate triangular decompositions in the presence of asymptotes*, in "Journal of Symbolic Computation", 2017, vol. 82, p. 123 – 133 [*DOI :* 10.1016/J.JSC.2017.01.004], https://hal.inria.fr/hal-01468796

[116] V. MILLER.*Use of elliptic curves in cryptography*, in "Advances in Cryptology — CRYPTO'85", H. WILLIAMS (editor), LNCS, Springer, 1986, vol. 218, p. 417–428

[117] B. MOURRAIN.*The 40 Generic Positions of a Parallel Robot*, in "Proceedings of the 1993 International Symposium on Symbolic and Algebraic Computation", New York, NY, USA, ISSAC '93, ACM, 1993, p. 173–182, http://doi.acm.org/10.1145/164081.164120

[118] D. NIANG DIATTA, F. ROUILLIER, M.-F. ROY.*On the computation of the topology of plane curves*, in "International Symposium on Symbolic and Algebraic Computation", Kobe, Japan, K. NABESHIMA (editor), ACM Press, July 2014, p. 130-137 [*DOI :* 10.1145/2608628.2608670], https://hal.archives-ouvertes.fr/hal-00935728

[119] U. OBERST.*Multidimensional constant linear systems*, in "Acta Appl. Math.", 1990, vol. 20, p. 1–175

[120] C. POMERANCE.*Analysis and comparison of some integer factoring methods*, in "Computational methods in number theory – Part I", Amsterdam, J. HENDRIK W. LENSTRA, R. TIJDEMAN (editors), Mathematical centre tracts, Mathematisch Centrum, 1982, vol. 154, p. 8–139

[121] POMMARET.*Systems of Partial Differential Equations and Lie Pseudogroups*, Ellis Horwood Series in Mathematics and its Applications, Gordon and Breach Science Publishers, 1978

[122] A. QUADRAT.*Grade filtration of linear functional systems*, in "Acta Applicandæ Mathematicæ", October 2013, vol. 127, n^o 1, p. 27–86 [*DOI :* 10.1007/S10440-012-9791-2], https://hal-supelec.archives-ouvertes.fr/hal-00925510

[123] A. QUADRAT.*Noncommutative geometric structures on stabilizable infinite-dimensional linear systems*, in "ECC 2014", Strasbourg, France, June 2014, p. 2460 – 2465 [*DOI :* 10.1109/ECC.2014.6862563], https://hal-supelec.archives-ouvertes.fr/hal-01108019

[124] A. QUADRAT.*A constructive algebraic analysis approach to Artstein's reduction of linear time-delay systems*, in "12th IFAC Workshop on Time Delay Systems", Ann Arbor, United States, Proceedings of 12th IFAC Workshop on Time Delay Systems, University of Michigan, May 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01259862

[125] A. QUADRAT.*Towards an effective study of the algebraic parameter estimation problem*, in " IFAC 2017 Workshop Congress", Toulouse, France, July 2017, https://hal.inria.fr/hal-01415300

[126] A. QUADRAT, G. REGENSBURGER.*Computing Polynomial Solutions and Annihilators of Integro-Differential Operators with Polynomial Coefficients*, Inria Lille - Nord Europe ; Institute for Algebra, Johannes Kepler University Linz, December 2016, n^o RR-9002, 24, https://hal.inria.fr/hal-01413907

[127] A. QUADRAT, D. ROBERTZ.*A constructive study of the module structure of rings of partial differential operators*, in "Acta Applicandæ Mathematicæ", 2014, vol. 133, p. 187–243 [*DOI :* 10.1007/S10440-013-9864-X], https://hal-supelec.archives-ouvertes.fr/hal-00925533

[128] A. QUADRAT, R. USHIROBIRA.*Algebraic analysis for the Ore extension ring of differential time-varying delay operators*, in " 22nd International Symposium on Mathematical Theory of Networks and Systems (MTNS)", Minneapolis, United States, July 2016, 8, https://hal.inria.fr/hal-01415256

[129] G. RANCE, Y. BOUZIDI, A. QUADRAT, A. QUADRAT.*A symbolic-numeric method for the parametric H∞ loop-shaping design problem*, in "22nd International Symposium on Mathematical Theory of Networks and Systems (MTNS) ", Minneapolis, United States, July 2016, 8, https://hal.inria.fr/hal-01415294

[130] G. RANCE, Y. BOUZIDI, A. QUADRAT, A. QUADRAT.*Explicit H∞ controllers for 1st to 3rd order single-input single-output systems with parameters*, in " IFAC 2017 Workshop Congress ", Toulouse, France, July 2017, https://hal.inria.fr/hal-01667410

[131] G. RANCE, Y. BOUZIDI, A. QUADRAT, A. QUADRAT, F. ROUILLIER.*Explicit H∞ controllers for 4th order single-input single-output systems with parameters and their applications to the two mass-spring system with damping*, in " IFAC 2017 Workshop Congress ", Toulouse, France, July 2017, https://hal.inria.fr/hal-01667368

[132] G. RANCE.*Parametric H_∞ control and its application to gyrostabilized sights*, Université Paris-Saclay, July 2018, https://tel.archives-ouvertes.fr/tel-01904086

[133] J. RITT.*Differential Algebra*, Colloquium publications, American Mathematical Society, 1950

[134] R. RIVEST, A. SHAMIR, L. ADLEMAN.*A method for obtaining digital signatures and public-key cryptosystems*, in "Commun. ACM", 1978, vol. 21, n^{o} 2, p. 120–126

[135] D. ROBERTZ.*Formal Algorithmic Elimination for PDEs*, Lecture Notes in Mathematics 2121, Springer, 2014

[136] J. ROTMAN.*An Introduction to Homological Algebra*, Universitext, Springer New York, 2008

[137] J. T. STAFFORD.*Module structure of Weyl algebras*, in "J. London Math. Soc.", 1978, vol. 18, p. 429–442

[138] V. A. VASSILIEV.*Cohomology of knot spaces*, in "Theory of singularities and its applications", Adv. Soviet Math., Amer. Math. Soc., Providence, RI, 1990, vol. 1, p. 23–69

[139] J. WEEKS.*Chapter 10 - Computation of Hyperbolic Structures in Knot Theory*, in "Handbook of Knot Theory", Amsterdam, W. MENASCO, M. THISTLETHWAITE (editors), Elsevier Science, 2005, p. 461 – 480 [*DOI :* 10.1016/B978-044451452-3/50011-3], http://www.sciencedirect.com/science/article/pii/B9780444514523500113

[140] P. WENGER.*A new general formalism for the kinematic analysis of all nonredundant manipulators*, in "ICRA", 1992

[141] J. WILLEMS, J. POLDERMAN.*Introduction to Mathematical Systems Theory: A Behavioral Approach*, Texts in Applied Mathematics, Springer New York, 2013

Project-Team PARKAS

Parallélisme de Kahn Synchrone

IN COLLABORATION WITH: Département d'Informatique de l'Ecole Normale Supérieure

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Paris

RESEARCH CENTER
Paris

THEME
Embedded and Real-time Systems

Table of contents

Project-Team PARKAS

Creation of the Team: 2011 April 01, updated into Project-Team: 2012 January 01

Keywords:

Computer Science and Digital Science:

A1.1.1. - Multicore, Manycore
A1.1.3. - Memory models
A2.1.1. - Semantics of programming languages
A2.1.4. - Functional programming
A2.1.6. - Concurrent programming
A2.1.9. - Synchronous languages
A2.2.2. - Memory models
A2.2.4. - Parallel architectures
A2.2.5. - Run-time systems
A2.2.6. - GPGPU, FPGA...
A2.2.7. - Adaptive compilation
A2.3. - Embedded and cyber-physical systems
A2.3.1. - Embedded systems
A2.3.2. - Cyber-physical systems
A2.3.3. - Real-time systems
A2.4.3. - Proofs
A3.1.3. - Distributed data
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.7. - High performance computing

Other Research Topics and Application Domains:

B5.2.1. - Road vehicles
B5.2.2. - Railway
B5.2.3. - Aviation
B6.4. - Internet of things
B6.6. - Embedded systems
B9.5.1. - Computer science
B9.5.2. - Mathematics

1. Team, Visitors, External Collaborators

Research Scientists

Timothy Bourke [Inria, Researcher]
Francesco Zappa Nardelli [Inria, Senior Researcher, HDR]

Faculty Member

Marc Pouzet [Team leader, Univ Pierre et Marie Curie, Professor]

External Collaborator

Paul Feautrier [Univ de Lyon]

Technical Staff

Andi Drebes [Inria, Engineer]
Sophie Kaleba [Inria, Engineer, from Mar 2019 until Aug 2019]
Chandan Reddy Gopal [Inria, Engineer, until Aug 2019]
Adilla Susungi [Inria, Engineer, until Mar 2019]
Nicolas Tollenaere [Inria, Engineer, until Mar 2019]

PhD Students

Ulysse Beaugnon [Ecole Normale Supérieure Paris, until Apr 2019]
Lelio Brun [Ecole Normale Supérieure Paris]
Basile Clément [Inria, until Nov 2019]
Ismail Lahkim Bennani [Inria]

Post-Doctoral Fellow

Guillaume Iooss [Inria]

2. Overall Objectives

2.1. Overall Objectives

Research in PARKAS focuses on the design, semantics, and compilation of programming languages which allow going from parallel deterministic specifications to target embedded code executing on sequential or multi-core architectures. We are driven by the ideal of a mathematical and executable language used both to program and simulate a wide variety of systems, including real-time embedded controllers in interaction with a physical environment (e.g., fly-by-wire, engine control), computationally intensive applications (e.g., video), and compilers that produce provably correct and efficient code.

The team bases its research on the foundational work of Gilles Kahn on the semantics of deterministic parallelism, the theory and practice of synchronous languages and typed functional languages, synchronous circuits, modern (polyhedral) compilation, and formal models to prove the correctness of low-level code running on weak-memory processors.

To realize our research program, we develop languages (LUCID SYNCHRONE, REACTIVEML, LUCY-N, ZELUS), compilers (PPCG), contributions to open-source projects (isl, LLVM, gcc, SundialsML), tools to study language semantics (Ott) and to test optimization compilers in the presence of threads (cmmtest), and formalizations in Interactive Theorem Provers of language semantics (Vélus, *n*-synchrony, quasi-synchrony). These software projects constitute essential “laboratories”: they ground our scientific contributions, guide and validate our research through experimentation, and are an important vehicle for long-standing collaborations with industry.

3. Research Program

3.1. Programming Languages for Cyber-Physical Systems

We study the definition of languages for reactive and Cyber-Physical Systems in which distributed control software interacts closely with physical devices. We focus on languages that mix discrete-time and continuous-time; in particular, the combination of synchronous programming constructs with differential equations, relaxed models of synchrony for distributed systems communicating via periodic sampling or through buffers, and the embedding of synchronous features in a general purpose ML language.

The synchronous language SCADE, [0] based on synchronous languages principles, is ideal for programming embedded software and is used routinely in the most critical applications. But embedded design also involves modeling the control software together with its environment made of physical devices that are traditionally defined by differential equations that evolve on a continuous-time basis and approximated with a numerical solver. Furthermore, compilation usually produces single-loop code, but implementations increasingly involve multiple and multi-core processors communicating via buffers and shared-memory.

The major player in embedded design for cyber-physical systems is undoubtedly SIMULINK, [0] with MODELICA[0] a new player. Models created in these tools are used not only for simulation, but also for test-case generation, formal verification, and translation to embedded code. That said, many foundational and practical aspects are not well-treated by existing theory (for instance, hybrid automata), and current tools. In particular, features that mix discrete and continuous time often suffer from inadequacies and bugs. This results in a broken development chain: for the most critical applications, the model of the controller must be reprogrammed into either sequential or synchronous code, and properties verified on the source model have to be reverified on the target code. There is also the question of how much confidence can be placed in the code used for simulation.

We attack these issues through the development of the ZELUS research prototype, industrial collaborations with the SCADE team at ANSYS/Esterel-Technologies, and collaboration with Modelica developers at Dassault-Systèmes and the Modelica association. Our approach is to develop a *conservative extension* of a synchronous language capable of expressing in a single source text a model of the control software and its physical environment, to simulate the whole using off-the-shelf numerical solvers, and to generate target embedded code. Our goal is to increase faithfulness and confidence in both what is actually executed on platforms and what is simulated. The goal of building a language on a strong mathematical basis for hybrid systems is shared with the Ptolemy project at UC Berkeley; our approach is distinguished by building our language on a synchronous semantics, reusing and extending classical synchronous compilation techniques.

Adding continuous time to a synchronous language gives a richer programming model where reactive controllers can be specified in idealized physical time. An example is the so called quasi-periodic architecture studied by Caspi, where independent processors execute periodically and communicate by sampling. We have applied ZELUS to model a class of quasi-periodic protocols and to analyze an abstraction proposed for model-checking such systems.

Communication-by-sampling is suitable for control applications where value timeliness is paramount and lost or duplicate values tolerable, but other applications—for instance, those involving video streams—seek a different trade-off through the use of bounded buffers between processes. We developed the n-synchronous model and the programming language LUCY-N to treat this issue.

3.2. Efficient Compilation for Parallel and Distributed Computing

We develop compilation techniques for sequential and multi-core processors, and efficient parallel run-time systems for computationally intensive real-time applications (e.g., video and streaming). We study the generation of parallel code from synchronous programs, compilation techniques based on the polyhedral model, and the exploitation of synchronous Single Static Assignment (SSA) representations in general purpose compilers.

We consider distribution and parallelism as two distinct concepts.

- Distribution refers to the construction of multiple programs which are dedicated to run on specific computing devices. When an application is designed for, or adapted to, an embedded multiprocessor, the distribution task grants fine grained—design- or compilation-time—control over the mapping and interaction between the multiple programs.

[0] http://www.esterel-technologies.com/products/scade-suite
[0] http://www.mathworks.com/products/simulink
[0] https://www.modelica.org

- Parallelism is about generating code capable of efficiently exploiting multiprocessors. Typically this amounts to maing (in)dependence properties, data transfers, atomicity and isolation explicit. Compiling parallelism translates these properties into low-level synchronization and communication primitives and/or onto a runtime system.

We also see a strong relation between the foundations of synchronous languages and the design of compiler intermediate representations for concurrent programs. These representations are essential to the construction of compilers enabling the optimization of parallel programs and the management of massively parallel resources. Polyhedral compilation is one of the most popular research avenues in this area. Indirectly, the design of intermediate representations also triggers exciting research on dedicated runtime systems supporting parallel constructs. We are particularly interested in the implementation of non-blocking dynamic schedulers interacting with decoupled, deterministic communication channels to hide communication latency and optimize local memory usage.

While distribution and parallelism issues arise in all areas of computing, our programming language perspective pushes us to consider four scenarios:

1. designing an embedded system, both hardware and software, and codesign;
2. programming existing embedded hardware with functional and behavioral constraints;
3. programming and compiling for a general-purpose or high-performance, best-effort system;
4. programming large scale distributed, I/O-dominated and data-centric systems.

We work on a multitude of research experiments, algorithms and prototypes related to one or more of these scenarios. Our main efforts focused on extending the code generation algorithms for synchronous languages and on the development of more scalable and widely applicable polyhedral compilation methods.

3.3. Validation and Proof of Compilers

Compilers are complex software and not immune from bugs. We work on validation and proof tools for compilers to relate the semantics of executed code and source programs. We develop techniques to formally prove the correctness of compilation passes for synchronous languages (Lustre), and to validate compilation optimization for C code in the presence of threads.

3.3.1. Lustre:

The formal validation of a compiler for a synchronous language (or more generally for a language based on synchronous block diagrams) promises to reduce the likelihood of compiler-introduced bugs, the cost of testing, and also to ensure that properties verified on the source model hold of the target code. Such a validation would be complementary to existing industrial qualifications which certify the development process and not the functional correctness of a compiler. The scientific interest is in developing models and techniques that both facilitate the verification and allow for convenient reasoning over the semantics of a language and the behavior of programs written in it.

3.3.2. C/C++:

The recently approved C11 and C++11 standards define a concurrency model for the C and C++ languages, which were originally designed without concurrency support. Their intent is to permit most compiler and hardware optimizations, while providing escape mechanisms for writing portable, high-performance, low-level code. Mainstream compilers are being modified to support the new standards. A subtle class of compiler bugs is the so-called concurrency compiler bugs, where compilers generate correct sequential code but break the concurrency memory model of the programming language. Such bugs are observable only when the miscompiled functions interact with concurrent contexts, making them particularly hard to detect. All previous techniques to test compiler correctness miss concurrency compiler bugs.

4. Application Domains

4.1. Embedded Control Software

Embedded control software defines the interactions of specialized hardware with the physical world. It normally ticks away unnoticed inside systems like medical devices, trains, aircraft, satellites, and factories. This software is complex and great effort is required to avoid potentially serious errors, especially over many years of maintenance and reuse.

Engineers have long designed such systems using block diagrams and state machines to represent the underlying mathematical models. One of the key insights behind synchronous programming languages is that these models can be executable and serve as the base for simulation, validation, and automatic code generation. This approach is sometimes termed Model-Based Development (MBD). The SCADE language and associated code generator allow the application of MBD in safety-critical applications. They incorporate ideas from LUSTRE, LUCID SYNCHRONE, and other programming languages.

4.2. Hybrid Systems Design and Simulation

Modern embedded systems are increasingly conceived as rich amalgams of software, hardware, networking, and physical processes. The terms Cyberphysical System (CPS) or Internet-of-Things (IoT) are sometimes used as labels for this point of view.

In terms of modeling languages, the main challenges are to specify both discrete and continuous processes in a single *hybrid* language, give meaning to their compositions, simulate their interactions, analyze the behavior of the overall system, and extract code either for target control software or more efficient, possibly online, simulation. Languages like Simulink and Modelica are already used in the design and analysis of embedded systems; it is more important than ever to understand their underlying principles and to propose new constructs and analyses.

5. New Software and Platforms

5.1. Cmmtest

FUNCTIONAL DESCRIPTION: Cmmtest is a tool for hunting concurrency compiler bugs. The Cmmtest tool performs random testing of C and C++ compilers against the C11/C++11 memory model. A test case is any well-defined, sequential C program, for each test case, cmmtest:

compiles the program using the compiler and compiler optimisations that are being tested,

runs the compiled program in an instrumented execution environment that logs all memory accesses to global variables and synchronisations,

compares the recorded trace with a reference trace for the same program, checking if the recorded trace can be obtained from the reference trace by valid eliminations, reorderings and introductions.

Cmmtest identified several mistaken write introductions and other unexpected behaviours in the latest release of the gcc compiler. These have been promptly fixed by the gcc developers.

- Participants: Anirudh Kumar, Francesco Zappa Nardelli, Pankaj More, Pankaj Pawan, Pankaj Prateek Kewalramani and Robin Morisset
- Contact: Francesco Zappa Nardelli
- URL: http://www.di.ens.fr/~zappa/projects/cmmtest/

5.2. GCC

KEYWORDS: Compilation - Polyhedral compilation

FUNCTIONAL DESCRIPTION: The GNU Compiler Collection includes front ends for C, C++, Objective-C, Fortran, Java, Ada, and Go, as well as libraries for these languages (libstdc++, libgcj,...). GCC was originally written as the compiler for the GNU operating system. The GNU system was developed to be 100% free software, free in the sense that it respects the user's freedom.

- Participants: Albert Cohen, Feng Li, Nhat Minh Le, Riyadh Baghdadi and Tobias Grosser
- Contact: Albert Cohen
- URL: http://gcc.gnu.org/

5.3. Heptagon

KEYWORDS: Compilers - Synchronous Language - Controller synthesis

FUNCTIONAL DESCRIPTION: Heptagon is an experimental language for the implementation of embedded real-time reactive systems. It is developed inside the Synchronics large-scale initiative, in collaboration with Inria Rhones-Alpes. It is essentially a subset of Lucid Synchrone, without type inference, type polymorphism and higher-order. It is thus a Lustre-like language extended with hierchical automata in a form very close to SCADE 6. The intention for making this new language and compiler is to develop new aggressive optimization techniques for sequential C code and compilation methods for generating parallel code for different platforms. This explains much of the simplifications we have made in order to ease the development of compilation techniques.

The current version of the compiler includes the following features: - Inclusion of discrete controller synthesis within the compilation: the language is equipped with a behavioral contract mechanisms, where assumptions can be described, as well as an "enforce" property part. The semantics of this latter is that the property should be enforced by controlling the behaviour of the node equipped with the contract. This property will be enforced by an automatically built controller, which will act on free controllable variables given by the programmer. This extension has been named BZR in previous works. - Expression and compilation of array values with modular memory optimization. The language allows the expression and operations on arrays (access, modification, iterators). With the use of location annotations, the programmer can avoid unnecessary array copies.

- Participants: Adrien Guatto, Brice Gelineau, Cédric Pasteur, Eric Rutten, Gwenaël Delaval, Léonard Gérard and Marc Pouzet
- Partners: UGA - ENS Paris - Inria - LIG
- Contact: Gwenaël Delaval
- URL: http://heptagon.gforge.inria.fr

5.4. isl

FUNCTIONAL DESCRIPTION: isl is a library for manipulating sets and relations of integer points bounded by linear constraints. Supported operations on sets include intersection, union, set difference, emptiness check, convex hull, (integer) affine hull, integer projection, transitive closure (and over-approximation), computing the lexicographic minimum using parametric integer programming. It includes an ILP solver based on generalized basis reduction, and a new polyhedral code generator. isl also supports affine transformations for polyhedral compilation, and increasingly abstract representations to model source and intermediate code in a polyhedral framework.

- Participants: Albert Cohen, Sven Verdoolaege and Tobias Grosser
- Contact: Sven Verdoolaege
- URL: http://freshmeat.net/projects/isl

5.5. Lem

lightweight executable mathematics

FUNCTIONAL DESCRIPTION: Lem is a lightweight tool for writing, managing, and publishing large scale semantic definitions. It is also intended as an intermediate language for generating definitions from domain-specific tools, and for porting definitions between interactive theorem proving systems (such as Coq, HOL4, and Isabelle). As such it is a complementary tool to Ott. Lem resembles a pure subset of Objective Caml, supporting typical functional programming constructs, including top-level parametric polymorphism, datatypes, records, higher-order functions, and pattern matching. It also supports common logical mechanisms including list and set comprehensions, universal and existential quantifiers, and inductively defined relations. From this, Lem generates OCaml, HOL4, Coq, and Isabelle code.

- Participants: Francesco Zappa Nardelli, Peter Sewell and Scott Owens
- Contact: Francesco Zappa Nardelli
- URL: http://www.cl.cam.ac.uk/~pes20/lem/

5.6. Lucid Synchrone

FUNCTIONAL DESCRIPTION: Lucid Synchrone is a language for the implementation of reactive systems. It is based on the synchronous model of time as provided by Lustre combined with features from ML languages. It provides powerful extensions such as type and clock inference, type-based causality and initialization analysis and allows to arbitrarily mix data-flow systems and hierarchical automata or flows and valued signals.

RELEASE FUNCTIONAL DESCRIPTION: The language is still used for teaching and in our research but we do not develop it anymore. Nonetheless, we have integrated several features from Lucid Synchrone in new research prototypes described below. The Heptagon language and compiler are a direct descendent of it. The new language Zélus for hybrid systems modeling borrows many features originaly introduced in Lucid Synchrone.

- Contact: Marc Pouzet
- URL: http://www.di.ens.fr/~pouzet/lucid-synchrone/

5.7. Lucy-n

Lucy-n: an n-synchronous data-flow programming language

FUNCTIONAL DESCRIPTION: Lucy-n is a language to program in the n-synchronous model. The language is similar to Lustre with a buffer construct. The Lucy-n compiler ensures that programs can be executed in bounded memory and automatically computes buffer sizes. Hence this language allows to program Kahn networks, the compiler being able to statically compute bounds for all FIFOs in the program.

- Participants: Adrien Guatto, Albert Cohen, Louis Mandel and Marc Pouzet
- Contact: Albert Cohen
- URL: https://www.lri.fr/~mandel/lucy-n/

5.8. Ott

FUNCTIONAL DESCRIPTION: Ott is a tool for writing definitions of programming languages and calculi. It takes as input a definition of a language syntax and semantics, in a concise and readable ASCII notation that is close to what one would write in informal mathematics. It generates output:

a LaTeX source file that defines commands to build a typeset version of the definition,

a Coq version of the definition,

an Isabelle version of the definition, and

a HOL version of the definition.

Additionally, it can be run as a filter, taking a LaTeX/Coq/Isabelle/HOL source file with embedded (symbolic) terms of the defined language, parsing them and replacing them by typeset terms.

The main goal of the Ott tool is to support work on large programming language definitions, where the scale makes it hard to keep a definition internally consistent, and to keep a tight correspondence between a definition and implementations. We also wish to ease rapid prototyping work with smaller calculi, and to make it easier to exchange definitions and definition fragments between groups. The theorem-prover backends should enable a smooth transition between use of informal and formal mathematics.

- Participants: Francesco Zappa Nardelli, Peter Sewell and Scott Owens
- Contact: Francesco Zappa Nardelli
- URL: http://www.cl.cam.ac.uk/~pes20/ott/

5.9. PPCG

FUNCTIONAL DESCRIPTION: PPCG is our source-to-source research tool for automatic parallelization in the polyhedral model. It serves as a test bed for many compilation algorithms and heuristics published by our group, and is currently the best automatic parallelizer for CUDA and OpenCL (on the Polybench suite).

- Participants: Albert Cohen, Riyadh Baghdadi, Sven Verdoolaege and Tobias Grosser
- Contact: Sven Verdoolaege
- URL: http://freshmeat.net/projects/ppcg

5.10. ReactiveML

FUNCTIONAL DESCRIPTION: ReactiveML is a programming language dedicated to the implementation of interactive systems as found in graphical user interfaces, video games or simulation problems. ReactiveML is based on the synchronous reactive model due to Boussinot, embedded in an ML language (OCaml).

The Synchronous reactive model provides synchronous parallel composition and dynamic features like the dynamic creation of processes. In ReactiveML, the reactive model is integrated at the language level (not as a library) which leads to a safer and a more natural programming paradigm.

- Participants: Cédric Pasteur, Guillaume Baudart and Louis Mandel
- Contact: Guillaume Baudart

5.11. SundialsML

Sundials/ML

KEYWORDS: Simulation - Mathematics - Numerical simulations

SCIENTIFIC DESCRIPTION: Sundials/ML is a comprehensive OCaml interface to the Sundials suite of numerical solvers (CVODE, CVODES, IDA, IDAS, KINSOL). Its structure mostly follows that of the Sundials library, both for ease of reading the existing documentation and for adapting existing source code, but several changes have been made for programming convenience and to increase safety, namely:

solver sessions are mostly configured via algebraic data types rather than multiple function calls,

errors are signalled by exceptions not return codes (also from user-supplied callback routines),

user data is shared between callback routines via closures (partial applications of functions),

vectors are checked for compatibility (using a combination of static and dynamic checks), and

explicit free commands are not necessary since OCaml is a garbage-collected language.

FUNCTIONAL DESCRIPTION: Sundials/ML is a comprehensive OCaml interface to the Sundials suite of numerical solvers (CVODE, CVODES, IDA, IDAS, KINSOL, ARKODE).

RELEASE FUNCTIONAL DESCRIPTION: Adds support for v3.1.x of the Sundials Suite of numerical solvers.

Notably this release adds support for the new generic matrix and linear solver interfaces. The OCaml interface changes but the library is backward compatible with Sundials 2.7.0.

OCaml 4.02.3 or greater is now required and optionally OCamlMPI 1.03.

* New Sundials.Matrix and Sundials.LinearSolver modules. * Better treatment of integer type used for matrix indexing. * Refactor Dls and Sls modules into Sundials.Matrix. * Add confidence intervals to performance graph. * Miscellaneous improvements to configure script. * Potential incompatibility: changes to some label names: comm_fn -> comm, iter_type -> iter. * Untangle the ARKODE mass-solver interface from the Jacobian interface.

- Participants: Jun Inoue, Marc Pouzet and Timothy Bourke
- Partner: UPMC
- Contact: Marc Pouzet
- URL: http://inria-parkas.github.io/sundialsml/

5.12. Zelus

SCIENTIFIC DESCRIPTION: The Zélus implementation has two main parts: a compiler that transforms Zélus programs into OCaml programs and a runtime library that orchestrates compiled programs and numeric solvers. The runtime can use the Sundials numeric solver, or custom implementations of well-known algorithms for numerically approximating continuous dynamics.

FUNCTIONAL DESCRIPTION: Zélus is a new programming language for hybrid system modeling. It is based on a synchronous language but extends it with Ordinary Differential Equations (ODEs) to model continuous-time behaviors. It allows for combining arbitrarily data-flow equations, hierarchical automata and ODEs. The language keeps all the fundamental features of synchronous languages: the compiler statically ensure the absence of deadlocks and critical races, it is able to generate statically scheduled code running in bounded time and space and a type-system is used to distinguish discrete and logical-time signals from continuous-time ones. The ability to combines those features with ODEs made the language usable both for programming discrete controllers and their physical environment.

- Participants: Marc Pouzet and Timothy Bourke
- Contact: Marc Pouzet

5.13. Telamon

KEYWORDS: Compilation - Monte-Carlo methods - Constraint Programming - GPU - Dense linear algebra

FUNCTIONAL DESCRIPTION: Telamon is a framework for the optimization of computational kernels for GPUs through efficient search in a well-behaved optimization space in which optimization decisions commute and satisfaction of constraints restricting legal optimizations.

- Contact: Ulysse Beaugnon
- URL: https://github.com/ulysseB/telamon/

5.14. Vélus

Verified Lustre Compiler

KEYWORDS: Synchronous Language - Compilation - Software Verification - Coq - Ocaml

FUNCTIONAL DESCRIPTION: Vélus is a prototype compiler from a subset of Lustre to assembly code. It is written in a mix of Coq and OCaml and incorporates the CompCert verified C compiler. The compiler includes formal specifications of the semantics and type systems of Lustre, as well as the semantics of intermediate languages, and a proof of correctness that relates the high-level dataflow model to the values produced by iterating the generated assembly code.

RELEASE FUNCTIONAL DESCRIPTION: First source-code release. Treatment of primitive reset construct. Clocks allowed for node arguments.

- Contact: Timothy Bourke
- URL: https://velus.inria.fr

5.15. Tensor Comprehensions

KEYWORDS: Machine learning - Matrix calculation - Polyhedral compilation - GPU - CUDA

FUNCTIONAL DESCRIPTION: Tensor Comprehensions (TC) is a notation based on generalized Einstein notation for computing on multi-dimensional arrays. TC greatly simplifies ML framework implementations by providing a concise and powerful syntax which can be efficiently translated to high-performance computation kernels, automatically.

RELEASE FUNCTIONAL DESCRIPTION: Integration of the loop tactics matching framework for identifying linear algebra operations and optimizing them.

- Partner: Eindhoven University of Technology
- Contact: Albert Cohen

5.16. Aftermath

KEYWORDS: Performance analysis - Runtime system - Parallel programming - High-performance calculation - Execution trace

FUNCTIONAL DESCRIPTION: Aftermath is a toolkit for building custom graphical tools for trace-based performance analysis of parallel programs, run-time systems and compilers.

- Partner: The University of Manchester
- Contact: Andi Drebes

5.17. MPPcodegen

Source-to-source loop tiling based on MPP

KEYWORDS: Source-to-source compiler - Polyhedral compilation

FUNCTIONAL DESCRIPTION: MPPcodegen applies a monoparametric tiling to a C program enriched with pragmas specifying the tiling and the scheduling function. The tiling can be generated by any convex polyhedron and translation functions, it is not necessarily a partition. The result is a C program depending on a scaling factor (the parameter). MPPcodegen relies on the MPP mathematical library to tile the iteration sets.

- Partner: Colorado State University
- Contact: Christophe Alias
- URL: http://foobar.ens-lyon.fr/mppcodegen/

5.18. MPP

MonoParametric Partitionning transformation

KEYWORDS: Compilation - Polyhedral compilation

FUNCTIONAL DESCRIPTION: This library applies a monoparametric partitioning transformation to polyhedra and affine functions. This transformation is a subset of the parametric sized tiling transformation, specialized for the case where shapes depend only on a single parameter. Unlike in the general case, the resulting sets and functions remain in the polyhedral model.

- Contact: Guillaume Iooss
- URL: https://github.com/guillaumeiooss/MPP

5.19. Obelisk

KEYWORDS: LaTeX - HTML - Ocaml

FUNCTIONAL DESCRIPTION: Obelisk is a simple tool which produces pretty-printed output from a Menhir parser file (.mly).

It is inspired from yacc2latex and it is also written in OCaml, but it is aimed at supporting features from Menhir instead of only those of ocamlyacc.

- Contact: Lelio Brun
- URL: https://github.com/Lelio-Brun/Obelisk

6. New Results

6.1. Efficiently Subtyping Union Types

Participant: Francesco Zappa Nardelli.

Julia is a programming language recently designed at MIT to support the needs of the scientific community. Julia occupies a unique position in the design landscape, it is a dynamic language with no type system, yet it has a surprisingly rich set of types and type annotations used to specify multimethod dispatch. The types that can be expressed in function signatures include parametric union types, covariant tuple types, parametric user-defined types with single inheritance, invariant type application, and finally types and values can be reified to appear in signatures. In 2017 with Vitek we started a research project to study the design and the pragmatic use of the Julia language, and formalised the Julia subtyping algorithm. In 2018 we have pursued this study, and we have proved correct the clever and space efficient algorithm relied upon by the Julia runtime. This has been published in [17].

6.2. Fast and reliable unwinding via DWARF tables

Participants: Theophile Bastian, Rémy Oudin, Francesco Zappa Nardelli.

DWARF is a widely-used debugging data format. DWARF is obviously relied upon by debuggers, but it plays an unexpected role in the runtime of high-level programming languages and in the implementation of program analysis tools. The debug information itself can be pervaded by subtle bugs, making the whole infrastructure unreliable. In this project we are investigating techniques and tools to perform validation and synthesis of the DWARF stack unwinding tables, to speedup DWARF-based unwinding, as well as exploring adventurous projects that can be built on top of reliable DWARF information.

We have built a tool that can validate DWARF unwind tables generated by mainstream compilers; the approach is effective, we found a problem in Clang table generation and several in GLIBC inline-assembly snippets. We also designed and implemented a tool that can synthesise DWARF unwind tables from binary that lacks them (e.g. because the compiler did not generate them - immediate applications: JITs assembly, inline assembly, ...). Additionally we have designed and implemented a ahead-of-time compiler of DWARF unwind tables to assembly, and an ad-hoc unwinder integrated with the defacto standard unwinder libuwind. It can speed up unwinding by a factor between 12x and 25x (depending on application), with a 2.5x size overhead for unwind information.

This work has been published in [13].

6.3. Verified compilation of Lustre

Participants: Timothy Bourke, Lélio Brun, Paul Jeanmaire, Marc Pouzet.

Vélus is a compiler for a subset of LUSTRE and SCADE that is specified in the Coq [28] Interactive Theorem Prover (ITP). It integrates the CompCert C compiler [34], [29] to define the semantics of machine operations (integer addition, floating-point multiplication, etcetera) and to generate assembly code for different architectures. The research challenges are to

- to mechanize, i.e., put into Coq, the semantics of the programming constructs used in modern languages for MBD;
- to implement compilation passes and prove them correct;
- to interactively verify source programs and guarantee that the obtained invariants also hold of the generated code.

This year we created a website for the project (https://velus.inria.fr) and made an initial release under an Inria non-commerical license (https://github.com/Inria/velus). T. Bourke's JCJC ("Jeune Chercheuse Jeune Chercheur") project *FidelR* was accepted for funding by the ANR: it aims to develop ITP-based techniques for treating state machines and interactive program verification. We also made progress on the compilation of the modular reset construct, the treatment of (non-normalized) Lustre, and our longer term goal of strengthening the main correctness theorem. These results are detailed below.

6.3.1. Compiling the modular reset construct.

In the original LUSTRE language, the only way to reset the internal state of an instantiated function is to propagate and test explicit reset signals. Later languages, like LUCID SYNCHRONE and SCADE, provide a construct for resetting an instance modularly (it works for any function) and efficiently (testing occurs only at the point of instantiation). Last year we showed how to encode the semantics of this construct in Coq. This year we focused on its compilation and the associated proof of correctness. We designed and implemented a new intermediate language that exposes different *step* and *reset* actions on node instances. This language facilitates the optimization of conditional statements in the generated code and permits the transformation to imperative code and its proof of correctness to be treated in two steps: one to introduce named memories and another to fix the sequential order of execution. This work forms the core of L. Brun's thesis, to be defended early next year, and an article accepted at the ACM SIGBED international conference on Principles of Programming Languages (POPL 2020).

6.3.2. Non-normalized Lustre.

Our previous work has focused on a subset of "normalized" programs where the form of expressions and equations is constrained to facilitate the compilation. We have generalized the definitions of syntax and semantics in our prototype compiler to accept non-normalized programs. This included simplifying and generalizing the formalization of clocks presented in [20]. With P. Jeanmaire (M2 internship), we have implemented a compilation pass to translate normalized programs from one syntactical form to another. The main challenge was to formally prove an alignment property (signals are present iff their clocks are true) that had been assumed until now. The proof is finished except for the inductive case for the reset construct which we hope to complete soon.

6.3.3. Strengthening the correctness theorem.

The current correctness theorem assumes that an accepted program can be given a semantics in terms of the mechanized model. It should be possible to prove this fact for programs that pass the initial type-checking and clock-checking algorithms, that can be scheduled, and which never invoke an undefined operation (such as a division by zero). We made good progress on this problem by defining an interpreter for normalized Lustre programs and showing that the results it calculates satisfy the semantic predicates. This initial work gives some useful insights into how to proceed. We presented it at the Synchron 2019 workshop.

6.4. Specifying multi-clock Lustre programs

Participants: Timothy Bourke, Guillaume Iooss, Baptiste Pauget, Marc Pouzet.

It is sometimes desirable to compile a single synchronous language program into multiple tasks for execution by a real-time operating system. We have been investigating this question from three different perspectives.

6.4.1. Harmonic clocks

We studied the extension of a synchronous language with periodic harmonic clocks based on the work of Mandel et al. [31], [37], [32], [35], [36] on n-synchrony and the extension proposed by Forget et al. [33]

Mandel et al. considered a language with periodic clocks expressed as ultimately periodic binary sequences. The decision procedures (equality, inclusion, precedence) for such an expressive language can be very costly. It is thus sometimes useful to apply an envelope-based abstraction, that is, one where sets of clocks are represented by a rational slope and an interval. Forget considered simpler "harmonic" clocks. His decision procedures conincide with those for the envelope-based abstraction but without any loss of information. During his M2 ineternship, B. Pauget continued this line of work by extending the input language of the Vélus Lustre compiler with harmonic clocks. This work was the starting point for the proposal of a new intermediate language for a synchronous compiler that is capable of exploiting clock information to apply agressive optimizations and generate parallel code.

6.4.2. New Intermediate Language MObc (Multi Object Code)

This intermediate language is reminiscent of the intermediate Obc language used in the Vélus and Heptagon compiler, but with some important differences and new features. MObc permits a synchronous function to be represented as a set of named state variables and possibly nested blocks with a partial ordering which express the way blocks can and must be called. In comparison, Obc represents a synchronous function as a set of state variables and a transition function that is itself written in a sequential language. Each block comprises a set of equations in Single Static Assignment (SSA) form, that is, exactly one equation per variable, so as to simplify the implementation of a number of classic optimizations (for example, constant propagation, inlining, common sub-expression elimination, code specialisation). Then, every block is translated into a step function (e.g., a C function). This intermediate language has been designed to facilitate the generation of code for a real-time OS and a multi-core target. This work exploits two older results: the article of Caspi et al. [30] that introduces an object representation for synchronous nodes and a "scheduling policy" that specifies how their methods may be called, and; the work of Pouzet et al. [38] on the calculation of input/output relations to merge calculations. We are preparing and article on this subject.

6.4.3. Clocking constraints, communication latencies, and constraint solving

In this approach, the top-level node of a Lustre program is distinguished from inner nodes. It may contain special annotations to specify the triggering and other details of node instances from which separate "tasks" are to be generated. Special operators are introduced to describe the buffering between top-level instances. Notably, different forms of the `fby` and `current` operators are provided. Some of the operators are under-specified and a constraint solver is used to determine their exact meaning, that is, whether the signal is delayed by zero, one, or more cycles of the receiving clock, which depends on the scheduling of the source and destination nodes. Scheduling is formalized as a constraint solving problem based on latency constraints between some pairs of input/outputs that are specified by the designer. G. Iooss has been prototyping these ideas in the academic Heptagon compiler.

This work is funded by a direct industrial contract with Airbus. In collaboration (this year) with Michel Angot Vincent Bregeon Jean Souyris (Airbus, R&D) and Matthieu Boitrel (Airbus BE).

6.5. The Zelus Language

Participants: Timothy Bourke, Ismail Lakhim-Bennani, Marc Pouzet.

Zelus is our laboratory to experiment our research on programming languages for hybrid systems. It is devoted to the design and implementation of systems that may mix discrete-time/continuous-time signals and systems between those signals. It is essentially a synchronous language reminiscent of Lustre and Lucid Synchrone but with the ability to define functions that manipulate continuous-time signals defined by Ordinary Differential Equations (ODEs). The language is functional in the sense that a system is a function from signals to signals (not a relation). It provides some features from ML languages like higher-order and parametric polymorphism as well as dedicated static analyses.

This year, we have pursued our work on the design, semantics and implementation of hybrid modeling language, in particular the treatment of Differential Algebraic Equations (DAEs) [23].

6.5.1. Compiler Internals: Static Typing and Compiler Organisation

The distribution with manual and examples is distributed at http://zelus.di.ens.fr (only Version 1, in binary form). Version 2 (the current active branch) is available in source form on Inria GitLab https://gitlab.inria.fr/parkas/zelus, on simple demand.

Several new experimentations have been done this year, in particular on the type system and an extensive rewritting of some compilation internals to simplify the code and make the generated code more shorter (in size) and more efficient.

6.5.2. Co-simulation as Function Lifting

Hybrid models in Zelus (that is, programs that mix discrete and continuous-time signals) are simulated using a single ODE and zero-crossing solvers only. All hybrid modeling languages (e.g., Simulink, Modelica, Ptolemy) act the same way, at least, single solver simulation is the default mechanism.

Its weaknesses are well known: any change of the dynamic, even local, calls for a global reset of the solver, making it slower for that later steps; the mix of a slow and fast signals slows down the whole simulation. Co-simulation is about running several solvers (or instances of the same) at the same time.

We proposed a limited (but useful) manner, by proving a way to internalize the solver to obtain, from a continuous-time function, a synchronous stream function. A preliminary experiment done this year was surprisingly and pleasingly simple to implement in Zelus. It consisted in defining a (higher-order) function *solve* that, given a continuous-time function *f* returns a stream function *solve f*. Given an input stream *x* and an increasing stream of time horizons *h*, *solve f(x, t)* returns the stream of approximated values. This function internalizes the ODE solver and the zero-crossing detection mechanism. The overall model is then a purely discrete-time, synchronous model. In particular, classical synchronization protocols between solvers can be programmed in the language itself, hence benefiting from the static checks that track typing, causality and initialization errors, properties that would be more difficult to ensure if programmed directly in C, for example. We think that it is even possible to write a formal synchronous specification of the simulation engine itself, that is, to program the function solve directly in Zelus. This experiment on co-simulation gives new insight on the semantics based on non standard analysis that we proposed and, more interestingly, to relate it to the proven and more classical semantics based on super-dense time studied and exploited by Edward Lee.

6.5.3. QSS-based Simulation

Quantized State Systems simulation (QSS) was introduced in the early 2000's by F. Cellier and E. Kofman as an alternative to time-based simulation, which is the dominant approach to ODE/DAE systems simulation.

Rather than linking QSS to Discrete Event Simulation, we have made a preliminary experiment to relate it to Synchronous Programming and its continuous time extension Zelus. Zelus is used to give a formal description of the QSS method that can be executed. We have described the very basic scheme called QSS (or QSS1) for which we can give a Zelus (hence executable) specification. Higher order schemes QSS2, 3, etc. can also be given an Zelus specification. Implicit schemes were also proposed by Kofman for a better handling of stiff systems. Higher order versions of BQSS are nontrivial; they are called LIQSS1, 2, 3, etc. and they can also be specified in Zelus. This preliminary work is done in collaboration with Albert Benveniste (Inria Hycomes, Rennes) and funded by the Modeliscale FUI project.

6.5.4. Property Based Testing of Hybrid Programs

Property-based program testing involves checking an executable specification by running many tests. We build on the work of Georgios Fainekos and Alexandre Donzé, and take inspiration from earlier work by Nicolas Halbwachs, to write a Zélus library of synchronous observers with a quantitative semantics that can be used to specify properties of a system under test. We implemented several optimization algorithms for producing test cases, some of which are gradient-based. To compute the gradients, we use Automatic Differentiation (AD) of the system under test and its specification. Together with François Bidet, we ported the well-known FADBAD++ library for AD written by Ole Stauning in 1997 to OCaml—the target language of Zélus. Our port is called FADBADml and is now released under an Inria license [0] and is available on opam.

6.6. Reactive Probabilistic Programming

Participant: Marc Pouzet.

Synchronous languages were introduced to design and implement real-time embedded systems with a (justified) enphasis on determinacy. Yet, they interact with a physical environment that is partially known and are implemented on architectures subject to failures and noise (e.g., channels, variable communication delays or computation time). Dealing with uncertainties is useful online monitoring, learning, statistical testing or to build simplified models for faster simulation. Actual synchronous and languages provide limited support for modeling the non-deterministic behaviors that are omnipresent in embedded systems.

In 2019, we started a new topic on *reactive probabilistic programming* under the initiative of Guillaume Baudart and Louis Mandel (IBM Research, Watson); in collaboration with Erik Atkinson, Michael Carbin and Benjamin Sherman (MIT). We have designed ProbZelus, an extension of Zelus with probabilistic programming constructs. The language makes it possible to describe probabilistic models in interaction with an observable environment. At runtime, a set of inference techniques can be used to learn the distributions of model parameters from observed data. The main results are (1) the design of the ProbZelus compiler, the formalization of the static and dynamic semantics of the language that mixes deterministic and probabilistic components, (2) the design and implementation of inference methods which can be executed with bounded resources, (3) the evaluation of ProbZelus on a set of examples and case studies.

For the moment, ProbZelus is mainly a library of Zelus, with minor changes of the language itself (essentially the type system) [0] It exploits heavilly the higher-order nature of Zelus. E.g., if `f` is a stream function (with type `f: 'a -D-> 'b`) and `x` is a stream (with type `'a`), `Particule.infer f x: 'a -D-> 'b Distribution.t` implements an inference algorithm which computes a distribution for the result of type `'b` with a particule filter algorithm.

A prelimary report describes a part of this work [24] and a presentation at JFLA will be given in January 2020. ProbZelus is available in open source at https://github.com/IBM/probzelus since december 2019. Our purpose is to go beyond the library approach with a closer integration of probabilistic constructs and reactive constructs, with dedicated static analyses and compilation techniques to give static guaranties on the result of inference, efficient inference techniques tuned for reactive applications and that ensure execution in bounded time and space; efficient dedicated compilation techniques for probabilistic programs. Finally, the treatment of both discrete-time and continuous-time signals and systems must be investigated (only discrete-time is considered at the moment).

6.7. Identification of matrix operations for Compute-In-Memory architectures from a high-level Machine Learning framework

Participant: Andi Drebes.

Compute-In-Memory (CIM) architectures are capable of performing certain performance-critical operations directly in memory (e.g., matrix multiplications) and represent a promising approach to partially eliminate the bottleneck of traditional von Neumann-based architectures resulting from long-distance communication between main memory and processing units.

In order for applications to benefit from such architectures, their operations must be divided into highly parallel, uniform operations eligible for in-memory computation and control logic that cannot benefit from CIM and that must be carried out by conventional computing devices. It is crucial for this process that as many eligible operations as possible are identified and effectively processed in memory, resulting only in as few computations as possible carried out on the conventional cores.

[0] https://fadbadml-dev.github.io/FADBADml/

[0] Yet, higher-order was only used occasionally since then; hence an important implementation effort has been spent this year to make it work well.

The programmability of CIM architectures is a key factor for its overall success. Manual identification of eligible operations and mapping to hardware resources is tedious, error-prone and requires detailed knowledge of the target architecture and therefore does not represent a viable approach to program CIM architectures.

With our partners from the MNEMOSENE project, we have developed a compilation toolchain that unburdens programmers from technical details of CIM architectures by allowing them to express algorithms at a high level of abstraction and that automates parallelization, orchestration and the mapping of operations to the CIM architecture. The solution integrates the Loop Tactics [40] declarative polyhedral pattern recognition and transformation framework into Tensor Comprehensions [39], a framework generating highly optimized kernels for accelerators from an abstract, mathematical notation for tensor operations. The compilation flow performs a set of dedicated optimizations aiming at enabling the reliable detection of computational patterns and their efficient mapping to CIM accelerators.

The results of this work have been submitted to the 10th International Workshop on Polyhedral Compilation Techniques (IMPACT).

6.8. Applying reinforcement learning to improve a branch-and-bound optimizing compiler

Participant: Basile Clement.

Frameworks for image processing, deep learning, etc., work with Directed Acyclic Graphs (DAGs) of computational operators that wrap high-performance libraries. The production of highly optimized, target-specific implementations of the library functions come at a high engineering cost: languages and compilers have failed to deliver performances competitive with expert written code, notably on GPUs and other hardware accelerators. Moreover, library implementations may not offer optimal performance for a specific use case. They may lack inter-operator optimizations and specializations to specific data sizes and shapes.

In his thesis, Ulysse Beaugnon, a former PhD student in the team, proposed to formulate this compilation problem as an optimization research problem using a combination of analytical modeling, experimental search, constraint programming and branch-and-bound optimization techniques. Basile Clement started a PhD to extend this idea, exploring the improvements required to make it fully competitive with handwritten code. In 2019, he evaluated reinforcement learning techniques such as multi-armed bandit schemes to improve the performance and efficiency of the search procedure; extended the analytical model with generic sizes, making it more precise before selecting tiling parameters; and made various improvements to the code generation procedure.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. Collaboration with Airbus

Our work on multi-clock Lustre programs is funded by a contract with Airbus.

7.2. Bilateral Grants with Industry

7.2.1. Google Research Fellowship: DWARF unwinding

Francesco Zappa Nardelli benefits from a Google Research Fellowship to pursue the work on DWARF unwinding, about 50k euros.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

The ANR JCJC project “FidelR” was awarded to Timothy Bourke this year and will begin in 2020.

8.1.1.1. ANR/CHIST-ERA DIVIDEND project, 2013-2019.

This project continues.

8.1.2. FUI: Fonds unique interministériel

8.1.2.1. Modeliscale contract (AAP-24)

Using Modelica at scale to model and simulate very large Cyber-Physical Systems. Principal industrial partner: Dassault-Systèmes. Inria contacts are Benoit Caillaud (HYCOMES, Rennes) and Marc Pouzet (PARKAS, Paris).

8.1.3. Programme d’Investissements d’Avenir (PIA)

8.1.3.1. ES3CAP collaborative project (Bpifrance)

Develop a software and hardware platform for tomorrow’s intelligent systems. PARKAS collaborates with the industrial participants ANSYS/Esterel Technologies, Kalray, and Safran Electronics & Defense. Inria contacts are Marc Pouzet (PARKAS, Paris) and Fabrice Rastello (CORSE, Grenoble).

8.1.4. Others

8.1.4.1. Inria Project Lab (IPL) Modeliscale

This project treats the modelling and analysis of Cyber-Physical Systems at large scale. The PARKAS team contributes their expertise in programming language design for reactive and hybrid systems to this multi-team effort.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

- MNEMOSENE is a project with funding from the European Union’s Horizon 2020 Research and Innovation Programme. Its objectives include the improvement of the energy-delay product, the computational efficiency and performance density by several orders of magnitude compared to state-of-the-art architectures. A cornerstone of the proposed solution is the memristor-based Compute-in-Memory (CIM) architecture, which eliminates long-distance, high-latency data transfers between memory and computing units required in conventional Von Neumann-based architectures by carrying out computations for performance-critical operations directly in memory.
- TETRAMAX, *Technology Transfer via Multinational Application Experiments*, is funded by the H2020 “Smart Anything Everywhere (SAE)” initiative. The overall ambition is to build and leverage a European Competence Center Network in customized low-energy computing, providing easy access for SMEs and mid-caps to novel CLEC technologies via local contact points. This is a bidirectional interaction: SMEs can demand CLEC technologies and solutions via the network, and vice versa academic research institutions can actively and effectively offer their new technologies to European industries. Furthermore, TETRAMAX wants to support 50+ industry clients and 3rd parties with innovative technologies, using different kinds of Technology Transfer Experiments (TTX) to accelerate innovation within European industries and to create a competitive advantage in the global economy.

8.3. International Initiatives

8.3.1. Participation in Other International Programs

- VerticA (Francesco Zappa Nardelli), 2017-2020, joint project with Northeastern University, USA, financed by the ONR (Office of Naval Research), $1.5M (subcontract for $150k).

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Selection

9.1.1.1. Chair of Conference Program Committees

- T. Bourke co-chaired the program committee of the ACM/IEEE international conference on Embedded Software (EMSOFT 2019).

9.1.1.2. Member of the Conference Program Committees

- T. Bourke served on the PC of the Euromicro Conference on Real-Time Systems (ECRTS 2019).
- T. Bourke served on the PC of the international workshop on Software and Compilers for Embedded Systems (SCOPES 2019).
- T. Bourke served on the PC of the international Modelica conference (MODELICA 2019).
- T. Bourke served on the PC of the OCaml Users and Developers Meeting (OCaml 2019).
- M. Pouzet served on the PC of the program committee of the ACM/IEEE international conference on Embedded Software (EMSOFT 2019).
- M. Pouzet served on the PC of the international workshop on Software and Compilers for Embedded Systems (SCOPES 2019).
- M. Pouzet served on the PC of the international workshop on Cyber-Physical Systems (CyPhy), a satellite event of ESWEEK.
- F. Zappa Nardelli served on the PC of the international conference on Object-Oriented Programming, Systems, Languages & Applications (OOPSLA 2019).
- F. Zappa Nardelli served on the PC of the ACM Workshop on Gradual Types (WGC 2020).

9.1.1.3. Reviewer

- T. Bourke reviewed submissions for the international conference on Computer Aided Verification (CAV 2019)
- T. Bourke reviewed submisssions for the international conference on Object-Oriented Programming, Systems, Languages & Applications (OOPSLA 2019).

9.1.2. Journal

9.1.2.1. Reviewer - Reviewing Activities

- T. Bourke reviewed articles for the Journal of Logical and Algebraic Methods in Programming (JLAMP).

9.1.3. Research Administration

- F. Zappa Nardelli chaired the part-time assistant professor recruitment committee at École Polytechnique.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Marc Pouzet is Director of Studies for the CS department, at ENS.

Licence : M. Pouzet & T. Bourke: “Operating Systems” (L3), Lectures and TDs, ENS, France.

Licence : T. Bourke, “Digital Systems” (L3), Lectures and TDs, ENS, France

Master: M. Pouzet & T. Bourke: “Synchronous Systems” (M2), Lectures and TDs, MPRI, France

Master: M. Pouzet: “Synchronous Reactive Languages” (M2), Lectures, Master COMASIC (École Polytechnique) and FIL (Université Paris-Sud, Saclay), France

Master: T. Bourke, lab classes for “A Programmer’s introduction to Computer Architectures and Operating Systems" (M1), École Polytechnique, France

Master: F. Zappa Nardelli: “A Programmer’s introduction to Computer Architectures and Operating Systems" (M1), 45h, École Polytechnique, France

Bachelor: F. Zappa Nardelli: “A Programmer’s introduction to Computer Architectures and Operating Systems", 20h, École Polytechnique, France

Internships T. Bourke participated in reviewing the L3 and M1 internships of students at the ENS, France.

M1 Projects: T. Bourke is directing two M1 research projects.

9.2.2. Supervision

- PhD in progress: Lélio Brun, under review, supervised by T. Bourke and M. Pouzet.
- PhD in progress: Ismail Lakhim-Bennani, 1st year, supervised by M. Pouzet, G. Frehse, and T. Bourke
- PhD in progress: Basile Clément, 1st year, supervised by F. Zappa Nardelli and M. Pouzet.
- PhD: Chandan Reddy, supervised by A. Cohen, defended in March 2019.
- PhD: Ulysse Beaugnon, supervised by A. Cohen and M. Pouzet, defended in June 2019.

9.2.3. Juries

- T. Bourke was a jury member for the for the PhD thesis of Steven Varoumas, November 2019.
- F. Zappa Nardelli was a jury member for the PhD thesis of Francois Ginraud, Grenoble, Jan 2019.

9.3. Popularization

9.3.1. Internal or external Inria responsibilities

- T. Bourke participated in the *commission d’emplois scientifiques* (postdocs, delegations, and PhDs) for the Inria Paris research centre.
- T. Bourke participated in the PhD progress committees (*suivi doctorale*) for B. Medeiros de Barros and N. Kulatova.
- T. Bourke and M. Pouzet participated in meetings with the *Ministère de la Transition écologique et solidaire* on the validation of autonomous driving systems.

10. Bibliography

Major publications by the team in recent years

[1] T. BOURKE, L. BRUN, P.-E. DAGAND, X. LEROY, M. POUZET, L. RIEG.*A Formally Verified Compiler for Lustre*, in "PLDI 2017 - 38th ACM SIGPLAN Conference on Programming Language Design and Implementation", Barcelone, Spain, ACM, June 2017, https://hal.inria.fr/hal-01512286

[2] T. BOURKE, F. CARCENAC, J.-L. COLAÇO, B. PAGANO, C. PASTEUR, M. POUZET.*A Synchronous Look at the Simulink Standard Library*, in "EMSOFT 2017 - 17th International Conference on Embedded Software", Seoul, South Korea, ACM Press, October 2017, 23, https://hal.inria.fr/hal-01575631

[3] T. BOURKE, J.-L. COLAÇO, B. PAGANO, C. PASTEUR, M. POUZET.*A Synchronous-based Code Generator For Explicit Hybrid Systems Languages*, in "International Conference on Compiler Construction (CC)", London, United Kingdom, LNCS, July 2015, https://hal.inria.fr/hal-01242732

[4] L. GÉRARD, A. GUATTO, C. PASTEUR, M. POUZET.*A modular memory optimization for synchronous data-flow languages: application to arrays in a lustre compiler*, in "Proceedings of the 13th ACM SIGPLAN/SIGBED International Conference on Languages, Compilers, Tools and Theory for Embedded Systems", Beijing, China, ACM, June 2012, p. 51–60 [*DOI :* 10.1145/2248418.2248426], https://hal.inria.fr/hal-00728527

[5] J. C. JUEGA, S. VERDOOLAEGE, A. COHEN, J. I. GÓMEZ, C. TENLLADO, F. CATTHOOR.*Patterns for parallel programming on GPUs*, in "Patterns for parallel programming on GPUs", F. MAGOULÈS (editor), Saxe-Cobourg, 2013, vol. Evaluation of State-of-the-Art Parallelizing Compilers Generating CUDA Code for Heterogeneous CPU/GPU Computing, ISBN 978-1-874672-57-9, https://hal.archives-ouvertes.fr/hal-01257261

[6] L. MANDEL, F. PLATEAU, M. POUZET.*Static Scheduling of Latency Insensitive Designs with Lucy-n*, in "FMCAD 2011 - Formal Methods in Computer Aided Design", Austin, TX, United States, October 2011, https://hal.inria.fr/hal-00654843

[7] R. MORISSET, P. PAWAN, F. ZAPPA NARDELLI.*Compiler testing via a theory of sound optimisations in the C11/C++11 memory model*, in "PLDI 2013 - 34th ACM SIGPLAN conference on Programming language design and implementation", Seattle, WA, United States, ACM, June 2013, p. 187-196 [*DOI :* 10.1145/2491956.2491967], https://hal.inria.fr/hal-00909083

[8] A. POP, A. COHEN.*OpenStream: Expressiveness and Data-Flow Compilation of OpenMP Streaming Programs*, in "ACM Transactions on Architecture and Code Optimization", 2013, vol. 9, n^{o} 4, Selected for presentation at the HiPEAC 2013 Conf [*DOI :* 10.1145/2400682.2400712], https://hal.inria.fr/hal-00786675

[9] J. SEVCIK, V. VAFEIADIS, F. ZAPPA NARDELLI, S. JAGANNATHAN, P. SEWELL.*CompCertTSO: A Verified Compiler for Relaxed-Memory Concurrency*, in "Journal of the ACM (JACM)", 2013, vol. 60, n^{o} 3, p. art. 22:1-50 [*DOI :* 10.1145/2487241.2487248], https://hal.inria.fr/hal-00909076

[10] V. VAFEIADIS, T. BALABONSKI, S. CHAKRABORTY, R. MORISSET, F. ZAPPA NARDELLI.*Common compiler optimisations are invalid in the C11 memory model and what we can do about it*, in "POPL 2015 - 42nd ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages", Mumbai, India, January 2015, https://hal.inria.fr/hal-01089047

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] U. BEAUGNON.*Efficient Code Generation for Hardware Accelerators by Refining Partially Specified Implementations*, Ecole Normale Superieure de Paris - ENS Paris, June 2019, https://hal.archives-ouvertes.fr/tel-02385303

[12] C. REDDY.*Polyhedral Compilation for Domain Specific Languages*, Ecole normale supérieure, March 2019, https://hal.inria.fr/tel-02385670

Articles in International Peer-Reviewed Journal

[13] T. BASTIAN, S. KELL, F. ZAPPA NARDELLI.*Reliable and Fast DWARF-Based Stack Unwinding*, in "Proceedings of the ACM on Programming Languages", 2019, vol. OOPSLA [*DOI :* 10.1145/3360572], https://hal.inria.fr/hal-02297690

[14] T. BOURKE, L. BRUN, M. POUZET.*Mechanized semantics and verified compilation for a dataflow synchronous language with reset*, in "Proceedings of the ACM on Programming Languages", January 2020, vol. 4, n^o^ POPL, p. 1-29 [*DOI :* 10.1145/3371112], https://hal.inria.fr/hal-02426573

[15] K. DIDIER, D. POTOP-BUTUCARU, G. IOOSS, A. COHEN, J. SOUYRIS, P. BAUFRETON, A. GRAILLAT.*Correct-by-Construction Parallelization of Hard Real-Time Avionics Applications on Off-the-Shelf Predictable Hardware*, in "ACM Transactions on Architecture and Code Optimization", August 2019, vol. 16, n^o^ 3, p. 1-27 [*DOI :* 10.1145/3328799], https://hal.inria.fr/hal-02422789

Invited Conferences

[16] G. BAUDART, T. BOURKE, M. POUZET.*Symbolic Simulation of Dataflow Synchronous Programs with Timers*, in "12th Forum on Specification and Design Languages (FDL 2017)", Vérone, Italy, D. GROSSE, S. VINCO, H. PATEL (editors), Languages, Design Methods, and Tools for Electronic System Design: Selected Contributions from FDL 2017, Springer, January 2019, vol. 530, 25 [*DOI :* 10.1007/978-3-030-02215-0_3], https://hal.inria.fr/hal-01575621

International Conferences with Proceedings

[17] B. CHUNG, F. ZAPPA NARDELLI, J. VITEK.*Julia's efficient algorithm for subtyping unions and covariant tuples (Pearl)*, in "ECOOP 2019 - 33rd European Conference of Object-Oriented Programming", London, United Kingdom, July 2019 [*DOI :* 10.4230/LIPIcs.ECOOP.2019.6], https://hal.inria.fr/hal-02297696

[18] K. DIDIER, A. COHEN, D. POTOP-BUTUCARU, A. GAUFFRIAU.*Sheep in wolf's Clothing: Implementation Models for Dataflow Multi-Threaded Software*, in "ACSD 2019 - 19th International Conference on Application of Concurrency to System Design", Aachen, Germany, IEEE, June 2019, p. 43-52 [*DOI :* 10.1109/ACSD.2019.00009], https://hal.inria.fr/hal-02422787

[19] J. PROY, K. HEYDEMANN, A. BERZATI, F. MAJERIC, A. COHEN.*A First ISA-Level Characterization of EM Pulse Effects on Superscalar Microarchitectures*, in "ARES 2019 - 14th International Conference on Availability, Reliability and Security", Canterbury, United Kingdom, ACM Press, 2019, p. 7:1–7:10 [*DOI :* 10.1145/3339252.3339253], https://hal.sorbonne-universite.fr/hal-02373088

National Conferences with Proceeding

[20] T. BOURKE, M. POUZET.*Clocked arguments in a verified Lustre compiler*, in "JFLA 2019 - Les Trentièmes Journées Francophones des Langages Applicatifs", Les Rousses, France, Les actes des trentièmes Journées Francophones des Langages Applicatifs (JFLA 2019), January 2019, 16, https://hal.inria.fr/hal-02005639

Conferences without Proceedings

[21] G. BAUDART, L. MANDEL, M. POUZET, E. ATKINSON, B. SHERMAN, M. CARBIN.*Programmation d'Applications Réactives Probabilistes*, in "JLFA 2020 - Journées Francophones des Langages Applicatifs", Gruissan, France, January 2020, https://hal.inria.fr/hal-02430070

[22] N. M. NOBRE, A. DREBES, G. RILEY, A. POP.*Bounded Stream Scheduling in Polyhedral OpenStream*, in "IMPACT 2020 - 10th International Workshop on Polyhedral Compilation Techniques", Bologna, Italy, January 2020, https://hal.inria.fr/hal-02441182

Scientific Books (or Scientific Book chapters)

[23] A. BENVENISTE, B. CAILLAUD, H. ELMQVIST, K. GHORBAL, M. OTTER, M. POUZET.*Multi-Mode DAE Models - Challenges, Theory and Implementation*, in "Computing and Software Science: State of the Art and Perspectives", Lecture Notes in Computer Science, October 2019, vol. 10000, p. 283-310 [*DOI :* 10.1007/978-3-319-91908-9_16], https://hal.inria.fr/hal-02333603

Other Publications

[24] G. BAUDART, L. MANDEL, E. ATKINSON, B. SHERMAN, M. POUZET, M. CARBIN.*Reactive Probabilistic Programming*, August 2019, https://arxiv.org/abs/1908.07563 - working paper or preprint, https://hal.inria.fr/hal-02426533

[25] A. DREBES, L. CHELINI, O. ZINENKO, A. COHEN, H. CORPORAAL, T. GROSSER, K. VADIVEL, N. VASILACHE.*TC-CIM: Empowering Tensor Comprehensions for Computing-In-Memory*, January 2020, IMPACT 2020 - 10th International Workshop on Polyhedral Compilation Techniques, https://hal.inria.fr/hal-02441163

[26] N. M. NOBRE, A. DREBES, G. RILEY, A. POP.*Beyond Polyhedral Analysis of OpenStream Programs*, January 2019, 9th International Workshop on Polyhedral Compilation Techniques, https://hal.inria.fr/hal-02370558

[27] J. PROY, K. HEYDEMANN, F. MAJERIC, A. COHEN, A. BERZATI.*Studying EM Pulse Effects on Superscalar Microarchitectures at ISA Level*, April 2019, https://arxiv.org/abs/1903.02623 - working paper or preprint, https://hal.sorbonne-universite.fr/hal-02102373

References in notes

[28] *The Coq proof Assistant*, 2019, http://coq.inria.fr

[29] S. BLAZY, Z. DARGAYE, X. LEROY.*Formal Verification of a C Compiler Front-End*, in "FM 2006: Int. Symp. on Formal Methods", Lecture Notes in Computer Science, Springer-Verlag, 2006, vol. 4085, p. 460–475, http://gallium.inria.fr/~xleroy/publi/cfront.pdf

[30] P. CASPI, J.-L. COLAÇO, L. GÉRARD, M. POUZET, P. RAYMOND.*Synchronous Objects with Scheduling Policies: Introducing safe shared memory in Lustre*, in "ACM International Conference on Languages, Compilers, and Tools for Embedded Systems (LCTES)", Dublin, June 2009

[31] A. COHEN, M. DURANTON, C. EISENBEIS, C. PAGETTI, F. PLATEAU, M. POUZET.*N-Synchronous Kahn Networks: a Relaxed Model of Synchrony for Real-Time Systems*, in "ACM International Conference on Principles of Programming Languages (POPL'06)", Charleston, South Carolina, USA, January 2006

[32] A. COHEN, L. MANDEL, F. PLATEAU, M. POUZET.*Abstraction of Clocks in Synchronous Data-flow Systems*, in "The Sixth ASIAN Symposium on Programming Languages and Systems (APLAS)", Bangalore, India, December 2008

[33] J. FORGET.*Un Langage Synchrone pour les Systèmes Embarqués Critiques Soumis à des Contraintes Temps Réel Multiples*, Université de Toulouse, November 2009

[34] X. LEROY.*The Compcert verified compiler*, 2009, http://compcert.inria.fr/doc/index.html

[35] L. MANDEL, F. PLATEAU, M. POUZET.*Lucy-n: a n-Synchronous Extension of Lustre*, in "Tenth International Conference on Mathematics of Program Construction (MPC 2010)", Québec, Canada, June 2010, http://www.lri.fr/~mandel/papiers/MandelPlateauPouzet-MPC-10.pdf

[36] L. MANDEL, F. PLATEAU, M. POUZET.*Static Scheduling of Latency Insensitive Designs with Lucy-n*, in "International Conference on Formal Methods in Computer-Aided Design (FMCAD)", Austin, Texas, USA, October 30 – November 2 2011

[37] F. PLATEAU.*Modèle n-synchrone pour la programmation de réseaux de Kahn à mémoire bornée*, Université Paris-Sud 11, Orsay, France, 6 janvier 2010, https://www.lri.fr/~mandel/lucy-n/~plateau/these/

[38] M. POUZET, P. RAYMOND.*Modular Static Scheduling of Synchronous Data-flow Networks: An efficient symbolic representation*, in "ACM International Conference on Embedded Software (EMSOFT'09)", Grenoble, France, October 2009

[39] N. VASILACHE, O. ZINENKO, T. THEODORIDIS, P. GOYAL, Z. DEVITO, W. S. MOSES, S. VERDOOLAEGE, A. ADAMS, A. COHEN.*The Next 700 Accelerated Layers: From Mathematical Expressions of Network Computation Graphs to Accelerated GPU Kernels, Automatically*, in "ACM Transactions on Architecture and Code Optimization (TACO)", October 2019, vol. 16, n^o 4, Article 38 [*DOI :* 10.1145/3355606]

[40] O. ZINENKO, L. CHELINI, T. GROSSER.*Declarative Transformations in the Polyhedral Model*, Inria ; ENS Paris - Ecole Normale Supérieure de Paris ; ETH Zurich ; TU Delft ; IBM Zürich, December 2018, n^o RR-9243, https://hal.inria.fr/hal-01965599

Project-Team PI.R2

Design, study and implementation of languages for proofs and programs

IN COLLABORATION WITH: Institut de Recherche en Informatique Fondamentale

IN PARTNERSHIP WITH:
CNRS

Université Denis Diderot (Paris 7)

RESEARCH CENTER
Paris

THEME
Proofs and Verification

Table of contents

Project-Team PI.R2

Creation of the Team: 2009 January 01, updated into Project-Team: 2011 January 01

Keywords:

Computer Science and Digital Science:

A2.1.1. - Semantics of programming languages
A2.1.4. - Functional programming
A2.1.11. - Proof languages
A2.4.3. - Proofs
A7.2. - Logic in Computer Science
A7.2.3. - Interactive Theorem Proving
A7.2.4. - Mechanized Formalization of Mathematics
A8.1. - Discrete mathematics, combinatorics
A8.4. - Computer Algebra

Other Research Topics and Application Domains:

B6.1. - Software industry

1. Team, Visitors, External Collaborators

Research Scientists

Pierre-Louis Curien [Team leader until October 2019, CNRS, Senior Researcher, Emeritus since October 2019, HDR]
Thierry Coquand [University of Gothenburg, Senior Researcher, until November 2020]
Emilio Jesús Gallego Arias [Inria, Starting Research Position, from November 2019]
Yves Guiraud [Inria, Researcher, HDR]
Hugo Herbelin [Inria, Senior Researcher, HDR]
Jean-Jacques Lévy [Inria, Emeritus, HDR]
Alexis Saurin [Team leader since October 2019, CNRS, Researcher]
Matthieu Sozeau [Inria, Researcher]

Faculty Members

Pierre Letouzey [Université Paris Diderot, Associate Professor]
Yann Régis-Gianas [Université Paris Diderot, Associate Professor, HDR]

External Collaborator

Jovana Obradović [Institute of Mathematics, Csech Academy of Science]

Technical Staff

Daniel de Rauglaudre [Inria, Engineer]

PhD Students

Antoine Allioux [Université Paris Diderot, PhD Student, from January 2018]
Félix Castro [Université Paris Diderot, PhD Student, from October 2019]
Kostia Chardonnet [Université Paris-Saclay, PhD Student, from November 2019]
Abhishek De [Université Paris Diderot, PhD Student]
Alen Durić [Université Paris Diderot, PhD Student, from October 2019]
Colin Gonzalez [Nomadics Lab, PhD Student, from October 2019, granted by CIFRE]
Cédric Ho Thanh [Université Paris Diderot, PhD Student]
Farzad Jafar-Rahmani [Université Paris Diderot, PhD Student, from October 2019]

Hugo Moeneclaey [Université Paris-Saclay, PhD Student, from September 2019]
Théo Zimmermann [Université Paris Diderot, PhD Student, until December 2019]

Post-Doctoral Fellows

Amar Hadzihasanovic [Sorbonne Université, Post-Doctoral Fellow, from December 2019]
Kailiang Ji [Inria, Post-Doctoral Fellow, until March 2019]
Exequiel Rivas Gadda [Inria, Post-Doctoral Fellow, until February 2019]

2. Overall Objectives

2.1. Overall Objectives

The research conducted in πr^2 is devoted both to the study of foundational aspects of formal proofs and programs and to the development of the Coq proof assistant software, with a focus on the dependently typed programming language aspects of Coq. The team acts as one of the strongest teams involved in the development of Coq as it hosts in particular the current coordinator of the Coq development team.

Since 2012, the team has also extended its scope to the study of the homotopy of rewriting systems, which shares foundational tools with recent advanced works on the semantics of type theories.

3. Research Program

3.1. Proof theory and the Curry-Howard correspondence

3.1.1. Proofs as programs

Proof theory is the branch of logic devoted to the study of the structure of proofs. An essential contributor to this field is Gentsen [77] who developed in 1935 two logical formalisms that are now central to the study of proofs. These are the so-called "natural deduction", a syntax that is particularly well-suited to simulate the intuitive notion of reasoning, and the so-called "sequent calculus", a syntax with deep geometric properties that is particularly well-suited for proof automation.

Proof theory gained a remarkable importance in computer science when it became clear, after genuine observations first by Curry in 1958 [72], then by Howard and de Bruijn at the end of the 60's [89], [109], that proofs had the very same structure as programs: for instance, natural deduction proofs can be identified as typed programs of the ideal programming language known as λ-calculus.

This proofs-as-programs correspondence has been the starting point to a large spectrum of researches and results contributing to deeply connect logic and computer science. In particular, it is from this line of work that Coquand and Huet's Calculus of Constructions [69], [70] stemmed out – a formalism that is both a logic and a programming language and that is at the source of the Coq system [107].

3.1.2. Towards the calculus of constructions

The λ-calculus, defined by Church [67], is a remarkably succinct model of computation that is defined via only three constructions (abstraction of a program with respect to one of its parameters, reference to such a parameter, application of a program to an argument) and one reduction rule (substitution of the formal parameter of a program by its effective argument). The λ-calculus, which is Turing-complete, i.e. which has the same expressiveness as a Turing machine (there is for instance an encoding of numbers as functions in λ-calculus), comes with two possible semantics referred to as call-by-name and call-by-value evaluations. Of these two semantics, the first one, which is the simplest to characterise, has been deeply studied in the last decades [60].

To explain the Curry-Howard correspondence, it is important to distinguish between intuitionistic and classical logic: following Brouwer at the beginning of the 20^{th} century, classical logic is a logic that accepts the use of reasoning by contradiction while intuitionistic logic proscribes it. Then, Howard's observation is that the proofs of the intuitionistic natural deduction formalism exactly coincide with programs in the (simply typed) λ-calculus.

A major achievement has been accomplished by Martin-Löf who designed in 1971 a formalism, referred to as modern type theory, that was both a logical system and a (typed) programming language [98].

In 1985, Coquand and Huet [69], [70] in the Formel team of Inria-Rocquencourt explored an alternative approach based on Girard-Reynolds' system F [78], [102]. This formalism, called the Calculus of Constructions, served as logical foundation of the first implementation of Coq in 1984. Coq was called CoC at this time.

3.1.3. The Calculus of Inductive Constructions

The first public release of CoC dates back to 1989. The same project-team developed the programming language Caml (nowadays called OCaml and coordinated by the Gallium team) that provided the expressive and powerful concept of algebraic data types (a paragon of it being the type of lists). In CoC, it was possible to simulate algebraic data types, but only through a not-so-natural not-so-convenient encoding.

In 1989, Coquand and Paulin [71] designed an extension of the Calculus of Constructions with a generalisation of algebraic types called inductive types, leading to the Calculus of Inductive Constructions (CIC) that started to serve as a new foundation for the Coq system. This new system, which got its current definitive name Coq, was released in 1991.

In practice, the Calculus of Inductive Constructions derives its strength from being both a logic powerful enough to formalise all common mathematics (as set theory is) and an expressive richly-typed functional programming language (like ML but with a richer type system, no effects and no non-terminating functions).

3.2. The development of Coq

During 1984-2012 period, about 40 persons have contributed to the development of Coq, out of which 7 persons have contributed to bring the system to the place it was six years ago. First Thierry Coquand through his foundational theoretical ideas, then Gérard Huet who developed the first prototypes with Thierry Coquand and who headed the Coq group until 1998, then Christine Paulin who was the main actor of the system based on the CIC and who headed the development group from 1998 to 2006. On the programming side, important steps were made by Chet Murthy who raised Coq from the prototypical state to a reasonably scalable system, Jean-Christophe Filliâtre who turned to concrete the concept of a small trustful certification kernel on which an arbitrary large system can be set up, Bruno Barras and Hugo Herbelin who, among other extensions, reorganised Coq on a new smoother and more uniform basis able to support a new round of extensions for the next decade.

The development started from the Formel team at Rocquencourt but, after Christine Paulin got a position in Lyon, it spread to École Normale Supérieure de Lyon. Then, the task force there globally moved to the University of Orsay when Christine Paulin got a new position there. On the Rocquencourt side, the part of Formel involved in ML moved to the Cristal team (now Gallium) and Formel got renamed into Coq. Gérard Huet left the team and Christine Paulin started to head a Coq team bilocalised at Rocquencourt and Orsay. Gilles Dowek became the head of the team which was renamed into LogiCal. Following Gilles Dowek who got a position at École Polytechnique, LogiCal moved to the new Inria Saclay research center. It then split again, giving birth to ProVal. At the same time, the Marelle team (formerly Lemme, formerly Croap) which has been a long partner of the Formel team, invested more and more energy in the formalisation of mathematics in Coq, while contributing importantly to the development of Coq, in particular for what regards user interfaces.

After various other spreadings resulting from where the wind pushed former PhD students, the development of Coq got multi-site with the development now realised mainly by employees of Inria, the CNAM, and Paris Diderot.

In the last seven years, Hugo Herbelin and Matthieu Sozeau coordinated the development of the system, the official coordinator hat passed from Hugo to Matthieu in August 2016. The ecosystem and development model changed greatly during this period, with a move towards an entirely distributed development model, integrating contributions from all over the world. While the system had always been open-source, its development team was relatively small, well-knit and gathered regularly at Coq working groups, and many developments on Coq were still discussed only by the few interested experts.

The last years saw a big increase in opening the development to external scrutiny and contributions. This was supported by the "core" team which started moving development to the open GitHub platform (including since 2017 its bug-tracker [43] and wiki), made its development process public, starting to use public pull requests to track the work of developers, organising yearly hackatons/coding-sprints for the dissemination of expertise and developers & users meetings like the Coq Workshop and CoqPL, and, perhaps more anecdotally, retransmitting Coq working groups on a public YouTube channel.

This move was also supported by the hiring of Maxime Dénès in 2016 as an Inria research engineer (in Sophia-Antipolis), and the work of Matej Košík (2-year research engineer). Their work involved making the development process more predictable and streamlined and to provide a higher level of quality to the whole system. In se 2018, a second engineer, Vincent Laporte, was hired. Yves Bertot, Maxime Dénès and Vincent Laporte are developing the Coq consortium, which aims to become the incarnation of the global Coq community and to offer support for our users.

Today, the development of Coq involves participants from the Inria project-teams pi.r2 (Paris), Marelle (Sophia-Antipolis), Toccata (Saclay), Gallinette (Nantes), Gallium (Paris), and Camus (Strasboug), the LIX at École Polytechnique and the CRI Mines-ParisTech. Apart from those, active collaborators include members from MPI-Saarbrucken (D. Dreyer's group), KU Leuven (B. Jacobs group), MIT CSAIL (A. Chlipala's group, which hosted an Inria/MIT engineer, and N. Zeldovich's group), the Institute for Advanced Study in Princeton (from S. Awodey, T. Coquand and V. Voevodsky's Univalent Foundations program) and Intel (M. Soegtrop). The latest released version Coq 8.8.0 had 40 contributors (counted from the start of 8.8 development) and the upcoming Coq 8.9 has 54.

On top of the developer community, there is a much wider user community, as Coq is being used in many different fields. The Software Foundations series, authored by academics from the USA, along with the reference Coq'Art book by Bertot and Castéran [61], the more advanced Certified Programming with Dependent Types book by Chlipala [66] and the recent book on the Mathematical Components library by Mahboubi, Tassi et al. provide resources for gradually learning the tool.

In the programming languages community, Coq is being taught in two summer schools, OPLSS and the DeepSpec summer school. For more mathematically inclined users, there are regular Winter Schools in Nice and in 2017 there was a school on the use of the Univalent Foundations library in Birmingham.

Since 2016, Coq also provides a central repository for Coq packages, the Coq opam archive, relying on the OCaml opam package manager and including around 250 packages contributed by users. It would be too long to make a detailed list of the uses of Coq in the wild. We only highlight four research projects relying heavily on Coq. The Mathematical Components library has its origins in the formal proof of the Four Colour Theorem and has grown to cover many areas of mathematics in Coq using the now integrated (since Coq 8.7) SSREFLECT proof language. The DeepSpec project is an NSF Expedition project led by A. Appel whose aim is full-stack verification of a software system, from machine-checked proofs of circuits to an operating system to a web-browser, entirely written in Coq and integrating many large projects into one. The ERC CoqHoTT project led by N. Tabareau aims to use logical tools to extend the expressive power of Coq, dealing with the univalence axiom and effects. The ERC RustBelt project led by D. Dreyer concerns the development of rigorous formal foundations for the Rust programming language, using the Iris Higher-Order Concurrent Separation Logic Framework in Coq.

We next briefly describe the main components of Coq.

3.2.1. The underlying logic and the verification kernel

The architecture adopts the so-called de Bruijn principle: the well-delimited *kernel* of Coq ensures the correctness of the proofs validated by the system. The kernel is rather stable with modifications tied to the evolution of the underlying Calculus of Inductive Constructions formalism. The kernel includes an interpreter of the programs expressible in the CIC and this interpreter exists in two flavours: a customisable lazy evaluation machine written in OCaml and a call-by-value bytecode interpreter written in C dedicated to efficient computations. The kernel also provides a module system.

3.2.2. Programming and specification languages

The concrete user language of Coq, called *Gallina*, is a high-level language built on top of the CIC. It includes a type inference algorithm, definitions by complex pattern-matching, implicit arguments, mathematical notations and various other high-level language features. This high-level language serves both for the development of programs and for the formalisation of mathematical theories. Coq also provides a large set of commands. Gallina and the commands together forms the *Vernacular* language of Coq.

3.2.3. Standard library

The standard library is written in the vernacular language of Coq. There are libraries for various arithmetical structures and various implementations of numbers (Peano numbers, implementation of $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ with binary digits, implementation of $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ using machine words, axiomatisation of $\mathbb{R}$). There are libraries for lists, list of a specified length, sorts, and for various implementations of finite maps and finite sets. There are libraries on relations, sets, orders.

3.2.4. Tactics

The tactics are the methods available to conduct proofs. This includes the basic inference rules of the CIC, various advanced higher level inference rules and all the automation tactics. Regarding automation, there are tactics for solving systems of equations, for simplifying ring or field expressions, for arbitrary proof search, for semi-decidability of first-order logic and so on. There is also a powerful and popular untyped scripting language for combining tactics into more complex tactics.

Note that all tactics of Coq produce proof certificates that are checked by the kernel of Coq. As a consequence, possible bugs in proof methods do not hinder the confidence in the correctness of the Coq checker. Note also that the CIC being a programming language, tactics can have their core written (and certified) in the own language of Coq if needed.

3.2.5. Extraction

Extraction is a component of Coq that maps programs (or even computational proofs) of the CIC to functional programs (in OCaml, Scheme or Haskell). Especially, a program certified by Coq can further be extracted to a program of a full-fledged programming language then benefiting of the efficient compilation, linking tools, profiling tools, ... of the target language.

3.2.6. Documentation

Coq is a feature-rich system and requires extensive training in order to be used proficiently; current documentation includes the reference manual, the reference for the standard library, as well as tutorials, and related tooling [sphinx plugins, coqdoc]. The jsCoq tool allows writing interactive web pages were Coq programs can be embedded and executed.

3.2.7. Proof development infrastructure

Coq is used in large-scale proof developments, and provides users miscellaneous tooling to help with them: the coq_makefile and Dune build systems help with incremental proof-checking; the Coq OPAM repository contains a package index for most Coq developments; the CoqIDE, ProofGeneral, and VSCoq user interfaces are environments for proof writing; and the Coq's API does allow users to extend the system in many important ways. Among the current extensions we have QuickChik, a tool for property-based testing; STMCoq and CoqHammer integrating Coq with automated solvers; ParamCoq, providing automatic derivation of parametricity principles; MetaCoq for metaprogramming; Equations for dependently-typed programming; SerAPI, for data-centric applications; etc... This also includes the main open Coq repository living at Github.

3.3. Dependently typed programming languages

Dependently typed programming (shortly DTP) is an emerging concept referring to the diffuse and broadening tendency to develop programming languages with type systems able to express program properties finer than the usual information of simply belonging to specific data-types. The type systems of dependently-typed programming languages allow to express properties *dependent* of the input and the output of the program (for instance that a sorting program returns a list of same size as its argument). Typical examples of such languages were the Cayenne language, developed in the late 90's at Chalmers University in Sweden and the DML language developed at Boston. Since then, various new tools have been proposed, either as typed programming languages whose types embed equalities (Ωmega at Portland, ATS at Boston, ...) or as hybrid logic/programming frameworks (Agda at Chalmers University, Twelf at Carnegie, Delphin at Yale, OpTT at U. Iowa, Epigram at Nottingham, ...).

DTP contributes to a general movement leading to the fusion between logic and programming. Coq, whose language is both a logic and a programming language which moreover can be extracted to pure ML code plays a role in this movement and some frameworks combining logic and programming have been proposed on top of Coq (Concoqtion at Rice and Colorado, Ynot at Harvard, Why in the ProVal team at Inria, Iris at MPI-Saarbrucken). It also connects to Hoare logic, providing frameworks where pre- and post-conditions of programs are tied with the programs.

DTP approached from the programming language side generally benefits of a full-fledged language (e.g. supporting effects) with efficient compilation. DTP approached from the logic side generally benefits of an expressive specification logic and of proof methods so as to certify the specifications. The weakness of the approach from logic however is generally the weak support for effects or partial functions.

3.3.1. Type-checking and proof automation

In between the decidable type systems of conventional data-types based programming languages and the full expressiveness of logically undecidable formulae, an active field of research explores a spectrum of decidable or semi-decidable type systems for possible use in dependently typed programming languages. At the beginning of the spectrum, this includes, for instance, the system F's extension ML_F of the ML type system or the generalisation of abstract data types with type constraints (G.A.D.T.) such as found in the Haskell programming language. At the other side of the spectrum, one finds arbitrary complex type specification languages (e.g. that a sorting function returns a list of type "sorted list") for which more or less powerful proof automation tools exist – generally first-order ones.

3.4. Around and beyond the Curry-Howard correspondence

For two decades, the Curry-Howard correspondence has been limited to the intuitionistic case but since 1990, an important stimulus spurred on the community following Griffin's discovery that this correspondence was extensible to classical logic. The community then started to investigate unexplored potential connections between computer science and logic. One of these fields is the computational understanding of Gentzen's sequent calculus while another one is the computational content of the axiom of choice.

3.4.1. Control operators and classical logic

Indeed, a significant extension of the Curry-Howard correspondence has been obtained at the beginning of the 90's thanks to the seminal observation by Griffin [79] that some operators known as control operators were typable by the principle of double negation elimination ($\neg\neg A \Rightarrow A$), a principle that enables classical reasoning.

Control operators are used to jump from one location of a program to another. They were first considered in the 60's by Landin [96] and Reynolds [101] and started to be studied in an abstract way in the 80's by Felleisen *et al* [75], leading to Parigot's $\lambda\mu$-calculus [99], a reference calculus that is in close Curry-Howard correspondence with classical natural deduction. In this respect, control operators are fundamental pieces to establish a full connection between proofs and programs.

3.4.2. Sequent calculus

The Curry-Howard interpretation of sequent calculus started to be investigated at the beginning of the 90's. The main technicality of sequent calculus is the presence of *left introduction* inference rules, for which two kinds of interpretations are applicable. The first approach interprets left introduction rules as construction rules for a language of patterns but it does not really address the problem of the interpretation of the implication connective. The second approach, started in 1994, interprets left introduction rules as evaluation context formation rules. This line of work led in 2000 to the design by Hugo Herbelin and Pierre-Louis Curien of a symmetric calculus exhibiting deep dualities between the notion of programs and evaluation contexts and between the standard notions of call-by-name and call-by-value evaluation semantics.

3.4.3. Abstract machines

Abstract machines came as an intermediate evaluation device, between high-level programming languages and the computer microprocessor. The typical reference for call-by-value evaluation of λ-calculus is Landin's SECD machine [95] and Krivine's abstract machine for call-by-name evaluation [92], [91]. A typical abstract machine manipulates a state that consists of a program in some environment of bindings and some evaluation context traditionally encoded into a "stack".

3.4.4. Delimited control

Delimited control extends the expressiveness of control operators with effects: the fundamental result here is a completeness result by Filinski [76]: any side-effect expressible in monadic style (and this covers references, exceptions, states, dynamic bindings, ...) can be simulated in λ-calculus equipped with delimited control.

3.5. Effective higher-dimensional algebra

3.5.1. Higher-dimensional algebra

Like ordinary categories, higher-dimensional categorical structures originate in algebraic topology. Indeed, ∞-groupoids have been initially considered as a unified point of view for all the information contained in the homotopy groups of a topological space X: the *fundamental ∞-groupoid*$\Pi(X)$ of X contains the elements of X as 0-dimensional cells, continuous paths in X as 1-cells, homotopies between continuous paths as 2-cells, and so on. This point of view translates a topological problem (to determine if two given spaces X and Y are homotopically equivalent) into an algebraic problem (to determine if the fundamental groupoids $\Pi(X)$ and $\Pi(Y)$ are equivalent).

In the last decades, the importance of higher-dimensional categories has grown fast, mainly with the new trend of *categorification* that currently touches algebra and the surrounding fields of mathematics. Categorification is an informal process that consists in the study of higher-dimensional versions of known algebraic objects (such as higher Lie algebras in mathematical physics [59]) and/or of "weakened" versions of those objects, where equations hold only up to suitable equivalences (such as weak actions of monoids and groups in representation theory [74]).

The categorification process has also reached logic, with the introduction of homotopy type theory. After a preliminary result that had identified categorical structures in type theory [88], it has been observed recently that the so-called "identity types" are naturally equiped with a structure of ∞-groupoid: the 1-cells are the proofs of equality, the 2-cells are the proofs of equality between proofs of equality, and so on. The striking resemblance with the fundamental ∞-groupoid of a topological space led to the conjecture that homotopy type theory could serve as a replacement of set theory as a foundational language for different fields of mathematics, and homotopical algebra in particular.

3.5.2. Higher-dimensional rewriting

Higher-dimensional categories are algebraic structures that contain, in essence, computational aspects. This has been recognised by Street [106], and independently by Burroni [64], when they have introduced the concept of *computad* or *polygraph* as combinatorial descriptions of higher categories. Those are directed presentations of higher-dimensional categories, generalising word and term rewriting systems.

In the recent years, the algebraic structure of polygraph has led to a new theory of rewriting, called *higher-dimensional rewriting*, as a unifying point of view for usual rewriting paradigms, namely abstract, word and term rewriting [93], [97], [80], [81], and beyond: Petri nets [83] and formal proofs of classical and linear logic have been expressed in this framework [82]. Higher-dimensional rewriting has developed its own methods to analyse computational properties of polygraphs, using in particular algebraic tools such as derivations to prove termination, which in turn led to new tools for complexity analysis [62].

3.5.3. Squier theory

The homotopical properties of higher categories, as studied in mathematics, are in fact deeply related to the computational properties of their polygraphic presentations. This connection has its roots in a tradition of using rewriting-like methods in algebra, and more specifically in the works of Anick [57] and Squier [105], [104]: Squier has proved that, if a monoid M can be presented by a *finite*, *terminating* and *confluent* rewriting system, then its third integral homology group $H_3(M, \mathbb{Z})$ is finitely generated and the monoid M has *finite derivation type* (a property of homotopical nature). This allowed him to conclude that finite convergent rewriting systems were not a universal solution to decide the word problem of finitely generated monoids. Since then, Yves Guiraud and Philippe Malbos have shown that this connection was part of a deeper unified theory when formulated in the higher-dimensional setting [14], [15], [85], [86], [87].

In particular, the computational content of Squier's proof has led to a constructive methodology to produce, from a convergent presentation, *coherent presentations* and *polygraphic resolutions* of algebraic structures, such as monoids [14] and algebras [32]. A coherent presentation of a monoid M is a 3-dimensional combinatorial object that contains not only a presentation of M (generators and relations), but also higher-dimensional cells, corresponding each to two fundamentally different proofs of the same equality: this is, in essence, the same as the proofs of equality of proofs of equality in homotopy type theory. When this process of "unfolding" proofs of equalities is pursued in every dimension, one gets a polygraphic resolution of the starting monoid M. This object has the following desirable qualities: it is free and homotopically equivalent to M (in the canonical model structure of higher categories [94], [58]). A polygraphic resolution of an algebraic object X is a faithful formalisation of X on which one can perform computations, such as homotopical or homological invariants of X. In particular, this has led to new algorithms and proofs in representation theory [11], and in homological algebra [84][32].

4. Highlights of the Year

4.1. Highlights of the Year

A one-day scientific meeting in honour of Pierre-Louis Curien's retirement was held at Université Paris Diderot on se 6, 2019 (organisers Antonio Bucciarelli, Bérénice Delcroix-Oger and Thomas Ehrhard) (https://www.irif.fr/plcmeeting).

The paper [35] presents the results of a large collaborative work led by Matthieu Sozeau on the metatheory and implementation of Coq's type theory in Coq itself.

Yves Guiraud defended his habilitation thesis on the 18th of June 2019, entitled "Rewriting methods in higher algebra". Yann Régis-Gianas defended his habilitation thesis on the 22nd of November 2019 entitled "About some metamorphoses of computer programs".

Yves Guiraud was granted an Action Exploratoire, Réécriture Algébrique, to start in January 2020. Emilio Gallego Arias joined the team in November 2019 on a Starting Research Position.

5. New Software and Platforms

5.1. 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 8.10 contains two major new features: support for a native fixed-precision integer type and a new sort SProp of strict propositions. It is also the result of refinements and stabilization of previous features, deprecations or removals of deprecated features, cleanups of the internals of the system and API, and many documentation improvements. This release includes many user-visible changes, including deprecations that are documented in the next subsection, and new features that are documented in the reference manual.

Version 8.10 is the fifth release of Coq developed on a time-based development cycle. Its development spanned 6 months from the release of Coq 8.9. Vincent Laporte is the release manager and maintainer of this release. This release is the result of 2500 commits and 650 PRs merged, closing 150+ issues.

See the Zenodo citation for more information on this release: https://zenodo.org/record/3476303#.Xe54f5NKjOQ

NEWS OF THE YEAR: Coq 8.10.0 contains:

- some quality-of-life bug fixes, - a critical bug fix related to template polymorphism, - native 63-bit machine integers, - a new sort of definitionally proof-irrelevant propositions: SProp, - private universes for opaque polymorphic constants, - string notations and numeral notations, - a new simplex-based proof engine for the tactics lia, nia, lra and nra, - new introduction patterns for SSReflect, - a tactic to rewrite under binders: under, - easy input of non-ASCII symbols in CoqIDE, which now uses GTK3.

All details can be found in the user manual.

- Participants: Yves Bertot, Frédéric Besson, Maxime Denes, Emilio Jesús Gallego Arias, Gaëtan Gilbert, Jason Gross, Hugo Herbelin, Assia Mahboubi, Érik Martin-Dorel, Guillaume Melquiond, Pierre-Marie Pédrot, Michael Soegtrop, Matthieu Sozeau, Enrico Tassi, Laurent Théry, Théo Zimmermann, Theo Winterhalter, Vincent Laporte, Arthur Charguéraud, Cyril Cohen, Christian Doczkal and Chantal Keller
- Partners: CNRS - Université Paris-Sud - ENS Lyon - Université Paris-Diderot
- Contact: Matthieu Sozeau
- URL: http://coq.inria.fr/

5.2. Equations

KEYWORDS: Coq - Dependent Pattern-Matching - Proof assistant - Functional programming

SCIENTIFIC DESCRIPTION: Equations is a tool designed to help with the definition of programs in the setting of dependent type theory, as implemented in the Coq proof assistant. Equations provides a syntax for defining programs by dependent pattern-matching and well-founded recursion and compiles them down to the core type theory of Coq, using the primitive eliminators for inductive types, accessibility and equality. In addition to the definitions of programs, it also automatically derives useful reasoning principles in the form of propositional equations describing the functions, and an elimination principle for calls to this function. It realizes this using a purely definitional translation of high-level definitions to core terms, without changing the core calculus in any way, or using axioms.

FUNCTIONAL DESCRIPTION: Equations is a function definition plugin for Coq (supporting Coq 8.8 to 8.10, with special support for the Coq-HoTT library), that allows the definition of functions by dependent pattern-matching and well-founded, mutual or nested structural recursion and compiles them into core terms. It automatically derives the clauses equations, the graph of the function and its associated elimination principle.

Equations is based on a simplification engine for the dependent equalities appearing in dependent eliminations that is also usable as a separate tactic, providing an axiom-free variant of dependent destruction. The main features of Equations include:

Dependent pattern-matching in the style of Agda/Epigram, with inaccessible patterns, with and where clauses. The use of the K axiom or a proof of K is configurable, and it is able to solve unification problems without resorting to the K rule if not necessary.

Support for well-founded and mutual recursion using measure/well-foundedness annotations, even on indexed inductive types, using an automatic derivation of the subterm relation for inductive families.

Support for mutual and nested structural recursion using with and where auxilliary definitions, allowing to factor multiple uses of the same nested fixpoint definition. It proves the expected elimination principles for mutual and nested definitions.

Automatic generation of the defining equations as rewrite rules for every definition.

Automatic generation of the unfolding lemma for well-founded definitions (requiring only functional extensionality).

Automatic derivation of the graph of the function and its elimination principle. In case the automation fails to prove these principles, the user is asked to provide a proof.

A new dependent elimination tactic based on the same splitting tree compilation scheme that can advantageously replace dependent destruction and sometimes inversion as well. The as clause of dependent elimination allows to specify exactly the patterns and naming of new variables needed for an elimination.

A set of Derive commands for automatic derivation of constructions from an inductive type: its signature, no-confusion property, well-founded subterm relation and decidable equality proof, if applicable.

RELEASE FUNCTIONAL DESCRIPTION: This version of Equations is based on an improved simplification engine for the dependent equalities appearing during dependent eliminations that is also usable as a separate dependent elimination tactic, providing an axiom-free variant of dependent destruction and a more powerful form of inversion. See http://mattam82.github.io/Coq-Equations/equations/2019/01/28/1.2beta.html and the following release notes for more information.

NEWS OF THE YEAR: Equations 1.2 was first released in may this year, after 3 years of development. It provides a refined simplification engine based on the work published at ICFP'19 (see the "Equations Reloaded" paper for details). The system has been improved to also work in the setting of Homotopy Type Theory and provides a more expressive source language and robust dependent elimination tactics.

- Participants: Matthieu Sozeau and Cyprien Mangin
- Contact: Matthieu Sozeau
- Publications: Equations reloaded - Equations for Hereditary Substitution in Leivant's Predicative System F: A Case Study - Equations: A Dependent Pattern-Matching Compiler
- URL: http://mattam82.github.io/Coq-Equations/

5.3. Rewr

Rewriting methods in algebra

KEYWORDS: Computer algebra system (CAS) - Rewriting systems - Algebra

Functional Description: Rewr is a prototype of computer algebra system, using rewriting methods to compute resolutions and homotopical invariants of monoids. The library implements various classical constructions of rewriting theory (such as completion), improved by experimental features coming from Garside theory, and allows homotopical algebra computations based on Squier theory. Specific functionalities have been developed for usual classes of monoids, such as Artin monoids and plactic monoids.

News Of The Year: Rewr has been extended with the experimental KGB completion algorithm, based on Knuth-Bendix completion procedure improved by techniques coming from Garside theory.

- Participants: Yves Guiraud and Samuel Mimram
- Contact: Yves Guiraud
- Publications: Higher-dimensional categories with finite derivation type - Higher-dimensional normalisation strategies for acyclicity - Coherent presentations of Artin monoids - A Homotopical Completion Procedure with Applications to Coherence of Monoids - Polygraphs of finite derivation type - Quadratic normalisation in monoids
- URL: http://www.lix.polytechnique.fr/Labo/Samuel.Mimram/rewr

5.4. Catex

Keywords: LaTeX - String diagram - Algebra

Functional Description: Catex is a Latex package and an external tool to typeset string diagrams easily from their algebraic expression. Catex works similarly to Bibtex.

News Of The Year: It is now possible to add labels to objects and morphisms

- Participant: Yves Guiraud
- Contact: Yves Guiraud
- URL: https://www.irif.fr/~guiraud/catex/catex.zip

5.5. Cox

Keywords: Computer algebra system (CAS) - Rewriting systems - Algebra

Functional Description: Cox is a Python library for the computation of coherent presentations of Artin monoids, with experimental features to compute the lower dimensions of the Salvetti complex.

- Participant: Yves Guiraud
- Contact: Yves Guiraud
- Publications: Coherent presentations of Artin monoids - A Homotopical Completion Procedure with Applications to Coherence of Monoids
- URL: https://www.irif.fr/~guiraud/cox/cox.zip

5.6. jsCoq

Keywords: Coq - Program verification - Interactive - Formal concept analysis - Proof assistant - Ocaml - Education - JavaScript

Functional Description: jsCoq is an Online Integrated Development Environment for the Coq proof assistant and runs in your browser! It aims to enable new UI/interaction possibilities and to improve the accessibility of the Coq platform itself.

RELEASE FUNCTIONAL DESCRIPTION: - Coq 8.10 support - Much improved interaction and general experience - Open / Save dialogs - AST and full serialization of Coq's datatypes - NPM packaging - Timeout support

- Participant: Emilio Jesus Gallego Arias
- Partners: Mines ParisTech - Technion, Israel Institute of Technology
- Contact: Emilio Jesus Gallego Arias
- Publication: jsCoq: Towards Hybrid Theorem Proving Interfaces
- URL: https://github.com/ejgallego/jscoq

5.7. coq-serapi

KEYWORDS: Interaction - Coq - Ocaml - Data centric - User Interfaces - GUI (Graphical User Interface) - Toolkit

FUNCTIONAL DESCRIPTION: SerAPI is a library for machine-to-machine interaction with the Coq proof assistant, with particular emphasis on applications in IDEs, code analysis tools, and machine learning. SerAPI provides automatic serialization of Coq's internal OCaml datatypes from/to JSON or S-expressions (sexps).

RELEASE FUNCTIONAL DESCRIPTION: - Support Coq 8.10 - Serialization of extensive AST - Serialization of kernel structures - Support for kernel traces [dumping and replay] - Tokenization of Coq documents - Serialization to JSON - Improved protocol and printing - Bug fixes

- Participant: Karl Palmskog
- Partner: KTH Royal Institute of Technology
- Contact: Emilio Jesus Gallego Arias
- Publication: SerAPI: Machine-Friendly, Data-Centric Serialization for COQ : Technical Report
- URL: https://github.com/ejgallego/coq-serapi

6. New Results

6.1. Effects in proof theory and programming

Participants: Kostia Chardonnet, Emilio Jesús Gallego Arias, Hugo Herbelin, Yann Régis-Gianas, Alexis Saurin, Exequiel Rivas Gadda.

6.1.1. A theory of effects and resources

In collaboration with Thomas Letan (ANSSI), Yann Régis-Gianas developed and proved several properties of a simple web server implemented in Coq using FreeSpec. This work will be presented at CPP 2020.

6.1.2. Call-by-need with probabilistic effects

As a follow up of Chardonnet's Master 1 internship, Kostia Chardonnet and Alexis Saurin continued investigating call-by-need calculi extended with probabilistic choice and started preliminary discussions with Claudia Faggian.

6.1.3. Proof-search, algebraically and graphically

Alexis Saurin worked on proof search in a proof-net scenario, that is proof-net search. A key aspect of proof construction is a management of non-determinism in bottom-up sequent-proof construction, be it when the search succeeds or when facing a failure and the need for backtracking. This is partially dealt with by focussing proof-construction, which reduces drastically the search space while retaining completeness of the resulting proof space (both at the provability level and at the denotational level).

His approach consists in viewing proof-search and sequentialisation as dual aspects of partial proof structures (that is proof nets with open premisses). In particular, he builds on Lafont's parsing criterion to obtain a proof-construction algorithm in which the proof space is not a search tree, as in sequent-calculus, but a dag allowing to share proof-construction paths.

Emilio Jesús Gallego Arias collaborates with Jim Lipton from Wesleyan University on the development of algebraic models for proof search.

6.2. Reasoning and programming with infinite data

Participants: Kostia Chardonnet, Lucien David, Abhishek De, Farzad Jafar-Rahmani, Luc Pellissier, Yann Régis-Gianas, Alexis Saurin.

This theme is part of the ANR project Rapido (see the National Initiatives section) which ended octobre 1st 2019.

6.2.1. Proof theory of non-wellfounded and circular proofs

6.2.1.1. Validity conditions of infinitary and circular proofs

In collaboration with David Baelde, Amina Doumane and Denis Kuperberg, Alexis Saurin extended the proof theory of infinite and circular proofs for fixed-point logics in various directions by relaxing the validity condition necessary to distinguish sound proofs from invalid ones. The original validity condition considered by Baelde, Doumane and Saurin in CSL 2016 rules out lots of proofs which are computationally and semantically sound and does not account for the cut-axiom interaction in sequent proofs. In the setting of sequent calculus, Alexis Saurin studied together with David Baelde, Amina Doumane and Denis Kuperberg a relaxed validity condition to allow infinite branches to be supported by threads which may leave the infinite branch, visiting other parts of the proofs and bouncing on axioms and cuts. This allows for a much more flexible criterion, inspired from Girard's geometry of interaction. The most general form of this criterion does not ensure productivity in the sequent calculus due to a discrepancy between the sequential nature of proofs in sequent calculus and the parallel nature of threads. David Baelde, Amina Doumane, Denis Kuperberg and Alexis Saurin provided a slight restriction of the full bouncing validity which grants productivity and validity of the cut-elimination process. This restriction still strictly extends previous notions of validity and is actually expressive enough to be undecidable.

Several directions of research have therefore been investigated from that point:

- Decidability can be recovered by constraining the shapes of bounces (bounding the depth of bounces). They actually exhibited a hierarchy of criteria, all decidable and satisfying the fact that their union corresponds to bouncing validity (which is therefore semi-decidable)
- While the result originaly held only for the multiplicative fragment of linear logic, the result was extended to multiplicative and additive linear logic.

Those results are currently submitted.

6.2.1.2. On the complexity of the validity condition of circular proofs

Alexis Saurin, together with Rémi Nollet and Christine Tasson, characterised the complexity of deciding the validity of circular proofs. While deciding validity was known to be in PSPACE, they proved that, for $\mu MALL$ proof, it is in fact a PSPACE-complete problem.

The proof is based on a deeper exploration of the connection between thread-validity and the size-change termination principle, a standard tool to prove program termination.

This result has been presented and published at TABLEAUX 2019 [41].

6.2.1.3. Proof nets for non-wellfounded proofs

Abhishek De and Alexis Saurin set the basis of the theory of non-wellfounded and circular proofs nets (in the multiplicative setting). Non-wellfounded proof nets, aka infinets, were defined extending Curien's presentation of proof nets allowing for a smooth extension to fixed point logics. The aim of this work is to provide a notion of canonical proof obiects for circular proofs free from the irrelevant details of the syntax of the sequent calculus. The first results were published in TABLEAUX 2019 [38] and provide a correctness condition for an infinet to be sequentialisable in a sequent proof.

The results of the TABLEAUX paper are limited in that they only address the case of proofs with finitely many cuts inferences. Abhishek De and Alexis Saurin are currently investigating, with Luc Pellissier, the general case of infinitely many cut in order to then lift the results from straight thread validity to bouncing thread validity.

6.2.2. On the denotational semantics of non-wellfounded proofs

Farzad Jafar-Rahmani started his PhD under the supervision of Thomas Ehrhard and Alexis Saurin in October 2019. His PhD work will focus on the denotational semantics of circular proofs of linear logic with fixed points. After working on the denotational semantics of finitary proofs for linear logic with fixed points (with Kozen rules) during his master, he is currently working at understanding the denotational counterpart of the validity condition of circular proofs.

6.2.3. Towards inductive and coinductive types in quantum programming languages

Kostia Chardonnet started his PhD under the supervision of Alexis Saurin and Benoît Valiron in November 2019. Previously, he did his MPRI Master internship under their joint supervision on designing a calculus of reversible programs with inductive and coinductive types. His research focused on extending a languages of type isomorphisms with inductive and coinductive types and understanding the connections of those reversible programs with $\mu MALL$ type isomorphisms and more specifically with $\mu MALL$ focused circular proof isomorphisms. In his PhD, he shall extend this to the case of a quantum programming language with inductive and coinductive data types.

6.2.4. Theory of fixed points in the lambda-calculus

The results of Alexis Saurin in collaboration with Giulio Manzonetto, Andrew Polonsky and Jacob Grue Simonsen, on two long-standing conjectures on fixed points in the λ-calculus – the "fixpoint property" and the "double-fixpoint conjecture" – have now appeared in the Journal of Logic and Computation [34]. The former asserts that every λ-term admits either a unique or an infinite number of β-distinct fixpoints while the second, formulated by Statman, says that there is no fixpoint satisfying $Y\delta = Y$ for $\delta = \lambda y, x.x(yx)$. They proved the first conjecture in the case of open terms and refute it in the case of sensible theories (instead of β). Moreover, they provide sufficient conditions for both conjectures in the general case. Concerning the double-fixpoint conjecture, they propose a proof technique identifying two key properties from which the results would follow, while they leave as conjecture to prove that those actually hold.

6.3. Effective higher-dimensional algebra

Participants: Antoine Allioux, Pierre-Louis Curien, Alen Durić, Eric Finster, Yves Guiraud, Amar Hadzihasanović, Cédric Ho Thanh, Matthieu Sozeau.

6.3.1. Rewriting methods in higher algebra

Yves Guiraud has completed a four-year collaboration with Eric Hoffbeck (Univ. Paris 13) and Philippe Malbos (Univ. Lyon 1), whose aim was to develop a theory of rewriting in associative algebras, with a view towards applications in homological algebra. They adapted the known notion of polygraph [64] to higher-dimensional associative algebras, and used these objects to develop a rewriting theory on associative algebras that generalises the two major tools for computations in algebras: Gröbner bases [63] and Poincaré-Birkhoff-Witt bases [100]. Then, they transposed the construction of [14], based on an extension of Squier's theorem [104] in higher dimensions, to compute small polygraphic resolutions of associative algebras from convergent presentations. Finally, this construction has been related to the Koszul homological property, yielding necessary or sufficient conditions for an algebra to be Koszul. The resulting work was published in Mathematische Zeitschrift [32].

Yves Guiraud has written and defended his “Habilitation à diriger des recherches” manuscript, as a survey on rewriting methods in algebra based on Squier theory [26]. The defense was held in June 2019.

Yves Guiraud works with Dimitri Ara (Univ. Aix-Marseille), Albert Burroni, Philippe Malbos (Univ. Lyon 1), François Métayer (Univ. Nanterre) and Samuel Mimram (École Polytechnique) on a reference book on the theory of polygraphs and higher-dimensional categories, and their applications in rewriting theory and homotopical algebra.

Yves Guiraud works with Marcelo Fiore (Univ. Cambridge) on the theoretical foundations of higher-dimensional algebra, in order to develop a common setting to develop rewriting methods for various algebraic structures at the same time. Practically, they aim at a definition of polygraphic resolutions of monoids in monoidal categories, based on the recent notion of n-oid in an n-oidal category. This theory will subsume the known cases of monoids and associative algebras, and encompass a wide range of objects, such as Lawvere theories (for term rewriting), operads (for Gröbner bases) or higher-order theories (for the λ-calculus).

Building on [9], Yves Guiraud is currently finishing with Matthieu Picantin (Univ. Paris Diderot) a work that generalises already known constructions such as the bar resolution, several resolutions defined by Dehornoy and Lafont [73], and the main results of Gaussent, Guiraud and Malbos on coherent presentations of Artin monoids [11], to monoids with a Garside family. This allows an extension of the field of application of the rewriting methods to other geometrically interesting classes of monoids, such as the dual braid monoids.

Still with Matthieu Picantin, Yves Guiraud develops an improvement of the classical Knuth-Bendix completion procedure, called the KGB (for Knuth-Bendix-Garside) completion procedure. The original algorithm tries to compute, from an arbitrary terminating rewriting system, a finite convergent presentation, by adding relations to solve confluence issues. Unfortunately, this algorithm fails on standard examples, like most Artin monoids with their usual presentations. The KGB procedure uses the theory of Tietze transformations, together with Garside theory, to also add new generators to the presentation, trying to reach the convergent Garside presentation identified in [9]. The KGB completion procedure is partially implemented in the prototype Rewr, developed by Yves Guiraud and Samuel Mimram.

Yves Guiraud has started a collaboration with Najib Idrissi (IMJ-PRG, Univ. Paris Diderot) whose aim is to understand the relation between several different methods known to compute small resolutions of algebras and operads: those based on rewriting methods (Anick, Squier) and those that stem from Koszul duality theory.

6.3.2. Normalisation of monoids

Alen Durić started his Phd thesis (supervised by Yves Guiraud and Pierre-Louis Curien) in October 2019. His work so far has been mostly bibliographical. The goal is to combine methods from rewriting theory (and in particular the method of homotopical completion and reduction developed by Guiraud-Malbos-Mimram) and methods developed by Dehornoy and his coauthors in the study of monoids with Garside families, and by Dehornoy-Guiraud in the study of normalisation for monoids. Alen Durić is currently experimenting with some examples taken from these latter works, with the goal of building coherent presentations for them using the former methods.

6.3.3. Topological aspects of polygraphs

Amar Hadzihasanović joined the team at the end of November 2019, as a one-year postdoc funded by FSMP. He has been working intensively on the study of shapes appropriate for the description of higher cells as needed in various approaches to higher categories and higher structures. Amar Hadzihasanović’s project is to recast his ideas in the framework of polygraphs, with the aim of bringing topological insights into the study of higher-dimensional rewriting.

6.3.4. Opetopes

The work of Pierre-Louis Curien, Cédric Ho Thanh and Samuel Mimram on syntactic and type-theoretic presentations of opetopes and opetopic sets has been submitted to a journal, and a short version has been presented at the LICS conference in Vancouver this year [45].

Cédric Ho Thanh, in collaboration with Chaitanya Leena Subramaniam, has defined the notion of "opetopic algebras" that leverages the subtle combinatorics of opetopes. This framework encompasses categories, planar operads, and Loday's combinads over planar trees. They have defined an opetopic nerve functor that fully embeds each category of opetopic algebras into the category of opetopic sets. In particular, they obtain fully faithful opetopic nerve functors for categories and for planar coloured operads. These results have been written up in [51]. This work is the first in a series aimed at using opetopic spaces as models for higher algebraic structures. In particular, the aim is to provide new models for infinity-categories and infinity-operads.

6.3.5. Foundations and formalisation of higher algebra

Antoine Allioux (PhD started in February 2018), Eric Finster, Yves Guiraud and Matthieu Sozeau are exploring the development of higher algebra in type theory. To formalise higher algebra, one needs a new source of coherent structure in type theory. During the first year of Allioux's PhD, they studied an internalisation of polynomial monads (of which opetopes and ∞-categories are instances) in type theory, which ought to provide such a coherent algebraic structure, inspired by the work of Kock et al [90]. They later realised that this internalisation is however incoherent as presented in pure type theory, essentially because of its reliance on equality types. Since then, they switched to a different view, describing opetopes as an external construction and relying on strict equalities in the metatheory to avoid the coherence problem. Opetopic type theory should then be, similarly to cubical type theory, a type theory indexed over these opetopic structures, where grafting and substitution are computional operations. They are now concentrating on showing that the modified inductive characterisation of opetopes and their algebras, still definable in type theory, gives rise to the standard notion of opetopes in mathematics, an original result in itself.

6.4. Incrementality

Participants: Thibaut Girka, Yann Régis-Gianas.

In collaboration with Paolo Giarrusso (EPFL, Switzerland), Philipp Shuster (Univ. of Tübingen, Germany), Yann Régis-Gianas developed a new method to incrementalise higher-order programs using formal derivatives and static caching. Yann Régis-Gianas has developed a mechanised proof for this transformation as well as a prototype language featuring efficient derivatives for functional programs. A paper has been presented at ESOP 2019 in Prague. Yann Régis-Gianas also presented this work at several places (Gallium seminar, Galinette seminar, and Chocola seminar).

In collaboration with Olivier Martinot (Paris Diderot), Yann Régis-Gianas studied a new technique to implement incrementalised operations on lists.

In collaboration with Faridath Akinotcho (Paris Diderot), Yann Régis-Gianas studied an incrementalisation of the Earley parsing algorithm.

6.5. Metatheory and development of Coq

Participants: Félix Castro, Emilio Jesús Gallego Arias, Gaëtan Gilbert, Hugo Herbelin, Pierre Letouzey, Cyprien Mangin, Thierry Martinez, Yann Régis-Gianas, Matthieu Sozeau, Théo Winterhalter, Théo Zimmermann.

6.5.1. Meta-programming and Metatheory of Coq

The MetaCoq project started last year, providing the means to program program transformations and general purpose plugins in Coq, using a quoting/unquoting mechanism. This year, they extended the framework to specify the theory, including the reduction, cumulativity and typing relations of the Polymorphic, Cumulative Calculus of Inductive Constructions at the basis of Coq. Matthieu Sozeau, together with Simon Boulier, Nicolas Tabareau and Théo Winterhalter at Galinette, Cyril Cohen at Marelle, Yannick Forster and Fabian Kunze at the University of Saarbrucken and Abhishek Anand and Gregory Malecha at BedRock Systems, Inc co-authored [54] a full description of the resulting theory (to appear in JAR). This allows for the verification of term manipulations with respect to typing: syntactic translations but also reflexive tactics glue code can hence be verified. The article also develops an alternative extraction mode to OCaml allowing the efficient compilation and execution of meta-programs written in the Template Monad. An example partial extraction of Coq programs to call-by-value pure lambda-calculus is developed this way.

Following up on this work, Matthieu Sozeau led a metatheoretical study of Coq in Coq, proving the basic metatheoretical properties of the typing relation, and developed together with Yannick Forster (Saarbrucken) and Simon Boulier, Nicolas Tabareau and Théo Winterhalter (Gallinette) verified correct versions of type-checking and erasure for a large subset of Coq. This work involved the production of a fully-precise specification for the type theory implemented by Coq, cleaning up the previously untested typing specification, and variants of the algorithms used in its kernel ammenable to proofs of correctness. The corresponding implementations can be extracted and provide an alternative, verified checker for Coq terms, that can run on medium-sized examples. This work will be presented [35] at POPL in New Orleans in January 2020.

6.5.2. Homotopy type theory

Hugo Moeneclaey started in September 2019 a PhD on the syntax of spheres in homotopy type theory, under the supervision of Hugo Herbelin.

Hugo Herbelin and Hugo Moeneclaey worked on the syntax of a variant of Cohen, Coquand, Huber and Mörtberg's Cubical Type Theory justified by an iterated parametricity model where equality on types is defined to be equivalence of types, thus satisfying univalence by construction.

6.5.3. Computational contents of the axiom of choice

Hugo Herbelin developed in collaboration with Nuria Brede (U. Potsdam) a unified logical structure for choice and bar induction principles.

6.5.4. Computational contents of Gödel's constructible universe

Félix Castro started his PhD under the supervision of Hugo Herbelin and Alexandre Miquel in September 2019. His PhD work will focus on the computational contents of Gödel's constructible universe. Previously, he worked on the formalisation of the ramified analytical hierarchy in classical second-order arithmetic.

6.5.5. Dependent pattern-matching and recursion

Together with Cyprien Mangin, Matthieu Sozeau refined the treatment of dependent pattern-matching in the Equations plugin. By carefully studying the type of equalities between indexed inductive types, he devised a new criterion for the elimination of equalities between inductive families based on the notion of forced arguments of constructors, resulting in a simplification of the setup of Cockx and Devriese [68] for simplification of dependent pattern-matching without K. This improved simplifier is part of the latest version of the Equations plugin, which also provides better support for the definition of mutual and well-founded recursive definitions on indexed inductive types. This work was presented at ICFP 2019 in Berlin [36]. A longer journal version is in preparation, along with a dedicated tutorial on Equations slated for inclusion in a new volume of the Software Foundations series dedicated to advanced tools.

Thierry Martinez continued part time the implementation of a dependent pattern-matching compilation algorithm in Coq based on the PhD thesis work of Pierre Boutillier and on the internship work of Meven Bertrand.

6.5.6. Software engineering aspects of the development of Coq

Théo Zimmermann has studied software engineering and open collaboration aspects of the development of Coq.

Following the migration of the Coq bug tracker from Bugzilla to GitHub which he conducted in 2017, he analysed data (extracted through the GitHub API), in collaboration with Annalí Casanueva Artís from the Paris School of Economics. The results show an increased number of bugs by core developers and an increased diversity of the people commenting bug reports. These quantitative results were completed with qualitative data coming from interviews with main Coq developers, which help interpret them. They validate *a posteriori* the usefulness of such a switch. A paper [43] has been published at ICSME 2019, which is the leading conference on the topic of Software Maintenance and Evolution.

Besides, Théo Zimmermann also studied and influenced the pull-based model that is now used for the development of Coq, he improved the release management process and tools, he studied package distribution and maintenance, in particular with the foundation of the coq-community organisation in 2018, which has taken off by attracting 19 maintainers, and hosting 25 projects. All of these topics are presented in the PhD thesis [28] that he defended in December 2019.

Emilio J. Gallego Arias and Théo Zimmerman took the roles of release managers for the Coq 8.12 and will oversee this release, planned for mid-2020.

Emilio J. Gallego Arias and Théo Zimmerman discussed on future plans for compositional proof checking using the Dune build system, which will include a new library format for Coq. The Dune team was informed, with Emilio J. Gallego Arias participating in the bi-weekly developer meetings. Emilio J. Gallego Arias also started discussion with the Debian OCaml maintainers (who are located at IRIF) as to see how to better integrate Dune with the Debian packaging workflow.

Emilio J. Gallego Arias designed the Coq instrumentation used in the [103] paper, which collects and analyses changes to proof scripts.

Emilio J. Gallego Arias and Karl Palmskog released a new version of the Coq SerAPI tool, which has been used in some recent proof engineering efforts, such as [65], the machine-learning environments CoqGYM and Proverbot9001 [108], [55], offering state of the art proof automation after training with proof data sets, and the educational user interface WaterProof [56]. SerAPI has also been used in some other works undergoing review and thus yet not public.

Emilio J. Gallego Arias and Shachar Itzhaky released a new version of the educational Coq frontend jsCoq [10], and assisted a few users who have been preparing courses using it.

Emilio J. Gallego Arias maintains an ongoing collaboration with the Deducteam group at Inria Saclay on the topic of interactive proof methods and standards; this has resulted in the release of an experimental LSP server for the Lambdapi theorem prover.

Emilio J. Gallego Arias, Hugo Herbelin, and Théo Zimmerman participate in the Logipedia project led by Gilles Dowek, which aims to develop a standard proof interchange format.

6.5.7. Software Infrastructure

Emilio J. Gallego Arias did significant work to refactor the Coq codebase in preparation for further work on incremental and multi-core aware type checking.

6.5.8. Dissemination activities

Emilio J. Gallego Arias and Théo Zimmerman organised the Coq meetup, an after-work event targeting industry and other communities outside academia.

6.5.9. Coordination of the development of Coq

Hugo Herbelin, Matthieu Sozeau, Emilio J. Gallego Arias and Théo Zimmermann, helped by members from Gallinette (Nantes) and Marelle (Sophia-Antipolis), devoted an important part of their time to coordinate the development, to review propositions of extensions of Coq from external and/or young contributors, and to propose themselves extensions.

6.6. Formalisation and verification

Participants: Pierre-Louis Curien, Lucien David, Emilio Jesús Gallego Arias, Kailiang Ji, Pierre Letouzey, Jean-Jacques Lévy, Cyprien Mangin, Daniel de Rauglaudre, Yann Régis-Gianas, Alexis Saurin, Matthieu Sozeau.

6.6.1. Proofs and surfaces

The joint work of Pierre-Louis Curien with Jovana Obradović (former PhD student of the team and now postdoc in Prague), Zoran Petrić and other Serbian colleagues on formalising proofs of incidence theorems (arising by repeated use of Menelaus theorem) by means of a cyclic sequent calculus, has been submitted to a journal, and has been presented at the conference Topology, Algebra, and Categories in Logic (TACL) 2019, Nice, in June 2019 [53].

6.6.2. A Coq formalisation of the first-order predicate calculus

In relation with a logic course for master students, Pierre Letouzey made a Coq formalisation of the first-order predicate calculus. The logical rules are expressed in a natural deduction style (with explicit contexts). Pierre Letouzey proposed two low-level representations of formulas : one based on quantifiers with names, the other using "locally nameless" techniques. The equivalence between the two settings has been proved correct. Using this deep embedding, Pierre Letouzey formalised in Coq the whole course notes (prepared some years ago by Alexandre Miquel), including the completeness theorem for this logic. This development is available at https://gitlab.math.univ-paris-diderot.fr/letouzey/natded.

6.6.3. A Coq formalisation of circular proofs and their validity condition

During the summer 2019, Alexis Saurin supervised Lucien David's M1 internship on formalizing in Coq circular proofs and their meta-theory. This work built on Xavier Onfroy's previous work as well as on Pierre Letouzey's formalisation of the predicate calculus in natural deduction mentioned above. While the previous work by Xavier Onfroy was both contributing to the proof theory part and the ω-automata part (which is need for the decidability theorem), Lucien David completely focused on the the proof theory side. In particular, he was able to improve significantly on Xavier Onfroy's formalisation by using ideas from Letouzey's formalisation of natural deduction and by interacting with Pierre Letouzey and Alexis Saurin. This development is available at https://github.com/LuluDavid/CircularProofsValidity.

6.6.4. Lexing and regular expressions in Coq

Pierre Letouzey and Yann Régis-Gianas revisited in Coq classical techniques about lexing and regular expressions. In particular, regular expressions (with complement and conjunction) have been formalised, as well as their Brzozowski derivatives, and the finiteness theorem due to Brzozowski : a given regular expression admits only a finite number of derivatives (up to some equivalence). Both the general equivalence (based on language identity) and practical approximations (similarities) has been considered (and proved decidable). From that, the algorithms building recognizing automata (with derivatives as states) have been formalised and proved, leading to the minimal automata when using the general equivalence (but at a high cost), or to practical approximations of the minimal automata when using various similarities. This work is still ongoing. For instance, the correctness proof of a particular similarity used in an existing implementation (ml-ulex) is quite elusive for the moment. They also plan to extend this development up to a full-scale tool a la ocamllex in Coq.

6.6.5. Real Numbers as sequences of digits in Coq

Daniel de Rauglaudre has been continuing the formalisation of real numbers defined as sequences of digits in any radix with the LPO axiom/oracle (Limited Principle of Omniscience). Although the operations (additions and multiplications) work with this method, the proof of associativity of addition needs more work to be achieved. This development is available at https://github.com/roglo/coq_real/.

6.6.6. Category theory in Coq

Daniel de Rauglaudre started an implementation in Coq of Category theory in Coq, using in particular theorems coming from HOTT (HOmotopy Type theory) that he implemeted some years ago. Several notions around Categories have been defined. For example, Yoneda Lemma, among others. This development is available at https://github.com/roglo/mycoqhott/.

6.6.7. Number theory in Coq

Daniel de Rauglaudre started and almost completed the formalisation in Coq of the proof of Euler's Product Formula, stating that the Riemann zeta function, which is a sum on all the natural numbers, is also a product on all the prime numbers. He also added several theorems about the prime numbers. This development is available at https://github.com/roglo/coq_euler_prod_form.

6.6.8. Proofs of algorithms on graphs

Jean-Jacques Lévy and Chen Ran (a PhD student at the Institute of Software, Beijing) pursued their work about formal proofs of graph algorithms. Their goal is to provide proofs of algorithms checked by computer and human readable. In 2019, they presented at ITP 2019 a joint paper with Cyril Cohen, Stephan Merz and Laurent Théry on this work [37]. This article compared formal proofs in three different systems (Why3, Coq, Isabelle/HOL) of Tarjan (1972) linear-time algorithm computing the strongly connected components in directed graphs.

The current work is to have a proof of the implementation of this algorithm with imperative programming and memory pointers. They also planed to produce formal proofs of other abstract algorithms such as the Hopcroft-Tarjan (1972) linear-time algorithm for planarity testing in undirected graphs.

6.6.9. Certified compilation and meta-programming

Matthieu Sozeau participates to the CertiCoq project led by Andrew Appel at Princeton (https://www.cs.princeton.edu/~appel/certicoq) whose aim is to verify a compiler from Coq's Gallina language down to CompCert C-light which provides itself a certified compilation path to assembly language. Together with Yannick Forster at the University of Saarbrucken and the MetaCoq team, Matthieu Sozeau focused the verification of type-checking and erasure which were previously trusted parts of the system. The new verified erasure function fills a gap in the proof of correctness of compilation from Gallina terms down to C-light. The whole compiler can be run on realistic examples (the erasure phase does take most of the compilation time and should be optimised further).

In collaboration with Xavier Denis (Paris Diderot), Yann Régis-Gianas formalised and built a compiler for Mtac2. A paper is in preparation.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

Theo Zimmermann will start a research engineer position in January 2020 to continue his research and development work about improving the Software Engineering practices of the development of Coq, especially to continue the improvement of the collaborative development processes and of its ecosystem. This position is funded by the Inria-NomadicLabs grant.

8. Partnerships and Cooperations

8.1. National Initiatives

Pierre-Louis Curien, Emilio J. Gallego Arias, Yves Guiraud, Hugo Herbelin, and Alexis Saurin are members of the GDR Informatique Mathématique, in the LHC (Logique, Homotopie, Catégories) and Scalp (Structures formelles pour le calcul et les preuves) working groups. Alexis Saurin is coordinator of the Scalp working group.

Pierre-Louis Curien, Yves Guiraud (local coordinator until Sept. 2019) and Matthieu Sozeau are members of the GDR Topologie Algébrique, federating French researchers working on classical topics of algebraic topology and homological algebra, such as homotopy theory, group homology, K-theory, deformation theory, and on more recent interactions of topology with other themes, such as higher categories and theoretical computer science.

Yves Guiraud is member of the GDR Tresses, federating French researchers working on algebraic, algorithmic and topological aspects of braid groups, low-dimensional topology, and connected subjects.

Yves Guiraud will coordinate the four-year Action Exploratoire Inria Réal (Réécriture Algébrique), starting in January 2020. Its aim is to continue the unification of rewriting-like methods in abtract and higher algebra, with a view toward applications in homological and higher algebra, and group and representation theory. This investigation is pursued in immersion at IMJ-PRG, the fundamental maths common laboratory of Sorbonne Université and Université Paris Diderot.

Emilio J. Gallego Arias is a member of the GDR Génie de la Programation et du Logiciel, in the LTP (Langages, Types et Preuves) group.

Yann Régis-Gianas collaborates with Mitsubishi Rennes on the topic of differential semantics. This collaboration led to the CIFRE grant for the PhD of Thibaut Girka.

Yann Régis-Gianas collaborates with ANSSI on the topic of certified ful programming in Coq.

Yann Régis-Gianas collaborates with Nomadic Labs on the topic of certified smart contract compilation.

Yann Régis-Gianas is a member of the ANR COLIS dedicated to the verification of Linux Distribution installation scripts. This project is joint with members of VALS (Univ Paris Sud) and LIFL (Univ Lille).

Yann Régis-Gianas and Alexis Saurin (coordinator) are members of the four-year RAPIDO ANR project, started in January 2015 and ended in September 2019. RAPIDO aims at investigating the use of proof-theoretical methods to reason and program on infinite data objects. The goal of the project is to develop logical systems capturing infinite proofs (proof systems with least and greatest fixpoints as well as infinitary proof systems), to design and to study programming languages for manipulating infinite data such as streams both from a syntactical and semantical point of view. Moreover, the ambition of the project is to apply the fundamental results obtained from the proof-theoretical investigations (i) to the development of software tools dedicated to the reasoning about programs computing on infinite data, *e.g.* stream programs (more generally coinductive programs), and (ii) to the study of properties of automata on infinite words and trees from a proof-theoretical perspective with an eye towards model-checking problems. Other permanent members of the project are Christine Tasson from IRIF (PPS team), David Baelde from LSV, ENS-Cachan, and Pierre Clairambault, Damien Pous and Colin Riba from LIP, ENS-Lyon.

Matthieu Sozeau is a member of the CoqHoTT project led by Nicolas Tabareau (Gallinette team, Inria Nantes & École des Mines de Nantes), funded by an ERC Starting Grant, ending in 2020. The PhD grant of Antoine Allioux is funded by the CoqHoTT ERC.

8.2. European Initiatives

8.2.1. Collaborations in European Programs, Except FP7 & H2020

- Program: COST
- Project acronym: EUTypes
- Project title: The European research network on types for programming and verification
- Duration: March 2016 - March 2020
- Coordinator: Herman Geuvers
- Other partners: 29 countries

- Abstract: This COST promotes (1) the synergy between theoretical computer scientists, logicians and mathematicians to develop new foundations for 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.

8.3. International Initiatives

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

Pierre-Louis Curien and Claudia Faggian are members of the CRECOGI associate team, coordinated on one side by Ugo dal Lago (research-team FoCUS, Inria Sophia and Bologna), and on the other side by Ichiro Hasuo (NII, Tokyo). The full name of the project is Concurrent, Resourceful and full Computation, by Geometry of Interaction. This project was renewed in 2019 for a duration of two years.

Presentation of CRECOGI: Game semantics and geometry of interaction (GoI) are two closely related frameworks whose strengh is to have the characters of both a denotational and an operational semantics. They offer a high-level, mathematical (denotational) interpretation, but are interactive in nature. The formalisation in terms of movements of tokens through which programs communicate with each other can actually be seen as a low-level program. The current limit of GoI is that the vast majority of the literature and of the software tools designed around it have a pure, sequential functional language as their source language. This project aims at investigating the application of GoI to concurrent, resourceful, and effectful computation, thus paving the way to the deployment of GoI-based correct-by-construction compilers in real-world software developments in fields like (massively parallel) high-performance computing, embedded and cyberphysical systems, and big data. The presence of both the Japanese GoI community (whose skills are centered around effects and coalgebras) and the French GoI community (more focused on linear logic and complexity analysis) bring essential, complementary, ingredients.

8.3.2. Inria International Partners

8.3.2.1. Participation in International Programs

Pierre-Louis Curien and Alexis Saurin are members of CNRS GDRI-LL a french-italian network on linear logic community in France and Italy.

8.3.2.2. International Initiatives

Pierre-Louis Curien is principal investigator on the French side for a joint project Inria - Chinese Academy of Sciences. The project's title is "Verification, Interaction, and Proofs" (December 2017 – December 2020). The principal investigator on the Chinese side is Ying Jiang, from the Institute of Software (ISCAS) in Beijing. The participants of the project on the French side are Pierre-Louis Curien and Jean-Jacques Lévy, as well as other members of IRIF (Thomas Ehrhard, Jean Krivine, Giovanni Bernardi, Ahmed Bouajjani, Mihaela Sighireanu, Constantin Enea, Gustavo Petri), and Gilles Dowek (Deducteam team of Inria Saclay). On the Chinese side, the participants are Ying Jiang, as well as other members of the ISCAS (Angsheng Li, Xinxin Liu, Yi Lü, Peng Wu, Yan Rongjie, Zhilin Wu, and Wenhui Zhang), and Yuxi Fu (from Shanghai Jiaotong University).

8.4. International Research Visitors

8.4.1. Research Stays Abroad

Matthieu Sozeau visited the Programming Languages group of Benjamin Pierce at the University of Pennsylvania in June and July 2019, along with visits at MIT and Princeton to other members of the NSF DeepSpec project.

Pierre-Louis Curien visited East China Normal University (ECNU), Shanghai, for a month from early October to early December 2019 as invited professor.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. Member of the Organizing Committees

Alexis Saurin, as coordinator of GT-Scalp is member of the organizing committee of the annual meeting of the working group. In 2019, the annual meeting took place in ENS Lyon in October 2019: https://www.irif.fr/gt-scalp/journees-2019.

9.1.1.2. Member of the Steering Committees

Pierre-Louis Curien is member of the steering committee of the international workshop Games for Logic and Programming Languages (GaLop).

9.1.2. Scientific Events: Selection

9.1.2.1. Member of the Conference Program Committees

Pierre-Louis Curien is member of the program committee of Formal Structures for Computation and Deduction 2020, Paris, June-July 2020.

Yann Régis-Gianas was member of the program committee of Software Language Engineering 2019.

Alexis Saurin was a member of the program committee of Circularity in Syntax and Semantics, held in Gothenbug in November 2019.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Pierre-Louis Curien is editor in chief of the Cambridge University Press journal Mathematical Structures in Computer Science (since January 2016). Alexis Saurin is a guest editor of the special issue of MSCS on structural proof theory, automated reasoning and computation in celebration of Dale Miller's 60th birthday published in September 2019.

9.1.3.2. Reviewer - Reviewing Activities

Pierre Letouzey has done several reviews for the Journal of Functional Programming (JFP) and for Science of Computer Programming. Yann Régis-Gianas reviewed a paper for Transactions on Programming Languages and Systems (TOPLAS).

9.1.4. Invited Talks

Hugo Herbelin gave an invited talk on the logical structure of choice and bar induction principles at the Proof, Computation and Complexity workshop (Mittag-Leffler institute). Yann Régis-Gianas gave an invited talk at the DeepSpec workshop of PLDI 2019 about FreeSpec.

9.1.5. Conferences and schools attended

Pierre-Louis Curien presented his joint work on Proofs and surfaces at the conference Topology, Algebra, and Categories in Logic (TACL) 2019, Nice, in June 2019.

Pierre-Louis Curien and Yves Guiraud attended the workshop Homotopy meets homology, Dublin, May 2019.

Pierre-Louis Curien and Yves Guiraud attended the *Higher Structures* conference, Marseille, January 2019.

Pierre-Louis Curien and Cédric Ho Thanh attended the Category Theory 2019 conference, Edinburgh, July 2019.

Pierre-Louis Curien attended Martin Hofmann Memorial Meeting, Munich, July 2019.

Abhishek De attended the Proof Society summer school at Swansea in September 2019.

Cédric Ho Thanh attended the Summer school on Topology, Algebra, and Categories in Logic at Ile de Porquerolles in June 2019.

Hugo Herbelin and Hugo Moeneclaey attended the Types Conference, Oslo, June 2019.

Hugo Moeneclaey attended the International Conference on Homotopy Type Theory (HoTT 2019) and the associated Homotopy Type Theory Summer School, Pittsburgh, August 2019.

Hugo Herbelin attended the Third Workshop on Mathematical Logic and its Applications, Nancy, March 2019.

Hugo Herbelin attended the Workshop Facets of Realizability, Cachan, July 2019.

9.1.6. Leadership within the Scientific Community

9.1.6.1. Scientific Expertise

Hugo Herbelin was a member of the HCERES committee evaluating the MICS laboratory of Centrale Supelec.

9.1.6.2. Research Administration

Pierre-Louis Curien has been a member of the Scientific Council of the CIRM (Centre International de Rencontres Mathématiques) until February 2019.

Pierre-Louis Curien was a member of the scientific council of the Computer Science department of University Paris Diderot until May 2019.

Yves Guiraud was the coordinator of the Proofs, Programs and Systems (PPS) pole at IRIF, and a member of the IRIF comité de direction, until August 2019. Since September 2019, Hugo Herbelin coordinates the PPS pole.

Yves Guiraud was a member of the Conseil Scientifique of the CS department of Univ. Paris Diderot until June 2019.

Yves Guiraud is a member of the Conseil Scientifique of the Math department of Univ. Paris Diderot since April 2019.

Alexis Saurin is a coordinator of the Scalp working group of GDR Informatique Mathématique: https://www.irif.fr/gt-scalp/index

Alexis Saurin is member of the "commission recherche" of both Faculté des Sciences and Faculté Société et humanités at Université de Paris.

Alexis Saurin and Hugo Herbelin are members of the IRIF conseil de laboratoire.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Master: Matthieu Sozeau, Assistants de Preuve, 18h ETD, MPRI, University Paris Diderot.
- Master: Pierre-Louis Curien, Models of programming languages: domains, categories, games (together with Thomas Ehrhard and Paul-André Melliès), 15h, MPRI, University Paris Diderot.
- Master level: Pierre-Louis Curien taught a course on the Foundations of Programming Languages at East China Normal University (4 hours, November-December 2019).
- Master: Hugo Herbelin taught a class on Homotopy Type Theory (together with Nicolas Tabareau), 24h, LMFI, University Paris Diderot.
- Master: Pierre Letousey teaches two short courses to the LMFI Master 2 students : "Programming in Coq" and "Introduction to computed-aided formal proofs". These two courses come in addition to Pierre Letousey's regular duty as teacher in the Computer Science department of University Paris Diderot (including a course on Compilation to M2-Pro students and a course on computed-aided formal proofs to M1 students).
- Master: Yann Régis-Gianas, Functional programming and Type Systems, 12h, MPRI (in addition to four other courses at the master level in University Paris Diderot).

- Master: Alexis Saurin, Outils classiques pour la correspondance de Curry-Howard, 24H, LMFI.
- Summer school: Pierre-Louis Curien gave an introductory mini-course on Homotopy type theory at the Fifteenth International Tbilisi Summer School in Logic and Language, Georgia, in September 2019 (3h).
- Summer school: Hugo Herbelin gave a lecture on type theory at the EUTypes Summer School, Ohrid, in September 2019 (4h).
- Yann Régis-Gianas organised the 4th session of the OCaml MOOC on the FUN platform.
- Yann Régis-Gianas organised the “Journée Francilienne de Programmation”, a programming contest between the first years students of University Paris Sorbonne, University Paris Saclay and University Paris Diderot.

9.2.2. Supervision

9.2.2.1. PhD Supervision

- PhD started: Félix Castro, Computational contents of the model of constructible sets, cotutelle between Université Paris Diderot and Universidad de la República (Montevideo, Uruguay), started in September 2019, supervised by Hugo Herbelin and Alexandre Miquel.
- PhD Started: Kostia Chardonnet, Inductive and coinductive types in quantum programming languages, Université Paris Saclay, started in November 2019, supervised by Alexis Saurin and Benoît Valiron.
- PhD started: Alen Durić, Normalisation for monoids and higher categories, Université Paris Diderot, started in October 2O19, supervised by Yves Guiraud and Pierre-Louis Curien.
- PhD started: Colin Gonzalez, Certified compilation of spreadsheets as Tezos smart contract, CIFRE, started in September 2019, supervised by Yann Régis-Gianas and Benjamin Canou (Nomadic Labs).
- PhD started: Farzad Jafar-Rahmani, Denotational semantics of circular and non-well-founded proofs, Université Paris Diderot, started in October 2019, supervised by Thomas Ehrhard and Alexis Saurin.
- PhD started: Hugo Moeneclaey, Syntax of spheres in homotopy type theory, Université Paris Diderot, started in September 2019, supervised by Hugo Herbelin.
- PhD in progress: Antoine Allioux, Opetopes in Type Theory, Université Paris Diderot, since March 2018, supervised by Yves Guiraud and Matthieu Soseau.
- PhD in progress: Abhishek De, Proof-nets for fixed-point logics and non-well-founded proofs, Université Paris Diderot, since October 2018, supervised by Alexis Saurin.
- PhD in progress: Cédric Ho Thanh, Opetopes for higher-dimensional rewriting and koszulity, Université Paris Diderot, since September 2017, supervised by Pierre-Louis Curien and Samuel Mimram.
- PhD in progress: Lucas Massoni Sguerra, Formal verification of mechanism design, Université PSL, since January 2018, supervised by Gérard Memmi, Fabien Coehlo, and Emilio J. Gallego Arias.
- PhD in progress: Rémi Nollet, Validity conditions for circular proofs, Université Paris Diderot, since October 2016, supervised by Alexis Saurin and Christine Tasson.
- PhD in progress: Théo Winterhalter, Metatheory and certified implementation of type theories, Université de Nantes, since September 2017, supervised by Nicolas Tabareau and Matthieu Sozeau.
- PhD ended: Gaëtan Gilbert, A type theory with definitional proof-irrelevance, École Nationale Supérieure Mines-Télécom Atlantique Bretagne Pays de la Loire - IMT Atlantique, defended on 20 December 2019, supervised by Nicolas Tabareau and Matthieu Sozeau.
- PhD ended: Théo Zimmermann, Challenges in the collaborative evolution of a proof language and its ecosystem, Université Paris Diderot, defended on 12 December 2019, supervised by Hugo Herbelin and Yann Régis-Gianas.

9.2.2.2. Internships

- Hugo Herbelin supervised the master internship of Adrien Champougny on a type-theoretic presentation of Primitive Recursive Arithmetic.
- Hugo Herbelin supervised the master internship of Hugo Moeneclaey on a syntax and semantics for cubical type theory with definitional univalence.
- Yann Régis-Gianas supervised the third-year licence internship of Alexandre Moine about the detection of clones in OCaml programs.
- Yann Régis-Gianas supervised the first-year master internship of Astyax Nourel and Sébastien Lecleire about the implementation of an instance server for Learn-OCaml.
- Yann Régis-Gianas cosupervised with Thomas Letan (ANSSI) the master internship of Vincent Tourneur about the certification of a Unix utility using FreeSpec.
- Alexis Saurin supervised the first-year master internship of Lucien David about Coq formalisation of circular proof systems.
- Alexis Saurin cosupervised with Benoît Valiron the MPRI internship of Kostia Chardonnet on (co)inductive types in reversible programming languages.

9.2.3. Juries

- Hugo Herbelin was reviewer of the PhD of Ulysse Gérard.
- Hugo Herbelin was president of the PhD committee of Kenji Maillard.
- Alexis Saurin was member of the jury of the entrance competition to Écoles normales supérieures, in charge of the computer science exam (informatique-mathématiques), for the last three years.

9.3. Popularisation

9.3.1. Internal or external Inria responsibilities

- Yves Guiraud is "correspondant communication" for IMJ-PRG at INSMI.
- Yann Régis-Gianas is "correspondant communication" for IRIF at INS2I.
- Jean-Jacques Lévy is member of the Inria-Alumni's executive committee (4 meetings in 2019).

9.3.2. Education

- Yann Régis-Gianas presented the Learn-OCaml project at the OCaml workshop, at the OCaml Meetup, at the University of Lisbon and at the JFLA.
- Yann Régis-Gianas gave a talk at the conference of CPGE teachers of computer science to present the Learn-OCaml project.

9.3.3. Interventions

- Jean-Jacques Lévy was invited by the French Academy of Sciences to participate to the Nanjing Tech Week (June 22-24) and to visit the new Qingdao International Academician Park (June 28). He also visited the Institute of Software of the Chinese Academy of Sciences (ISCAS) on July 1-3.
- Jean-Jacques Lévy as a member of the AEFC (Association des Experts France Chine) was invited to the 2019 Hangzhou International Human Resources Exchanges and Cooperation Conference (Hangzhou, November 9-11). He also visited the East China Normal University (ECNU) Shanghai (November 12-15) and Jinan University at Guangzhou (November 18-19).
- Jean-Jacques Lévy gave a talk about the Four Ages of Computer Science at the Huashang College in Guangzhou (November 18).

9.3.4. Internal action

Jean-Jacques Lévy talked about "L'informatique en 4 temps" at the Café des Sciences at Inria Rocquencourt (February 12) and to the Journées Européennes du Patrimoine at Inria-Rocquencourt (21 September).

10. Bibliography

Major publications by the team in recent years

[1] R. M. Amadio, Y. Régis-Gianas. *Certifying and reasoning about cost annotations of functional programs*, in "Higher-Order and Symbolic Computation", January 2013, https://hal.inria.fr/inria-00629473

[2] Z. Ariola, H. Herbelin, A. Sabry. *A Type-Theoretic Foundation of Delimited Continuations*, in "Higher Order and Symbolic Computation", 2007, http://dx.doi.org/10.1007/s10990-007-9006-0

[3] D. Baelde, A. Doumane, A. Saurin. *Infinitary proof theory : the multiplicative additive case* , in "Proceedings of CSL 2016", September 2016, https://hal.archives-ouvertes.fr/hal-01339037

[4] C. Chenavier. *The lattice of reduction operators: applications to noncommutative Gröbner bases and homological algebra*, Université paris Diderot, December 2016, https://tel.archives-ouvertes.fr/tel-01415910

[5] P.-L. Curien. *Operads, clones, and distributive laws*, in "Operads and Universal Algebra : Proceedings of China-France Summer Conference", Tianjin, China, C. Bai, L. Guo, J.-L. Loday (editors), Nankai Series in Pure, Applied Mathematics and Theoretical Physics, Vol. 9, World Scientific, July 2010, p. 25-50, https://hal.archives-ouvertes.fr/hal-00697065

[6] P.-L. Curien, R. Garner, M. Hofmann. *Revisiting the categorical interpretation of dependent type theory*, in "Theoretical computer Science", 2014, vol. 546, p. 99-119, http://dx.doi.org/10.1007/s10990-007-9006-0

[7] P.-L. Curien, H. Herbelin. *The duality of computation*, in "Proceedings of the Fifth ACM SIGPLAN International Conference on Functional Programming (ICFP '00)", Montreal, Canada, SIGPLAN Notices 35(9), ACM, September 18-21 2000, p. 233–243 [*DOI :* 10.1145/351240.351262], http://hal.archives-ouvertes.fr/inria-00156377/en/

[8] P.-L. Curien, H. Herbelin. *Abstract machines for dialogue games*, in "Interactive models of computation and program behavior", Panoramas et Synthèses, Société Mathématique de France, 2009, p. 231-275, https://hal.archives-ouvertes.fr/hal-00155295

[9] P. Dehornoy, Y. Guiraud. *Quadratic normalization in monoids*, in "Internat. J. Algebra Comput.", 2016, vol. 26, n^{o} 5, p. 935–972, https://doi.org/10.1142/S0218196716500399

[10] E. J. Gallego Arias, B. Pin, P. Jouvelot. *jsCoq: Towards Hybrid Theorem Proving Interfaces*, in "Proceedings of the 12th Workshop on User Interfaces for Theorem Provers, Coimbra, Portugal, 2nd July 2016", S. Autexier, P. Quaresma (editors), Electronic Proceedings in Theoretical Computer Science, Open Publishing Association, 2017, vol. 239, p. 15-27, http://dx.doi.org/10.4204/EPTCS.239.2

[11] S. Gaussent, Y. Guiraud, P. Malbos. *Coherent presentations of Artin monoids*, in "Compositio Mathematica", 2015, vol. 151, n^{o} 5, p. 957-998 [*DOI :* 10.1112/S0010437X14007842], https://hal.archives-ouvertes.fr/hal-00682233

[12] G. Gilbert, J. Cockx, M. Sozeau, N. Tabareau. *Definitional Proof-Irrelevance without K*, in "46th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2019", Lisbon, Portugal, POPL, January 2019, https://hal.inria.fr/hal-01859964

[13] T. GIRKA, D. MENTRÉ, Y. RÉGIS-GIANAS.*Oracle-based Dierential Operational Semantics (long version)*, Université Paris Diderot / Sorbonne Paris Cité, October 2016, https://hal.inria.fr/hal-01419860

[14] Y. GUIRAUD, P. MALBOS.*Higher-dimensional normalisation strategies for acyclicity*, in "Advances in Mathematics", 2012, vol. 231, n^{o} 3-4, p. 2294-2351 [*DOI :* 10.1016/J.AIM.2012.05.010], https://hal.archives-ouvertes.fr/hal-00531242

[15] Y. GUIRAUD, P. MALBOS, S. MIMRAM.*A Homotopical Completion Procedure with Applications to Coherence of Monoids*, in "RTA - 24th International Conference on Rewriting Techniques and Applications - 2013", Eindhoven, Netherlands, F. VAN RAAMSDONK (editor), Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, June 2013, vol. 21, p. 223-238 [*DOI :* 10.4230/LIPIcs.RTA.2013.223], https://hal.inria.fr/hal-00818253

[16] H. HERBELIN.*On the Degeneracy of Sigma-Types in Presence of Computational Classical Logic*, in "Proceedings of TLCA 2005", P. URZYCZYN (editor), Lecture Notes in Computer Science, Springer, 2005, vol. 3461, p. 209–220

[17] H. HERBELIN.*An intuitionistic logic that proves Markov's principle*, in "Logic In Computer Science", Edinburgh, Royaume-Uni, IEEE Computer Society, 2010, http://hal.inria.fr/inria-00481815/en/

[18] H. HERBELIN.*A Constructive Proof of Dependent Choice, Compatible with Classical Logic*, in "LICS 2012 - 27th Annual ACM/IEEE Symposium on Logic in Computer Science", Dubrovnik, Croatia, Proceedings of the 27th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2012, 25-28 June 2012, Dubrovnik, Croatia, IEEE Computer Society, June 2012, p. 365-374, https://hal.inria.fr/hal-00697240

[19] G. JABER, N. TABAREAU, M. SOZEAU.*Extending Type Theory with Forcing*, in "LICS 2012 : Logic In Computer Science", Dubrovnik, Croatia, June 2012, https://hal.archives-ouvertes.fr/hal-00685150

[20] P. LETOUZEY.*Hofstadter's problem for curious readers*, Université Paris Diderot ; Inria Paris-Rocquencourt, September 2015, 29, https://hal.inria.fr/hal-01195587

[21] G. MUNCH-MACCAGNONI.*Focalisation and Classical Realisability*, in "Computer Science Logic '09", E. GRÄDEL, R. KAHLE (editors), Lecture Notes in Computer Science, Springer-Verlag, 2009, vol. 5771, p. 409–423

[22] T. U. F. PROGRAM.*Homotopy type theory—univalent foundations of mathematics*, The Univalent Foundations Program, Princeton, NJ; Institute for Advanced Study (IAS), Princeton, NJ, 2013, xiv+589, http://homotopytypetheory.org/book

[23] Y. RÉGIS-GIANAS, F. POTTIER.*A Hoare Logic for Call-by-Value Functional Programs*, in "Proceedings of the Ninth International Conference on Mathematics of Program Construction (MPC'08)", Lecture Notes in Computer Science, Springer, July 2008, vol. 5133, p. 305–335, http://gallium.inria.fr/~fpottier/publis/regis-gianas-pottier-hoarefp.ps.gz

[24] A. SAURIN.*Separation with Streams in the $\Lambda\mu$-calculus*, in "Symposium on Logic in Computer Science (LICS 2005)", Chicago, IL, USA, Proceedings, IEEE Computer Society, 26-29 June 2005, p. 356-365

[25] B. ZILIANI, M. SOZEAU.*A comprehensible guide to a new unifier for CIC including universe polymorphism and overloading*, in "Journal of Functional Programming", 2017, vol. 27 [*DOI :* 10.1017/S0956796817000028], https://hal.inria.fr/hal-01671925

Publications of the year

Doctoral Dissertations and Habilitation Theses

[26] Y. GUIRAUD.*Rewriting methods in higher algebra*, Université Paris 7, June 2019, Habilitation à diriger des recherches, https://hal.archives-ouvertes.fr/tel-02161197

[27] Y. RÉGIS-GIANAS.*About some Metamorphoses of Computer Programs*, Université Paris Diderot, November 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02405839

[28] T. ZIMMERMANN.*Challenges in the collaborative evolution of a proof language and its ecosystem*, Université de Paris, December 2019, https://hal.inria.fr/tel-02451322

Articles in International Peer-Reviewed Journal

[29] C. CHENAVIER.*A Lattice Formulation of the F 4 Completion Procedure*, in "International Journal of Algebra and Computation", 2019, https://hal.archives-ouvertes.fr/hal-01489200

[30] C. CHENAVIER.*Syzygies among reduction operators*, in "Journal of Pure and Applied Algebra", 2019, https://arxiv.org/abs/1708.08709 , https://hal.archives-ouvertes.fr/hal-01578555

[31] G. GILBERT, J. COCKX, M. SOZEAU, N. TABAREAU.*Definitional Proof-Irrelevance without K*, in "Proceedings of the ACM on Programming Languages", January 2019, p. 1-28 [*DOI :* 10.1145/329031610.1145/3290316], https://hal.inria.fr/hal-01859964

[32] Y. GUIRAUD, E. HOFFBECK, P. MALBOS.*Convergent presentations and polygraphic resolutions of associative algebras*, in "Mathematische Zeitschrift", 2019, vol. 293, n^{o} 1-2, p. 113-179, https://arxiv.org/abs/1406.0815 [*DOI :* 10.1007/S00209-018-2185-Z], https://hal.archives-ouvertes.fr/hal-01006220

[33] Y. JIANG, J. LIU, G. DOWEK, K. JI.*Towards Combining Model Checking and Proof Checking*, in "The Computer Journal", 2019, forthcoming [*DOI :* 10.1093/COMJNL/BXY112], https://hal.inria.fr/hal-01970274

[34] G. MANZONETTO, A. POLONSKY, A. SAURIN, J. G. SIMONSEN.*The fixed point property and a technique to harness double fixed point combinators*, in "Journal of Logic and Computation", September 2019, vol. 29, n^{o} 5, p. 831-880 [*DOI :* 10.1093/LOGCOM/EXZ013], https://hal.archives-ouvertes.fr/hal-02408192

[35] M. SOZEAU, S. BOULIER, Y. FORSTER, N. TABAREAU, T. WINTERHALTER.*Coq Coq Correct! Verification of Type Checking and Erasure for Coq, in Coq*, in "Proceedings of the ACM on Programming Languages", January 2020, p. 1-28 [*DOI :* 10.1145/3371076], https://hal.archives-ouvertes.fr/hal-02380196

[36] M. SOZEAU, C. MANGIN.*Equations reloaded*, in "Proceedings of the ACM on Programming Languages", July 2019, vol. 3, n^{o} ICFP, p. 1-29 [*DOI :* 10.1145/3341690], https://hal.inria.fr/hal-01671777

International Conferences with Proceedings

[37] R. Chen, C. Cohen, J.-J. Levy, S. Merz, L. Théry. *Formal Proofs of Tarjan's Strongly Connected Components Algorithm in Why3, Coq and Isabelle*, in "ITP 2019 - 10th International Conference on Interactive Theorem Proving", Portland, United States, J. Harrison, J. O'Leary, A. Tolmach (editors), Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2019, vol. 141, p. 13:1 - 13:19 [*DOI :* 10.4230/LIPIcs.ITP.2019.13], https://hal.inria.fr/hal-02303987

[38] A. De, A. Saurin. *Infinets: The parallel syntax for non-wellfounded proof-theory*, in "TABLEAUX 2019 - 28th International Conference on Automated Reasoning with Analytic Tableaux and Related Methods", London, United Kingdom, TABLEAUX 2019 - 28th International Conference on Automated Reasoning with Analytic Tableaux and Related Methods, September 2019, https://hal.archives-ouvertes.fr/hal-02337286

[39] P. G. Giarrusso, Y. Régis-Gianas, P. Schuster. *Incremental λ-Calculus in Cache-Transfer Style Static Memoization by Program Transformation*, in "ESOP 2019 - European Symposium on Programming", Prague, Czech Republic, L. Caires (editor), Springer, April 2019, https://hal.inria.fr/hal-02405864

[40] T. Letan, Y. Régis-Gianas. *FreeSpec: Specifying, Verifying and Executing Impure Computations in Coq*, in "CPP 2020 - 9th ACM SIGPLAN International Conference on Certified Programs and Proofs", Nouvelle-Orléans, United States, ACM, January 2020, p. 1-15 [*DOI :* 10.1145/3372885.3373812], https://hal.inria.fr/hal-02422273

[41] R. Nollet, A. Saurin, C. Tasson. *PSPACE-Completeness of a Thread Criterion for Cyclic Proofs in Linear Logic with Least and Greatest Fixed Points*, in "TABLEAUX 2019 - 28th International Conference on Automated Reasoning with Analytic Tableaux and Related Methods", London, United Kingdom, Automated Reasoning with Analytic Tableaux and Related Methods - 28th International Conference, TABLEAUX 2019, London, UK, September 3-5, 2019, Proceedings, September 2019, https://hal.archives-ouvertes.fr/hal-02173207

[42] T. Winterhalter, M. Sozeau, N. Tabareau. *Eliminating Reflection from Type Theory : To the Legacy of Martin Hofmann*, in "CPP 2019 - 8th ACM SIGPLAN International Conference on Certified Programs and Proofs", Lisbonne, Portugal, ACM, January 2019, p. 91-103 [*DOI :* 10.1145/3293880.3294095], https://hal.archives-ouvertes.fr/hal-01849166

[43] T. Zimmermann, A. Casanueva Artís. *Impact of switching bug trackers: a case study on a medium-sized open source project*, in "ICSME 2019 - International Conference on Software Maintenance and Evolution", Cleveland, United States, September 2019, https://hal.inria.fr/hal-01951176

National Conferences with Proceeding

[44] C. Bozman, B. Canou, R. Di Cosmo, P. Couderc, L. Gesbert, G. Henry, F. Le Fessant, M. Mauny, C. Morel, L. Peyrot, Y. Régis-Gianas. *Learn-OCaml : un assistant à l'enseignement d'OCaml*, in "JFLA 2019 - Journées Francophones des Langages Applicatifs", Les Rousses, France, January 2019, https://hal.inria.fr/hal-02405876

Conferences without Proceedings

[45] C. Ho Thanh, P.-L. Curien, S. Mimram. *A Sequent Calculus for Opetopes*, in "LICS 2019 - Logic in computer science 2019", Vancouver, Canada, June 2019, https://hal.archives-ouvertes.fr/hal-02406569

Scientific Books (or Scientific Book chapters)

[46] D. BAELDE, A. FELTY, G. NADATHUR, A. SAURIN.*A special issue on structural proof theory, automated reasoning and computation in celebration of Dale Miller's 60th birthday*, Cambridge University Press, September 2019, vol. 29, n° 8, p. 1007-1008 [*DOI :* 10.1017/S0960129519000136], https://hal.archives-ouvertes.fr/hal-02408211

[47] Z. L. DARGAYE, Y. RÉGIS-GIANAS.*31ème Journées Francophones des Langages Applicatifs*, IRIF, January 2020, https://hal.inria.fr/hal-02427360

Research Reports

[48] N. JEANNEROD, C. MARCHÉ, Y. RÉGIS-GIANAS, M. SIGHIREANU, R. TREINEN.*Specification of UNIX Utilities*, ANR, October 2019, https://hal.inria.fr/hal-02321691

Other Publications

[49] P.-L. CURIEN, C. HO THANH, S. MIMRAM.*Syntactical approaches to opetopes*, March 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02064784

[50] H. HERBELIN, É. MIQUEY.*Continuation-and-environment-passing style translations: a focus on call-by-need*, January 2019, working paper or preprint, https://hal.inria.fr/hal-01972846

[51] C. HO THANH, C. LEENA SUBRAMANIAM.*Opetopic algebras I: Algebraic structures on opetopic sets*, November 2019, 38 pages, https://hal.archives-ouvertes.fr/hal-02343861

[52] C. HO THANH, C. LEENA SUBRAMANIAM.*Opetopic algebras III: Presheaf models of homotopy-coherent opetopic algebras*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02448208

[53] M. MILICEVIC, D. BARALIC, J. OBRADOVIC, Z. PETRIC, M. ZEKIC, R. ZIVALJEVIC, P.-L. CURIEN.*Proofs and surfaces*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02410910

[54] M. SOZEAU, A. ANAND, S. BOULIER, C. COHEN, Y. FORSTER, F. KUNZE, G. MALECHA, N. TABAREAU, T. WINTERHALTER.*The MetaCoq Project*, June 2019, working paper or preprint, https://hal.inria.fr/hal-02167423

References in notes

[55] *Proverbot, a bot for proving*, 2019, Accessed: 2019-12-09, https://github.com/UCSD-PL/proverbot9001

[56] *Waterproof: an educational environment for writing mathematical proofs in interactive notebooks*, 2019, Accessed: 2019-12-09, https://github.com/impermeable/waterproof

[57] D. J. ANICK.*On the Homology of Associative Algebras*, in "Trans. Amer. Math. Soc.", 1986, vol. 296, n° 2, p. 641–659

[58] D. ARA, F. MÉTAYER.*The Brown-Golasiński Model Structure on strict ∞-groupoids revisited*, in "Homology, Homotopy and Applications", 2011, vol. 13, n° 1, p. 121–142

[59] J. BAEZ, A. CRANS.*Higher-dimensional algebra. VI. Lie 2-algebras*, in "Theory Appl. Categ.", 2004, vol. 12, p. 492–538

[60] H. P. BARENDREGT.*The Lambda Calculus: Its Syntax and Semantics*, North Holland, Amsterdam, 1984

[61] Y. BERTOT, P. CASTÉRAN.*Interactive Theorem Proving and Program Development Coq'Art: The Calculus of Inductive Constructions*, Springer, 2004

[62] G. BONFANTE, Y. GUIRAUD.*Polygraphic Programs and Polynomial-Time Functions*, in "Logical Methods in Computer Science", 2009, vol. 5, n^{o} 2, p. 1–37

[63] B. BUCHBERGER.*Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal (An Algorithm for Finding the Basis Elements in the Residue Class Ring Modulo a Zero Dimensional Polynomial Ideal)*, Mathematical Institute, University of Innsbruck, Austria, 1965

[64] A. BURRONI.*Higher-dimensional word problems with applications to equational logic*, in "Theoretical Computer Science", jul 1993, vol. 115, n^{o} 1, p. 43–62

[65] A. CELIK, K. PALMSKOG, M. PAROVIC, E. J. GALLEGO ARIAS, M. GLIGORIC.MCOQ*: Mutation Proving for Analysis of Verification Projects*, in "Proceedings of the 34rd ACM/IEEE International Conference on Automated Software Engineering, ASE 2019", 2019

[66] A. CHLIPALA.*Certified Programming with Dependent Types - A Pragmatic Introduction to the Coq Proof Assistant*, MIT Press, 2013, http://mitpress.mit.edu/books/certified-programming-dependent-types

[67] A. CHURCH.*A set of Postulates for the foundation of Logic*, in "Annals of Mathematics", 1932, vol. 2, p. 33, 346-366

[68] J. COCKX, D. DEVRIESE.*Proof-relevant unification: Dependent pattern matching with only the axioms of your type theory*, in "J. Funct. Program.", 2018, vol. 28, e12, https://doi.org/10.1017/S095679681800014X

[69] T. COQUAND.*Une théorie des Constructions*, University Paris 7, January 1985

[70] T. COQUAND, G. HUET.*Constructions : A Higher Order Proof System for Mechanizing Mathematics*, in "EUROCAL'85", Linz, Lecture Notes in Computer Science, Springer Verlag, 1985, vol. 203

[71] 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

[72] H. B. CURRY, R. FEYS, W. CRAIG.*Combinatory Logic*, North-Holland, 1958, vol. 1, §9E

[73] P. DEHORNOY, Y. LAFONT.*Homology of Gaussian groups*, in "Ann. Inst. Fourier (Grenoble)", 2003, vol. 53, n^{o} 2, p. 489–540, http://aif.cedram.org/item?id=AIF_2003__53_2_489_0

[74] P. DELIGNE.*Action du groupe des tresses sur une catégorie*, in "Invent. Math.", 1997, vol. 128, n^{o} 1, p. 159–175

[75] M. FELLEISEN, D. P. FRIEDMAN, E. KOHLBECKER, B. F. DUBA.*Reasoning with continuations*, in "First Symposium on Logic and Computer Science", 1986, p. 131-141

[76] A. FILINSKI.*Representing Monads*, in "Conf. Record 21st ACM SIGPLAN-SIGACT Symp. on Principles of Programming Languages, POPL'94", Portland, OR, USA, ACM Press, 17-21 Jan 1994, p. 446-457

[77] G. GENTZEN.*Untersuchungen über das logische Schließen*, in "Mathematische Zeitschrift", 1935, vol. 39, p. 176–210,405–431

[78] J.-Y. GIRARD.*Une extension de l'interpretation de Gödel à l'analyse, et son application à l'élimination des coupures dans l'analyse et la théorie des types*, in "Second Scandinavian Logic Symposium", J. FENSTAD (editor), Studies in Logic and the Foundations of Mathematics, North Holland, 1971, n^o 63, p. 63-92

[79] T. G. GRIFFIN.*The Formulae-as-Types Notion of Control*, in "Conf. Record 17th Annual ACM Symp. on Principles of Programming Languages, POPL '90", San Francisco, CA, USA, 17-19 Jan 1990, ACM Press, 1990, p. 47–57

[80] Y. GUIRAUD.*Présentations d'opérades et systèmes de réécriture*, Univ. Montpellier 2, 2004

[81] Y. GUIRAUD.*Termination Orders for 3-Dimensional Rewriting*, in "Journal of Pure and Applied Algebra", 2006, vol. 207, n^o 2, p. 341–371

[82] Y. GUIRAUD.*The Three Dimensions of Proofs*, in "Annals of Pure and Applied Logic", 2006, vol. 141, n^o 1–2, p. 266–295

[83] Y. GUIRAUD.*Two Polygraphic Presentations of Petri Nets*, in "Theoretical Computer Science", 2006, vol. 360, n^o 1–3, p. 124–146

[84] Y. GUIRAUD, E. HOFFBECK, P. MALBOS.*Confluence of linear rewriting and homology of algebras*, in "3rd International Workshop on Confluence", Vienna, Austria, July 2014, https://hal.archives-ouvertes.fr/hal-01105087

[85] Y. GUIRAUD, P. MALBOS.*Higher-dimensional categories with finite derivation type*, in "Theory Appl. Categ.", 2009, vol. 22, n^o 18, p. 420-478

[86] Y. GUIRAUD, P. MALBOS.*Identities among relations for higher-dimensional rewriting systems*, in "Séminaires et Congrès, Société Mathématique de France", 2011, vol. 26, p. 145-161

[87] Y. GUIRAUD, P. MALBOS.*Coherence in monoidal track categories*, in "Math. Structures Comput. Sci.", 2012, vol. 22, n^o 6, p. 931–969

[88] M. HOFMANN, T. STREICHER.*The groupoid interpretation of type theory*, in "Twenty-five years of constructive type theory (Venice, 1995)", Oxford Logic Guides, Oxford Univ. Press, New York, 1998, vol. 36, p. 83–111

[89] W. A. HOWARD.*The formulae-as-types notion of constructions*, in "to H.B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism", Academic Press, 1980, Unpublished manuscript of 1969

[90] J. KOCK, A. JOYAL, M. BATANIN, J.-F. MASCARI.*Polynomial functors and opetopes*, in "Advances in Mathematics", 2010, vol. 224, n^o 6, p. 2690 - 2737 [*DOI :* 10.1016/J.AIM.2010.02.012], http://www.sciencedirect.com/science/article/pii/S0001870810000769

[91] J.-L. KRIVINE.*A call-by-name lambda-calculus machine*, in "Higher Order and Symbolic Computation", 2005

[92] J.-L. KRIVINE.*Un interpréteur du lambda-calcul*, 1986, Unpublished

[93] Y. LAFONT.*Towards an Algebraic Theory of Boolean Circuits*, in "Journal of Pure and Applied Algebra", 2003, vol. 184, p. 257-310

[94] Y. LAFONT, F. MÉTAYER, K. WORYTKIEWICZ.*A Folk Model Structure on Omega-Cat*, in "Advances in Mathematics", 2010, vol. 224, n^o 3, p. 1183–1231

[95] P. LANDIN.*The mechanical evaluation of expressions*, in "The Computer Journal", January 1964, vol. 6, n^o 4, p. 308–320

[96] P. LANDIN.*A generalisation of jumps and labels*, UNIVAC Systems Programming Research, August 1965, n^o ECS-LFCS-88-66, Reprinted in Higher Order and Symbolic Computation, 11(2), 1998

[97] P. MALBOS.*Critères de finitude homologique pour la non convergence des systèmes de réécriture de termes*, Univ. Montpellier 2, 2004

[98] P. MARTIN-LÖF.*A theory of types*, University of Stockholm, 1971, n^o 71-3

[99] M. PARIGOT.*Free Deduction: An Analysis of "Computations" in Classical Logic*, in "Logic Programming, Second Russian Conference on Logic Programming", St. Petersburg, Russia, A. VORONKOV (editor), Lecture Notes in Computer Science, Springer, September 11-16 1991, vol. 592, p. 361-380, http://www.informatik.uni-trier.de/~ley/pers/hd/p/Parigot:Michel.html

[100] S. B. PRIDDY.*Koszul resolutions*, in "Trans. Amer. Math. Soc.", 1970, vol. 152, p. 39–60

[101] J. C. REYNOLDS.*Definitional interpreters for higher-order programming languages*, in "ACM '72: Proceedings of the ACM annual conference", New York, NY, USA, ACM Press, 1972, p. 717–740

[102] J. C. REYNOLDS.*Towards a theory of type structure*, in "Symposium on Programming", B. ROBINET (editor), Lecture Notes in Computer Science, Springer, 1974, vol. 19, p. 408-423

[103] T. RINGER, A. SANCHEZ-STERN, D. GROSSMAN, S. LERNER.*REPLICA: REPL Analysis for Coq Instrumentation*, in "Certified Programs and Proofs (CPP 2019)", 2019

[104] C. SQUIER, F. OTTO, Y. KOBAYASHI.*A finiteness condition for rewriting systems*, in "Theoret. Comput. Sci.", 1994, vol. 131, n^o 2, p. 271–294

[105] C. C. SQUIER.*Word problems and a homological finiteness condition for monoids*, in "J. Pure Appl. Algebra", 1987, vol. 49, n^o 1-2, p. 201–217

[106] R. STREET.*Limits Indexed by Category-Valued 2-Functors*, in "Journal of Pure and Applied Algebra", 1976, vol. 8, p. 149–181

[107] T. C. D. TEAM.*The Coq Proof Assistant, version 8.7.1*, December 2017, https://doi.org/10.5281/zenodo.1133970

[108] K. YANG, J. DENG.*Learning to Prove Theorems via Interacting with Proof Assistants*, in "International Conference on Machine Learning", 2019

[109] N. DE BRUIJN.*AUTOMATH, a language for mathematics*, Technological University Eindhoven, November 1968, n^{o} 66-WSK-05

Project-Team POLSYS

Polynomial Systems

IN COLLABORATION WITH: Laboratoire d'informatique de Paris 6 (LIP6)

IN PARTNERSHIP WITH:
CNRS

Sorbonne Université (UPMC)

RESEARCH CENTER
Paris

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team POLSYS

Creation of the Team: 2012 January 01, updated into Project-Team: 2013 January 01, end of the Project-Team: 2019 December 31

Keywords:

Computer Science and Digital Science:

A2.4. - Formal method for verification, reliability, certification
A4.3. - Cryptography
A4.3.1. - Public key cryptography
A4.3.4. - Quantum Cryptography
A5.10.1. - Design
A6.1. - Methods in mathematical modeling
A6.2.3. - Probabilistic methods
A6.2.6. - Optimization
A6.2.7. - High performance computing
A6.4.3. - Observability and Controlability
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.3. - Geometry, Topology
A8.4. - Computer Algebra

Other Research Topics and Application Domains:

B5. - Industry of the future
B5.2. - Design and manufacturing
B5.2.3. - Aviation
B5.2.4. - Aerospace
B6. - IT and telecom
B6.3. - Network functions
B6.5. - Information systems
B9.5.1. - Computer science
B9.5.2. - Mathematics
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

Jean-Charles Faugère [Team leader, Inria, Senior Researcher, HDR]
Elias Tsigaridas [Inria, Researcher, until Aug 2019]
Dongming Wang [CNRS, Senior Researcher, HDR]

Faculty Members

Mohab Safey El Din [Team leader, Sorbonne Université, Professor, HDR]
Jérémy Berthomieu [Sorbonne Université, Associate Professor]
Ludovic Perret [Sorbonne Université, Associate Professor, HDR]

PhD Students

Matías Bender [Inria, PhD Student, until May 2019]
Olive Chakraborty [Sorbonne Université, PhD Student]
Solane El Hirch [Sorbonne Université, PhD Student]
Jocelyn Ryckeghem [DGA, PhD Student]
Thi Xuan Vu [Sorbonne Université, PhD Student]

Administrative Assistants
Christelle Guiziou [Inria, Administrative Assistant]
Mathieu Mourey [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall Objectives

The main focus of the POLSYS project is to solve systems of polynomial equations.

Our main objectives are:

- **Fundamental Algorithms and Structured Systems.** The objective is to propose fast exponential exact algorithms for solving polynomial equations and to identify large classes of structured polynomial systems which can be solved in polynomial time.
- **Solving Systems over the Reals and Applications.** For positive dimensional systems basic questions over the reals may be very difficult (for instance testing the existence of solutions) but also very useful in applications (e.g. global optimization problems). We plan to propose efficient algorithms and implementations to address the most important issues: computing sample points in the real solution sets, decide if two such sample points can be path-connected and, as a long term objective, perform quantifier elimination over the reals (computing a quantifier-free formula which is equivalent to a given quantified boolean formula of polynomial equations/inequalities).
- **Dedicated Algebraic Computation and Linear Algebra.** While linear algebra is a key step in the computation of Gröbner bases, the matrices generated by the algorithms F_4/F_5 have specific structures (quasi block triangular). The objective is to develop a dedicated efficient multi-core linear algebra package as the basis of a future open source library for computing Gröbner bases.
- **Solving Systems in Finite Fields, Applications in Cryptology and Algebraic Number Theory.** We propose to develop a systematic use of *structured systems* in Algebraic Cryptanalysis. We want to improve the efficiency and to predict the theoretical complexity of such attacks. We plan to demonstrate the power of algebraic techniques in new areas of cryptography such as Algebraic Number Theory (typically, in curve based cryptography).

3. Research Program

3.1. Introduction

Polynomial system solving is a fundamental problem in Computer Algebra with many applications in cryptography, robotics, biology, error correcting codes, signal theory, ... Among all available methods for solving polynomial systems, computation of Gröbner bases remains one of the most powerful and versatile method since it can be applied in the continuous case (rational coefficients) as well as in the discrete case (finite fields). Gröbner bases are also building blocks for higher level algorithms that compute real sample points in the solution set of polynomial systems, decide connectivity queries and quantifier elimination over the reals. The major challenge facing the designer or the user of such algorithms is the intrinsic exponential behaviour of the complexity for computing Gröbner bases. The current proposal is an attempt to tackle these issues in a number of different ways: improve the efficiency of the fundamental algorithms (even when the complexity is exponential), develop high performance implementation exploiting parallel computers, and investigate new classes of structured algebraic problems where the complexity drops to polynomial time.

3.2. Fundamental Algorithms and Structured Systems

Participants: Jérémy Berthomieu, Jean-Charles Faugère, Mohab Safey El Din, Elias Tsigaridas, Dongming Wang, Matías Bender, Thi Xuan Vu.

Efficient algorithms F_4/F_5[0] for computing the Gröbner basis of a polynomial system rely heavily on a connection with linear algebra. Indeed, these algorithms reduce the Gröbner basis computation to a sequence of Gaussian eliminations on several submatrices of the so-called Macaulay matrix in some degree. Thus, we expect to improve the existing algorithms by
(i) developing dedicated linear algebra routines performing the Gaussian elimination steps: this is precisely the objective 2 described below;
(ii) generating smaller or simpler matrices to which we will apply Gaussian elimination.
We describe here our goals for the latter problem. First, we focus on algorithms for computing a Gröbner basis of *general polynomial systems*. Next, we present our goals on the development of dedicated algorithms for computing Gröbner bases of *structured polynomial systems* which arise in various applications.

Algorithms for general systems. Several degrees of freedom are available to the designer of a Gröbner basis algorithm to generate the matrices occurring during the computation. For instance, it would be desirable to obtain matrices which would be almost triangular or very sparse. Such a goal can be achieved by considering various interpretations of the F_5 algorithm with respect to different monomial orderings. To address this problem, the tight complexity results obtained for F_5 will be used to help in the design of such a general algorithm. To illustrate this point, consider the important problem of solving boolean polynomial systems; it might be interesting to preserve the sparsity of the original equations and, at the same time, using the fact that overdetermined systems are much easier to solve.

Algorithms dedicated to *structured* polynomial systems. A complementary approach is to exploit the structure of the input polynomials to design specific algorithms. Very often, problems coming from applications are not random but are highly structured. The specific nature of these systems may vary a lot: some polynomial systems can be sparse (when the number of terms in each equation is low), overdetermined (the number of the equations is larger than the number of variables), invariants by the action of some finite groups, multi-linear (each equation is linear w.r.t. to one block of variables) or more generally multihomogeneous. In each case, the ultimate goal is to identify large classes of problems whose theoretical/practical complexity drops and to propose in each case dedicated algorithms.

3.3. Solving Systems over the Reals and Applications.

Participants: Mohab Safey El Din, Elias Tsigaridas, Daniel Lazard, Thi Xuan Vu.

We shall develop algorithms for solving polynomial systems over complex/real numbers. Again, the goal is to extend significantly the range of reachable applications using algebraic techniques based on Gröbner bases and dedicated linear algebra routines. Targeted application domains are global optimization problems, stability of dynamical systems (e.g. arising in biology or in control theory) and theorem proving in computational geometry.

The following functionalities shall be requested by the end-users:
(i) deciding the emptiness of the real solution set of systems of polynomial equations and inequalities,
(ii) quantifier elimination over the reals or complex numbers,
(iii) answering connectivity queries for such real solution sets.
We will focus on these functionalities.

[0] J.-C. Faugère. *A new efficient algorithm for computing Gröbner bases without reduction to zero (F5).* In Proceedings of ISSAC '02, pages 75-83, New York, NY, USA, 2002. ACM.

We will develop algorithms based on the so-called critical point method to tackle systems of equations and inequalities (problem *(i)*) . These techniques are based on solving 0-dimensional polynomial systems encoding "critical points" which are defined by the vanishing of minors of Jacobian matrices (with polynomial entries). Since these systems are highly structured, the expected results of Objective 1 and 2 may allow us to obtain dramatic improvements in the computation of Gröbner bases of such polynomial systems. This will be the foundation of practically fast implementations (based on singly exponential algorithms) outperforming the current ones based on the historical Cylindrical Algebraic Decomposition (CAD) algorithm (whose complexity is doubly exponential in the number of variables). We will also develop algorithms and implementations that allow us to analyze, at least locally, the topology of solution sets in some specific situations. A long-term goal is obviously to obtain an analysis of the global topology.

3.4. Low level implementation and Dedicated Algebraic Computation and Linear Algebra.

Participants: Jean-Charles Faugère, Mohab Safey El Din, Elias Tsigaridas, Olive Chakraborty, Jocelyn Ryckeghem.

Here, the primary objective is to focus on *dedicated* algorithms and software for the linear algebra steps in Gröbner bases computations and for problems arising in Number Theory. As explained above, linear algebra is a key step in the process of computing efficiently Gröbner bases. It is then natural to develop specific linear algebra algorithms and implementations to further strengthen the existing software. Conversely, Gröbner bases computation is often a key ingredient in higher level algorithms from Algebraic Number Theory. In these cases, the algebraic problems are very particular and specific. Hence dedicated Gröbner bases algorithms and implementations would provide a better efficiency.

Dedicated linear algebra tools. The FGB library is an efficient one for Gröbner bases computations which can be used, for instance, via MAPLE. However, the library is sequential. A goal of the project is to extend its efficiency to new trend parallel architectures such as clusters of multi-processor systems in order to tackle a broader class of problems for several applications. Consequently, our first aim is to provide a durable, long term software solution, which will be the successor of the existing FGB library. To achieve this goal, we will first develop a high performance linear algebra package (under the LGPL license). This could be organized in the form of a collaborative project between the members of the team. The objective is not to develop a general library similar to the LINBOX[0] project but to propose a dedicated linear algebra package taking into account the specific properties of the matrices generated by the Gröbner bases algorithms. Indeed these matrices are sparse (the actual sparsity depends strongly on the application), almost block triangular and not necessarily of full rank. Moreover, most of the pivots are known at the beginning of the computation. In practice, such matrices are huge (more than 10^6 columns) but taking into account their shape may allow us to speed up the computations by one or several orders of magnitude. A variant of a Gaussian elimination algorithm together with a corresponding C implementation has been presented. The main peculiarity is the order in which the operations are performed. This will be the kernel of the new linear algebra library that will be developed.

Fast linear algebra packages would also benefit to the transformation of a Gröbner basis of a zero–dimensional ideal with respect to a given monomial ordering into a Gröbner basis with respect to another ordering. In the generic case at least, the change of ordering is equivalent to the computation of the minimal polynomial of a so-called multiplication matrix. By taking into account the sparsity of this matrix, the computation of the Gröbner basis can be done more efficiently using a variant of the Wiedemann algorithm. Hence, our goal is also to obtain a dedicated high performance library for transforming (i.e. change ordering) Gröbner bases.

[0]http://www.linalg.org/

Dedicated algebraic tools for Algebraic Number Theory. Recent results in Algebraic Number Theory tend to show that the computation of Gröbner basis is a key step toward the resolution of difficult problems in this domain [0]. Using existing resolution methods is simply not enough to solve relevant problems. The main algorithmic bottleneck to overcome is to adapt the Gröbner basis computation step to the specific problems. Typically, problems coming from Algebraic Number Theory usually have a lot of symmetries or the input systems are very structured. This is the case, in particular, for problems coming from the algorithmic theory of Abelian varieties over finite fields [0] where the objects are represented by polynomial system and are endowed with intrinsic group actions. The main goal here is to provide dedicated algebraic resolution algorithms and implementations for solving such problems. We do not restrict our focus on problems in positive characteristic. For instance, tower of algebraic fields can be viewed as triangular sets; more generally, related problems (e.g. effective Galois theory) which can be represented by polynomial systems will receive our attention. This is motivated by the fact that, for example, computing small integer solutions of Diophantine polynomial systems in connection with Coppersmith's method would also gain in efficiency by using a dedicated Gröbner bases computations step.

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

Participants: Jérémy Berthomieu, Jean-Charles Faugère, Ludovic Perret, Olive Chakraborty, Nagardjun Chinthamani, Solane El Hirch, Jocelyn Ryckeghem.

Here, we focus on solving polynomial systems over finite fields (i.e. the discrete case) and the corresponding applications (Cryptology, Error Correcting Codes, ...). Obviously this objective can be seen as an application of the results of the two previous objectives. However, we would like to emphasize that it is also the source of new theoretical problems and practical challenges. We propose to develop a systematic use of *structured systems* in *algebraic cryptanalysis*.
(i) So far, breaking a cryptosystem using algebraic techniques could be summarized as modeling the problem by algebraic equations and then computing a, usually, time consuming Gröbner basis. A new trend in this field is to require a theoretical complexity analysis. This is needed to explain the behavior of the attack but also to help the designers of new cryptosystems to propose actual secure parameters.
(ii) To assess the security of several cryptosystems in symmetric cryptography (block ciphers, hash functions, ...), a major difficulty is the size of the systems involved for this type of attack. More specifically, the bottleneck is the size of the linear algebra problems generated during a Gröbner basis computation.

We propose to develop a systematic use of *structured systems* in *algebraic cryptanalysis*.

The first objective is to build on the recent breakthrough in attacking McEliece's cryptosystem: it is the first structural weakness observed on one of the oldest public key cryptosystem. We plan to develop a well founded framework for assessing the security of public key cryptosystems based on coding theory from the algebraic cryptanalysis point of view. The answer to this issue is strongly related to the complexity of solving bihomogeneous systems (of bidegree $(1, d)$). We also plan to use the recently gained understanding on the complexity of structured systems in other areas of cryptography. For instance, the MinRank problem – which can be modeled as an overdetermined system of bilinear equations – is at the heart of the structural attack proposed by Kipnis and Shamir against HFE (one of the most well known multivariate public cryptosystem). The same family of structured systems arises in the algebraic cryptanalysis of the Discrete Logarithmic Problem (DLP) over curves (defined over some finite fields). More precisely, some bilinear systems appear in the polynomial modeling the points decomposition problem. Moreover, in this context, a natural group action can also be used during the resolution of the considered polynomial system.

[0] P. Gaudry, *Index calculus for abelian varieties of small dimension and the elliptic curve discrete logarithm problem*, Journal of Symbolic Computation 44,12 (2009) pp. 1690-1702

[0] e.g. point counting, discrete logarithm, isogeny.

Dedicated tools for linear algebra problems generated during the Gröbner basis computation will be used in algebraic cryptanalysis. The promise of considerable algebraic computing power beyond the capability of any standard computer algebra system will enable us to attack various cryptosystems or at least to propose accurate secure parameters for several important cryptosystems. Dedicated linear tools are thus needed to tackle these problems. From a theoretical perspective, we plan to further improve the theoretical complexity of the hybrid method and to investigate the problem of solving polynomial systems with noise, i.e. some equations of the system are incorrect. The hybrid method is a specific method for solving polynomial systems over finite fields. The idea is to mix exhaustive search and Gröbner basis computation to take advantage of the over-determinacy of the resulting systems.

Polynomial system with noise is currently emerging as a problem of major interest in cryptography. This problem is a key to further develop new applications of algebraic techniques; typically in side-channel and statistical attacks. We also emphasize that recently a connection has been established between several classical lattice problems (such as the Shortest Vector Problem), polynomial system solving and polynomial systems with noise. The main issue is that there is no sound algorithmic and theoretical framework for solving polynomial systems with noise. The development of such framework is a long-term objective.

4. Highlights of the Year

4.1. Highlights of the Year

CryptoNext Security, a spinoff of the PolSys team, was founded in June 2019 and has already been selected in the Future 40 group by STATION F of the most promising young startups.

5. New Software and Platforms

5.1. Epsilon

FUNCTIONAL DESCRIPTION: Epsilon is a library of functions implemented in Maple and Java for polynomial elimination and decomposition with (geometric) applications.

- Contact: Dongming Wang
- URL: http://wang.cc4cm.org/epsilon/index.html

5.2. FGb

KEYWORDS: Gröbner bases - Nonlinear system - Computer algebra

FUNCTIONAL DESCRIPTION: FGb is a powerful software for computing Gröbner bases. It includes the new generation of algorihms for computing Gröbner bases polynomial systems (mainly the F4, F5 and FGLM algorithms). It is implemented in C/C++ (approximately 250000 lines), standalone servers are available on demand. Since 2006, FGb is dynamically linked with Maple software (version 11 and higher) and is part of the official distribution of this software.

- Participant: Jean Charles Faugere
- Contact: Jean-Charles Faugère
- URL: http://www-polsys.lip6.fr/~jcf/FGb/index.html

5.3. FGb Light

FUNCTIONAL DESCRIPTION: Gröbner basis computation modulo p (p is a prime integer of 16 bits).

- Participant: Jean-Charles Faugère
- Contact: Jean-Charles Faugère
- URL: http://www-polsys.lip6.fr/~jcf/FGb/index.html

5.4. GBLA

FUNCTIONAL DESCRIPTION: GBLA is an open source C library for linear algebra specialized for eliminating matrices generated during Gröbner basis computations in algorithms like F4 or F5.

- Contact: Jean-Charles Faugère
- URL: http://www-polsys.lip6.fr/~jcf/GBLA/index.html

5.5. HFEBoost

FUNCTIONAL DESCRIPTION: Public-key cryptography system enabling an authentification of dematerialized data.

- Authors: Jean-Charles Faugère and Ludovic Perret
- Partner: UPMC
- Contact: Jean-Charles Faugère
- URL: http://www-polsys.lip6.fr/Links/hfeboost.html

5.6. RAGlib

Real Algebraic Geometry library

FUNCTIONAL DESCRIPTION: RAGLib is a powerful library, written in Maple, dedicated to solving over the reals polynomial systems. It is based on the FGb library for computing Grobner bases. It provides functionalities for deciding the emptiness and/or computing sample points to real solution sets of polynomial systems of equations and inequalities. This library provides implementations of the state-of-the-art algorithms with the currently best known asymptotic complexity for those problems.

- Contact: Mohab Safey El Din
- URL: http://www-polsys.lip6.fr/~safey/RAGLib/

5.7. RealCertify

KEYWORDS: Polynomial or analytical systems - Univariate polynomial - Real solving

FUNCTIONAL DESCRIPTION: The package RealCertify aims at providing a full suite of hybrid algorithms for computing certificates of non-negativity based on numerical software for solving linear matrix inequalities. The module univsos handles the univariate case and the module multivsos is designed for the multivariate case.

- Contact: Mohab Safey El Din
- URL: https://gricad-gitlab.univ-grenoble-alpes.fr/magronv/RealCertify

5.8. SLV

KEYWORDS: Univariate polynomial - Real solving

FUNCTIONAL DESCRIPTION: SLV is a software package in C that provides routines for isolating (and subsequently refine) the real roots of univariate polynomials with integer or rational coefficients based on subdivision algorithms and on the continued fraction expansion of real numbers. Special attention is given so that the package can handle polynomials that have degree several thousands and size of coefficients hundrends of Megabytes. Currently the code consists of approx. 5000 lines.

- Contact: Elias Tsigaridas
- URL: https://who.paris.inria.fr/Elias.Tsigaridas/soft.html

5.9. SPECTRA

Semidefinite Programming solved Exactly with Computational Tools of Real Algebra

KEYWORD: Linear Matrix Inequalities

FUNCTIONAL DESCRIPTION: SPECTRA is a Maple library devoted to solving exactly Semi-Definite Programs. It can handle rank constraints on the solution. It is based on the FGb library for computing Gröbner bases and provides either certified numerical approximations of the solutions or exact representations thereof.

- Contact: Mohab Safey El Din
- URL: http://homepages.laas.fr/henrion/software/spectra/

6. New Results

6.1. Fundamental algorithms and structured polynomial systems

The Berlekamp–Massey–Sakata algorithm and the Scalar-FGLM algorithm both compute the ideal of relations of a multidimensional linear recurrent sequence. Whenever quering a single sequence element is prohibitive, the bottleneck of these algorithms becomes the computation of all the needed sequence terms. As such, having adaptive variants of these algorithms, reducing the number of sequence queries, becomes mandatory. A native adaptive variant of the Scalar-FGLM algorithm was presented by its authors, the so-called Adaptive Scalar-FGLM algorithm. In [3], our first contribution is to make the Berlekamp–Massey–Sakata algorithm more efficient by making it adaptive to avoid some useless relation test-ings. This variant allows us to divide by four in dimension 2 and by seven in dimension 3 the number of basic operations performed on some sequence family. Then, we compare the two adaptive algorithms. We show that their behaviors differ in a way that it is not possible to tweak one of the algorithms in order to mimic exactly the behavior of the other. We detail precisely the differences and the similarities of both algorithms and conclude that in general the Adaptive Scalar-FGLM algorithm needs fewer queries and performs fewer basic operations than the Adaptive Berlekamp–Massey–Sakata algorithm. We also show that these variants are always more efficient than the original algorithms.

The problem of finding $m \times s$ matrices (with $m \geq s$) of rank r in a real affine subspace of dimension n has many applications in information and systems theory, where low rank is synonymous of structure and parsimony. In [8], we design computer algebra algorithms to solve this problem efficiently and exactly: the input are the rational coefficients of the matrices spanning the affine subspace as well as the expected maximum rank, and the output is a rational parametrization encoding a finite set of points that intersects each connected component of the low rank real algebraic set. The complexity of our algorithm is studied thoroughly. It is essentially polynomial in $n + m(s - r)$; it improves on the state-of-the-art in the field. Moreover, computer experiments show the practical efficiency of our approach.

Gröbner bases is one the most powerful tools in algorithmic non-linear algebra. Their computation is an intrinsically hard problem with a complexity at least single exponential in the number of variables. However, in most of the cases, the polynomial systems coming from applications have some kind of structure. For example , several problems in computer-aided design, robotics, vision, biology , kinematics, cryptography, and optimization involve sparse systems where the input polynomials have a few non-zero terms. In [16], our approach to exploit sparsity is to embed the systems in a semigroup algebra and to compute Gröbner bases over this algebra. Up to now, the algorithms that follow this approach benefit from the sparsity only in the case where all the polynomials have the same sparsity structure, that is the same Newton polytope. We introduce the first algorithm that overcomes this restriction. Under regularity assumptions, it performs no redundant computations. Further, we extend this algorithm to compute Gröbner basis in the standard algebra and solve sparse polynomials systems over the torus $(\mathbb{C}^{\star})^n$. The complexity of the algorithm depends on the Newton polytopes.

In [10], we consider the problem of approximating numerically the moments and the supports of measures which are invariant with respect to the dynamics of continuous- and discrete-time polynomial systems, under semialgebraic set constraints. First, we address the problem of approximating the density and hence the support of an invariant measure which is absolutely continuous with respect to the Lebesgue measure. Then, we focus on the approximation of the support of an invariant measure which is singular with respect to the Lebesgue measure. Each problem is handled through an appropriate reformulation into a linear optimization problem over measures, solved in practice with two hierarchies of finite-dimensional semidefinite moment-sum-of-square relaxations, also called Lasserre hierarchies. Under specific assumptions, the first Lasserre hierarchy allows to approximate the moments of an absolutely continuous invariant measure as close as desired and to extract a sequence of polynomials converging weakly to the density of this measure. The second Lasserre hierarchy allows to approximate as close as desired in the Hausdorff metric the support of a singular invariant measure with the level sets of the Christoffel polynomials associated to the moment matrices of this measure. We also present some application examples together with numerical results for several dynamical systems admitting either absolutely continuous or singular invariant measures.

6.2. Solving systems over the reals and applications

It is well-known that every non-negative univariate real polynomial can be written as the sum of two polynomial squares with real coefficients. When one allows a weighted sum of finitely many squares instead of a sum of two squares, then one can choose all coefficients in the representation to lie in the field generated by the coefficients of the polynomial. In particular, this allows an effective treatment of polynomials with rational coefficients. In [11], we describe, analyze and compare both from the theoretical and practical points of view, two algorithms computing such a weighted sums of squares decomposition for univariate polynomials with rational coefficients. The first algorithm, due to the third author relies on real root isolation, quadratic approximations of positive polynomials and square-free decomposition but its complexity was not analyzed. We provide bit complexity estimates, both on the runtime and the output size of this algorithm. They are exponential in the degree of the input univariate polynomial and linear in the maximum bitsize of its complexity. This analysis is obtained using quantifier elimination and root isolation bounds. The second algorithm, due to Chevillard, Harrison, Joldes and Lauter, relies on complex root isolation and square-free decomposition and has been introduced for certifying positiveness of poly-nomials in the context of computer arithmetics. Again, its complexity was not analyzed. We provide bit complexity estimates, both on the runtime and the output size of this algorithm, which are polynomial in the degree of the input polynomial and linear in the maximum bitsize of its complexity. This analysis is obtained using Vieta's formula and root isolation bounds. Finally, we report on our implementations of both algorithms and compare them in practice on several application benchmarks. While the second algorithm is, as expected from the complexity result, more efficient on most of examples, we exhibit families of non-negative polynomials for which the first algorithm is better.

[9] describes our freely distributed Maple library SPECTRA, for Semidefinite Programming solved Exactly with Computational Tools of Real Algebra. It solves linear matrix inequalities with symbolic computation in exact arithmetic and it is targeted to small-size, possibly degenerate problems for which symbolic infeasibility or feasibility certificates are required.

Let $S \subset \mathbb{R}^n$ be a compact basic semi-algebraic set defined as the real solution set of multivariate polynomial inequalities with rational coefficients. In [19], we design an algorithm which takes as input a polynomial system defining S and an integer $p \geq 0$ and returns the n-dimensional volume of S at absolute precision 2^{-p}. Our algorithm relies on the relationship between volumes of semi-algebraic sets and periods of rational integrals. It makes use of algorithms computing the Picard-Fuchs differential equation of appropriate periods, properties of critical points, and high-precision numerical integration of differential equations. The algorithm runs in essentially linear time with respect to p. This improves upon the previous exponential bounds obtained by Monte-Carlo or moment-based methods. Assuming a conjecture of Dimca, the arithmetic cost of the algebraic subroutines for computing Picard-Fuchs equations and critical points is singly exponential in n and polynomial in the maximum degree of the input.

Let $\mathbf{f} = (f_1, ..., f_s)$ be a sequence of polynomials in $\mathbb{Q}[X_1, ..., X_n]$ of maximal degree D and $V \subset \mathbb{C}^n$ be the algebraic set defined by $\mathbf{f}$ and r be its dimension. The real radical $\sqrt{\langle \mathbf{f} \rangle}$ associated to $\mathbf{f}$ is the largest ideal which defines the real trace of V. When V is smooth, we show in [13], that $\sqrt[re]{\langle \mathbf{f} \rangle}$, has a finite set of generators with degrees bounded by $\deg V$. Moreover, we present a probabilistic algorithm of complexity $(snD^n)^{O(1)}$ to compute the minimal primes of $\sqrt[re]{\langle \mathbf{f} \rangle}$. When V is not smooth, we give a probabilistic algorithm of complexity $s^{O(1)}(nD)^{O(nr2^r)}$ to compute rational parametrizations for all irreducible components of the real algebraic set $V \cap \mathbb{R}^n$.

Let $(g_1, ..., g_p)$ in $\mathbb{Q}[X_1, ..., X_n]$ and S be the basic closed semi-algebraic set defined by $g_1 \geq 0, ..., g_p \geq 0$. The S-radical of $\langle \mathbf{f} \rangle$, which is denoted by $\sqrt[S]{\langle \mathbf{f} \rangle}$, is the ideal associated to the Zariski closure of $V \cap S$. We give a probabilistic algorithm to compute rational parametrizations of all irreducible components of that Zariski closure, hence encoding $\sqrt[S]{\langle \mathbf{f} \rangle}$. Assuming now that D is the maximum of the degrees of the f_i's and the g_i's, this algorithm runs in time $2^p(s+p)^{O(1)}(nD)^{O(rn2^r)}$.

Experiments are performed to illustrate and show the efficiency of our approaches on computing real radicals.

In [14], we consider the second-order discontinuous differential equation $y'' + \eta \mathrm{sgn}(y) = \theta y + \alpha \sin(\beta t)$ where the parameters $\eta, \theta, \alpha, \beta$ are real. The main goal is to discuss the existence of periodic solutions. Under explicit conditions, the number of such solutions is given. Furthermore, for each of these periodic solutions, an explicit formula is provided.

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

6.3.1. Algebraic Cryptanalysis of a Quantum Money Scheme – The Noisy Case.

At STOC 2012, Aaronson and Christiano proposed a noisy and a noiseless version of the first public-key quantum money scheme endowed with a security proof. [5] addresses the so-called noisy hidden subspaces problem, on which the noisy version of their scheme is based. The first contribution of this work is a non-quantum cryptanalysis of the above-mentioned noisy quantum money scheme extended to prime fields $\mathbb{F}$, with $|\mathbb{F}| \neq 2$, that runs in randomised polynomial time. This finding is supported with experimental results showing that, in practice, the algorithm presented is efficient and succeeds with overwhelming probability. The second contribution is a non-quantum randomised polynomial-time cryptanalysis of the noisy quantum money scheme over $\mathbb{F}_2$ succeeding with a certain probability for values of the noise lying within a certain range. This result disproves a conjecture made by Aaronson and Christiano about the non-existence of an algorithm that solves the noisy hidden subspaces problem over $\mathbb{F}_2$ and succeeds with such probability.

6.3.2. On the Complexity of MQ in the Quantum Setting.

In August 2015 the cryptographic world was shaken by a sudden and surprising announcement by the US National Security Agency NSA concerning plans to transition to post-quantum algorithms. Since this announcement post-quantum cryptography has become a topic of primary interest for several standardization bodies. The transition from the currently deployed public-key algorithms to post-quantum algorithms has been found to be challenging in many aspects. In particular the problem of evaluating the quantum-bit security of such post-quantum cryptosystems remains vastly open. Of course this question is of primarily concern in the process of standardizing the post-quantum cryptosystems. In [21] we consider the quantum security of the problem of solving a system of *m Boolean multivariate quadratic equations in n variables* (MQb); a central problem in post-quantum cryptography. When $n = m$, under a natural algebraic assumption, we present a Las-Vegas quantum algorithm solving MQb that requires the evaluation of, on average, $O(2^{0.462n})$ quantum gates. To our knowledge this is the fastest algorithm for solving MQb.

6.3.3. MQsoft.

In 2017, NIST shook the cryptographic world by starting a process for standardizing post-quantum cryptography. Sixty-four submissions have been considered for the first round of the on-going NIST Post-Quantum

Cryptography (PQC) process. Multivariate cryptography is a classical post-quantum candidate that turns to be the most represented in the signature category. At this stage of the process, it is of primary importance to investigate efficient implementations of the candidates. [17] presents `MQsoft`, an efficient library which permits to implement `HFE`-based multivariate schemes submitted to the NIST PQC process such as *GeMSS*, `Gui` and *DualModeMS*. The library is implemented in `C` targeting Intel 64-bit processors and using `avx2` set instructions. We present performance results for our library and its application to *GeMSS*, `Gui` and *DualModeMS*. In particular, we optimize several crucial parts for these schemes. These include root finding for `HFE` polynomials and evaluation of multivariate quadratic systems in $\mathbb{F}_2$. We propose a new method which accelerates root finding for specific `HFE` polynomials by a factor of two. For *GeMSS* and `Gui`, we obtain a speed-up of a factor between 2 and 19 for the keypair generation, between 1.2 and 2.5 for the signature generation, and between 1.6 and 2 for the verifying process. We have also improved the arithmetic in F_{2^n} by a factor of 4 compared to the `NTL` library. Moreover, a large part of our implementation is protected against timing attacks.

7. Bilateral Contracts and Grants with Industry

7.1. Technological Transfer

The group have had a continuous commitment into industrial transfer as well as a strong involvement into standardization bodies.

Industrial transfer. This activity is related to our long-standing activity in post-quantum cryptography. The transfer started at the beginning of the current evaluation period and culminated this year with the creation of a new spin-off, called CRYPTONEXT SECURITY[0], from Inria Paris and Sorbonne Université. The goal of CRYPTONEXT SECURITY is to propose security products that are resistant against the quantum computers. Its business model is based on B2B and targeted customers are Fortune 500 companies.

The activity has been partially founded and supervised by SATT-LUTECH who is specialized in the processing and transfer of technologies from research laboratories of its shareholders: Inria, CNRS, University of Technology of Compiègne, National Museum of Natural History, Institute Curie, University Panthéon-Assas, Paris Sorbonne University and National School of Industrial Creation). Typically, SATT-LUTECH has funded the post-quantum experiment described in the dedicated Section. The impact of such experiment can be partially measured by the press released covering out the test [0] (La Recherche, l'Usine Nouvelle, L'Informaticien).

As a preliminary step for launching the spin-off, two members of the team (J.-C. Faugère and L. Perret) followed two entrepreneurship programs for creating innovative companies : HEC Challenge plus [0] (1 week of courses by month, 9 months) and Deep Tech Founders [0] (3 months, 2 sessions by weeks). This was a necessary, but significant effort, before launching CRYPTONEXT in which the two members will work full-time from now on.

Post-quantum standardization. Besides the `NIST` PQC standardization process, we are involved in the on-going world effort for standardizing post-quantum cryptography. More precisely, the European Telecommunications Standards Institute (`ETSI`) has a strong standardization activity on post-quantum cryptography. `ETSI` is a EU standardization body with a worldwide scope. We are an active member of `ETSI` regarding post-quantum cryptography. In particular, we are the rapporteur of a technical document on post-quantum cryptosystems. The goal of our involvement is to bring our scientific expertise to define trustworthy post-quantum public-key standards; that are going to be the basis of our digital economy within $10/15$ years.

[0] https://cryptonext-security.com/
[0] https://www-polsys.lip6.fr/Links/index.html
[0] http://entrepreneurship-center.hec.edu/learn-program/hec-challenge-plus/
[0] https://deeptechfounders.com/

We are also involved in the Cloud Security Alliance (CSA). This is a large non-profit organization (80.000 member worldwide) whose main goal is to promote the best practices with the secure usage of cloud computing. CSA has a group dedicated to quantum-safe security. The group is ideation catalyzer for promoting the transition of companies to a quantum-safe security. In particular, the group has a significant educational activity in order to increase awareness regarding the quantum risk and the techniques to mitigate this risk. We are co-chairing this group and participated to several white papers. The group is probably now the main channel for promoting quantum-safe security. Standardization is a long-term effort.

8. Partnerships and Cooperations

8.1. Regional Initiatives

- **Grant CAMiSAdo** (funded by PGMO).

  COMPUTER ALGEBRA METHODS FOR SEMI-ALGEBRAIC PROGRAMMING

  `Participants:` J. Berthomieu [contact], M. Safey El Din.

  Semi-Algebraic Programming is the art of optimizing some quantity subject to semi-algebraic constraints. The very basic and natural instance of semi-algebraic programming is the problem of optimizating a polynomial function subject to polynomial inequalities and is known as the polynomial optimization problem (POP). More general instances of semi-algebraic programming are as follows: given a system of polynomial equations/inequalities depending on parameters, what are the parameters' values which maximize the dimension of the semi-algebraic set defined by the instantiated system? And when the number of solutions is finite, what is this maximum number of solutions? Hence Semi-Algebraic Programming encompasses a wide range of computational issues related to semi-algebraic sets. It finds applications in many engineering sciences. Let us mention the few ones that we target in CAMiSAdo: Path-planning optimization in robotics, Mobility properties of manipulators in mechanism design, Stability analysis for sensor-based controllers.

8.2. National Initiatives

- **ANR SESAME (Singularités Et Stabilité des AsservisseMEnts référencés capteurs)**

  `Duration:` 2018–2022

  `Participants:` J.-C. Faugère, M. Safey El Din [contact].

  The demand for flexible, adaptable robots capable of interacting with their environment (e.g. navigation, handling, cooperation) is growing. This is why the sensor-based controllers, which make it possible to include external sensory feedback in robot control, have been widely developed in recent years, both for industrial, medical, air, space and marine robotics and in the context of autonomous vehicles (ground mobile robotics).

  The first research on sensor-based control techniques took place at the end of the 1980s, with the use of proximal and force and vision sensors, and much work has been done to improve the performance of this type of controllers, in particular by modelling various sensor primitives.

  Despite the fact that, empirically, sensor-based controllers have shown that they have interesting performances, these performances are by no means guaranteed, which is a major obstacle to the widespread use of their large-scale use. This is related to the fact that, despite three decades of research on the subject, two broad classes of problems have been little explored:

  – The study of the singularities of sensor-based controllers
  – The study of their stability.

  The objectives of the project SESAME are take advantage on recent mathematical advances in order to:

  - study singularities and stability of certain classes of sensor-based controllers
  - synthesize globally asymptotically stable sensor-based controllers, whose performance (i.e. convergence properties towards the desired configuration, abseance of local singularities and minima) are guaranteed in all object/sensor related configurations.

  Many of the computational tools SESAME relies on involve computer algebra and polynomial system solving.

- **ANR Jeunes Chercheurs GALOP (Games through the lens of ALgebra and OPptimization)**

  `Duration:` 2018–2022

  `Participants:` E. Tsigaridas [contact], F. Johansson, H. Gimbert, J.-C. Faugère, M. Safey El Din.

  GALOP [0] is a Young Researchers (JCJC) project with the purpose of extending the limits of the state-of-the-art algebraic tools in computer science, especially in stochastic games. It brings original and innovative algebraic tools, based on symbolic-numeric computing, that exploit the geometry and the structure and complement the state-of-the-art. We support our theoretical tools with a highly efficient open-source software for solving polynomials. Using our algebraic tools we study the geometry of the central curve of (semi-definite) optimization problems. The algebraic tools and our results from the geometry of optimization pave the way to introduce algorithms and precise bounds for stochastic games.

- **ANR ECARP (Efficient Certified Algorithms for Robot Motion Planning)**

  `Duration:` 2020–2024

  `Participants:` J. Berthomieu, J.-C. Faugère, M. Safey El Din [contact].

  ECARP is an international project, jointly funded by ANR and FWF (the funding agency of Austria). It targets the design and implementation of high-performance computer algebra algorithms for semi-algebraic sets in order to answer connectivity queries over those sets. This is applied to motion planning issues in robotics, e.g. for analyzing kinematic singularities ; parallel and serial manipulators will be investigated. The consortium gathers experts in geometry and robotics from J. Kepler Univ. (Austria) and LS2N (Nantes).

- **ANR DRN (DeRerumNatura)**

  `Duration:` 2020–2024

  `Participants:` J. Berthomieu [contact], M. Safey El Din.

  Classifying objects, determining their nature is more often than not the endgame of a theory. Yet, even the most established theory can be impractible on a concrete instance, either because of a lack of efficiency or because of a computational wall. In both cases, an algorithm is lacking: we need to systematize efficiently and automatically. This is what DRN proposes to do to solve classification problems related to numbers, analytic functions and combinactorics generating series. The consortium gathers experts in computer algebra (Inria Saclay, Limoges, Lyon, POLSYS), Combinactorics (Inria Saclay, Lyon) and Galois Theory (Toulouse, Strasbourg, Versailles).

8.2.1. Programme d'investissements d'avenir (PIA)

- **PIA grant RISQ: Regroupement of the Security Industry for Quantum-Safe security (2017-2020).** The goal of the RISQ project is to prepare the security industry to the upcoming shift of classical cryptography to quantum-safe cryptography. (J.-C. Faugère [contact], and L. Perret).

  The RISQ [0] project is certainly the biggest industrial project ever organized in quantum-safe cryptography. RISQ is one of few projects accepted in the call Grands Défis du Numérique which is managed by BPI France, and will be funded thanks to the so-called Plan d'Investissements d'Avenir.

[0] https://project.inria.fr/galop/
[0] http://risq.fr/

The RISQ project is a natural continuation of POLSYS commitment to the industrial transfert of quantum-safe cryptography. RISQ is a large scale version of the HFEBoost project; which demonstrated the potential of quantum-safe cryptography.

POLSYS actively participated to shape the RISQ project. POLSYS is now a member of the strategic board of RISQ, and is leading the task of designing and analyzing quantum-safe algorithms. In particular, a first milestone of this task was to prepare submissions to NIST's quantum-safe standardisation process.

8.3. European Initiatives

- Innovative Training Network POEMA (Polynomial Optimization, Efficiency through Moments and Algebra) - ITN Marie Curie H2020 program.

  `Duration:` 2019–2023

  `Participants:` J. Berthomieu, J.-C. Faugère, M. Safey El Din [contact].

  POEMA is part of the Marie Sklodowska-Curie Actions — Innovative Training Networks (ITN) funding scheme.

  POEMA aims to train scientists at the interplay of algebra, geometry and computer science for polynomial optimization problems and to foster scientific and technological advances, stimulating interdisciplinary and intersectoriality knowledge exchange between algebraists, geometers, computer scientists and industrial actors facing real-life optimization problems.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. General Chair, Scientific Chair

D. Wang was General co-chair of the 44th International Symposium on Symbolic and Algebraic Computation (ISSAC 2019) , Beijing, China, July 15-18, 2019).

9.1.1.2. Member of the Organizing Committees

J. Berthomieu was Poster chair of the 44th International Symposium on Symbolic and Algebraic Computation (ISSAC 2019), Beijing, China, July 15-18, 2019).

9.1.2. Scientific Events: Selection

9.1.2.1. Member of the Conference Program Committees

E. Prouff was member of the program committees of the following conferences:

- Smart Card Research and Advanced Application Conference (CARDIS) 2019
- IACR Transactions on Cryptographic Hardware and Embedded Systems (TCHES) 2019

D. Wang was member of the program committee of the International Symposium on Wen-Tsun Wu's Academic Thought and Mathematics Mechanization (Wu 2019) (Beijing, China, May 12-17, 2019). He was member of the scientific committee of the 4th International Conference on Numerical and Symbolic Computation (SYMCOMP 2019) (Porto, Portugal, April 11-12, 2019).

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

E. Prouff is member of the editorial board of Journal of Cryptographic Engineering.

M. Safey El Din is member of the editorial board of Journal of Symbolic Computation.

D. Wang is Editor-In-Chief of Mathematics in Computer Science (published by Birkhäuser- Springer, Basel) and member of the Advisory Board for the journal • SCIENCE CHINA Information Sciences (published by Science China Press, Beijing and Springer, Berlin).

D. Wang is member of the Editorial Boards for the Journal of Symbolic Computation (published by Academic Press/Elsevier, London), SN Computer Science (published by Springer Nature, Switzerland) , Texts and Monographs in Symbolic Computation (published by Springer, Wien New York).

D. Wang is member of the International Advisory Board for Communications of JSSAC (Japan Society for Symbolic and Algebraic Computation) (published by JSSAC).

9.1.4. Invited Talks

E. Prouff was invited speaker at Spring School on Hardware Security (UK), the NIST Threshold Cryptography Workshop (USA), the Summer School on Real World Crypto and Privacy and the conference on Boolean Functions and Applications (BFA).

M. Safey El Din was invited speaker at the Arctic Applied Algebra conference (Norway), the workshop on Geometry of Real Polynomials, Convexity and Optimization at BIRS Banff International Research Station (Canada), the and the Conference in memory of Wen-tsun Wu's centennial birthday (China).

9.1.5. Scientific Expertise

M. Safey El Din was expert for the Czech Republic funding agency.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Jérémy Berthomieu had the following teaching activities:

Master : Polynomial System Solving, 13 hours, M2, Sorbonne Université & Polytech' Sorbonne, France.

Master : Computation Modeling, 42 hours, M1, Sorbonne Université, France.

Master : In charge of Basics of Algebraic Algorithms, 50 hours, M1, Sorbonne Université & Polytech' Sorbonne, France.

Master : Projects supervision, 7 hours, M1, Sorbonne Université, France.

Licence : Introduction to Cryptology, 15,75 hours, L3, Sorbonne Université , France.

Licence : Introduction to Numerical Algorithmic, 4,75 hours, L3, Sorbonne Université , France.

Licence : Introduction to Shell, 38,75 hours, L2, Sorbonne Université , France.

Licence : Projects supervision, 10 hours, L2, Sorbonne Université, France.

Jean-Charles Faugère had the following teaching activities:

Master : Polynomial Systems solving, 12 hours, M2, MPRI, France.

Ludovic Perret had the following teaching activities:

Master : Polynomial Systems solving, 12 hours, M2, MPRI, France.

Mohab Safey El Din has the following teaching activities:

Master : Polynomial System Solving, 40 hours, M1, Sorbonne Université & Polytech' Sorbonne, France.

Master : In charge of the curriculum on Security, Reliability of Performance in Computing, 30 hours, M1, Sorbonne Université , France.

Master : Projects management, 20 hours, M1, Sorbonne Université, France.

Licence : Projects supervision, 10 hours, L2, Sorbonne Université, France.

9.2.2. Supervision

PhD : Matías Bender, Algorithms for sparse polynomial systems: Grbner basis and resultants, June 2019, Jean-Charles Faugère and Elias Tsigaridas.

PhD in progress : Thi Xuan Vu, Faster algorithms for structured polynomial systems, started in Oct. 2017, Jean-Charles Faugère and Mohab Safey El Din.

PhD in progress : Phuoc Le, Real root classification and polar varieties, started in Oct. 2018, Jean-Charles Faugère and Mohab Safey El Din.

PhD in progress : Olive Chakraborty, Conception and cryptanalysis in secure quantic cryptography, started in May 2017, Jean-Charles Faugère and Ludovic Perret.

9.2.2.1. Juries

E. Prouff was member of the PhD committes of:

- Nicolas Belleville, Université Grenoble Alpes, 21-11-2019, "Compilation pour l'application de contre-mesures contre les attaques par canal auxiliaire"
- Joey Green, Université de Bristol, 20-10-2019, "A Study of Inference-Based Attacks with Neural Networks"
- Aberrahman DAIF, Université Paris 8, 03-05-2019, "Les codes correcteurs au service de la cryptographie symétrique et asymétrique"

J.C. Faugère was member of the PhD committes of:

- Robin Larrieu, École Polytechnique, dec. 2019, "Arithmetique rapide pour des corps finis"

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] M. BENDER.*Algorithms for sparse polynomial systems : Gröbner basis and resultants*, Sorbonne Université, 6 2019, http://page.math.tu-berlin.de/ mbender/etc/thesis.pdf

Articles in International Peer-Reviewed Journal

[2] E. BARTZOS, I. Z. EMIRIS, J. LEGERSKÝ, E. TSIGARIDAS.*On the maximal number of real embeddings of minimally rigid graphs in* $\mathbb{R}^2$, $\mathbb{R}^3$ *and* S^2, in "Journal of Symbolic Computation", 2019, https://arxiv.org/abs/1811.12800 , forthcoming [*DOI :* 10.1016/J.JSC.2019.10.015], https://hal.archives-ouvertes.fr/hal-02271782

[3] J. BERTHOMIEU, J.-C. FAUGÈRE.*In-depth comparison of the Berlekamp–Massey–Sakata and the Scalar-FGLM algorithms: the adaptive variants*, in "Journal of Symbolic Computation", 2019, https://arxiv.org/abs/1806.00978 , forthcoming [*DOI :* 10.1016/J.JSC.2019.09.001], https://hal.inria.fr/hal-01805478

[4] L. BUSÉ, A. MANTZAFLARIS, E. TSIGARIDAS.*Matrix formulae for Resultants and Discriminants of Bivariate Tensor-product Polynomials*, in "Journal of Symbolic Computation", June 2020, vol. 98, p. 65-83 [*DOI :* 10.1016/J.JSC.2019.07.007], https://hal.inria.fr/hal-01654263

[5] M. CONDE PENA, R. DURÁN DÍAZ, J.-C. FAUGÈRE, L. HERNÁNDEZ ENCINAS, L. PERRET.*Non-quantum cryptanalysis of the noisy version of Aaronson–Christiano's quantum money scheme*, in "IET Information Security", July 2019, vol. 13, n^{o} 4, p. 362-366 [*DOI :* 10.1049/IET-IFS.2018.5307], https://hal.inria.fr/hal-02395072

[6] I. Z. Emiris, A. Mantzaflaris, E. Tsigaridas.*Multilinear Polynomial Systems: Root Isolation and Bit Complexity*, in "Journal of Symbolic Computation", 2019, Special Issue of the Journal of Symbolic Computation on Milestones in Computer Algebra (MICA 2016), forthcoming, https://hal.inria.fr/hal-02099556

[7] I. Z. Emiris, B. Mourrain, E. Tsigaridas.*Separation bounds for polynomial systems*, in "Journal of Symbolic Computation", 2019 [*DOI :* 10.1016/j.jsc.2019.07.001], https://hal.inria.fr/hal-01105276

[8] D. Henrion, S. Naldi, M. Safey El Din.*Real root finding for low rank linear matrices*, in "Applicable Algebra in Engineering, Communication and Computing", 2019, https://arxiv.org/abs/1506.05897 [*DOI :* 10.1007/s00200-019-00396-w], https://hal.archives-ouvertes.fr/hal-01159210

[9] D. Henrion, S. Naldi, M. Safey El Din.*SPECTRA – a Maple library for solving linear matrix inequalities in exact arithmetic*, in "Optimization Methods and Software", January 2019, vol. 34, n^{o} 1, p. 62-78, https://arxiv.org/abs/1611.01947 - Significantly extended version [*DOI :* 10.1080/10556788.2017.1341505], https://hal.laas.fr/hal-01393022

[10] V. Magron, M. Forets, D. Henrion.*Semidefinite Approximations of Invariant Measures for Polynomial Systems*, in "Discrete and Continuous Dynamical Systems - Series B", December 2019, vol. 24, n^{o} 12, p. 6745-6770, https://arxiv.org/abs/1807.00754 - 28 pages, 14 figures [*DOI :* 10.3934/dcdsb.2019165], https://hal.archives-ouvertes.fr/hal-01828443

[11] V. Magron, M. Safey El Din, M. Schweighofer.*Algorithms for Weighted Sums of Squares Decomposition of Non-negative Univariate Polynomials*, in "Journal of Symbolic Computation", 2019, vol. 93, p. 200-220 [*DOI :* 10.1016/j.jsc.2018.06.005], https://hal.archives-ouvertes.fr/hal-01538729

[12] L. Masure, C. Dumas, E. Prouff.*A Comprehensive Study of Deep Learning for Side-Channel Analysis*, in "IACR Transactions on Cryptographic Hardware and Embedded Systems", November 2019, vol. 2020, n^{o} 1, p. 348-375 [*DOI :* 10.13154/tches.v2020.i1.348-375], https://hal.archives-ouvertes.fr/hal-02425261

[13] M. Safey El Din, Z.-H. Yang, L. Zhi.*Computing real radicals and S-radicals of polynomial systems*, in "Journal of Symbolic Computation", November 2019 [*DOI :* 10.1016/j.jsc.2019.10.018], https://hal.sorbonne-universite.fr/hal-02388168

[14] C. da Silva, A. Jacquemard, M. Teixeira.*Periodic Solutions of a Class of Non-autonomous Discontinuous Second-Order Differential Equations*, in "Journal of Dynamical and Control Systems", 2019, p. 1-28, forthcoming [*DOI :* 10.1007/s10883-018-9426-7], https://hal-univ-bourgogne.archives-ouvertes.fr/hal-02094522

International Conferences with Proceedings

[15] A. Bauer, H. Gilbert, G. Renault, M. Rossi.*Assessment of the Key-Reuse Resilience of NewHope*, in "CT-RSA 2019 - The Cryptographers' Track at the RSA Conference", San Francisco, United States, M. Matsui (editor), Lecture Notes in Computer Science, Springer, February 2019, vol. 11405, p. 272-292 [*DOI :* 10.1007/978-3-030-12612-4_14], https://hal.archives-ouvertes.fr/hal-02139910

[16] M. R. Bender, J.-C. Faugère, E. P. Tsigaridas.*Gröbner Basis over Semigroup Algebras: Algorithms and Applications for Sparse Polynomial Systems*, in "ISSAC 2019 - 44th International Symposium on Symbolic and Algebraic Computation", Beijing, China, ACM, July 2019, p. 42-49, https://arxiv.org/abs/1902.00208 [*DOI :* 10.1145/3326229.3326248], https://hal.inria.fr/hal-02002689

[17] J.-C. FAUGÈRE, L. PERRET, J. RYCKEGHEM.*Software Toolkit for HFE-based Multivariate Schemes*, in "CHES 2019 : International Conference on Cryptographic Hardware and Embedded Systems", Atlanta, United States, IACR Transactions on Cryptographic Hardware and Embedded Systems, May 2019, vol. 2019, n^{o} 3, p. 257-304 [*DOI :* 10.13154/TCHES.V2019.I3.257-304], https://hal.sorbonne-universite.fr/hal-02389747

[18] E. KUSHILEVITZ, R. OSTROVSKY, E. PROUFF, A. ROSÉN, A. THILLARD, D. VERGNAUD.*Lower and Upper Bounds on the Randomness Complexity of Private Computations of AND*, in "TCC 2019 - 17th International Conference on Theory of Cryptography", Nuremberg, Germany, D. HOFHEINZ, A. ROSEN (editors), Lecture Notes in Computer Science, Springer, November 2019, vol. 11892, p. 386-406 [*DOI :* 10.1007/978-3-030-36033-7_15], https://hal.archives-ouvertes.fr/hal-02395052

[19] P. LAIREZ, M. MEZZAROBBA, M. SAFEY EL DIN.*Computing the volume of compact semi-algebraic sets*, in "ISSAC 2019 - International Symposium on Symbolic and Algebraic Computation", Beijing, China, ACM, July 2019, https://arxiv.org/abs/1904.11705 , https://hal.archives-ouvertes.fr/hal-02110556

Other Publications

[20] M. R. BENDER, J.-C. FAUGÈRE, L. PERRET, E. TSIGARIDAS.*A nearly optimal algorithm to decompose binary forms*, April 2019, https://arxiv.org/abs/1810.12588 - Accepted to JSC, https://hal.archives-ouvertes.fr/hal-01907777

[21] J.-C. FAUGÈRE, K. HORAN, D. KAHROBAEI, M. KAPLAN, E. KASHEFI, L. PERRET.*Fast Quantum Algorithm for Solving Multivariate Quadratic Equations*, August 2019, https://arxiv.org/abs/1712.07211 - working paper or preprint, https://hal.inria.fr/hal-01995374

[22] T.-H. VU.*Solution maps of polynomial variational inequalities*, February 2020, working paper or preprint, https://hal.sorbonne-universite.fr/hal-02468977

Project-Team PROSECCO

Programming securely with cryptography

RESEARCH CENTER
Paris

THEME
Security and Confidentiality

Table of contents

Project-Team PROSECCO

Creation of the Team: 2012 January 01, updated into Project-Team: 2012 July 01

Keywords:

Computer Science and Digital Science:

A1.1. - Architectures
A1.1.8. - Security of architectures
A1.2. - Networks
A1.2.8. - Network security
A1.3. - Distributed Systems
A2. - Software
A2.1. - Programming Languages
A2.1.1. - Semantics of programming languages
A2.1.4. - Functional programming
A2.1.7. - Distributed programming
A2.1.11. - Proof languages
A2.2. - Compilation
A2.2.1. - Static analysis
A2.2.5. - Run-time systems
A2.4. - Formal method for verification, reliability, certification
A2.4.2. - Model-checking
A2.4.3. - Proofs
A2.5. - Software engineering
A4. - Security and privacy
A4.3. - Cryptography
A4.3.3. - Cryptographic protocols
A4.5. - Formal methods for security
A4.6. - Authentication
A4.8. - Privacy-enhancing technologies

Other Research Topics and Application Domains:

B6. - IT and telecom
B6.1. - Software industry
B6.1.1. - Software engineering
B6.3. - Network functions
B6.3.1. - Web
B6.3.2. - Network protocols
B6.4. - Internet of things
B9. - Society and Knowledge
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

Karthikeyan Bhargavan [Team leader, Inria, Senior Researcher, HDR]
Bruno Blanchet [Inria, Senior Researcher, HDR]
Harry Halpin [Inria, Advanced Research Position, until Aug 2019]
Catalin Hritcu [Inria, Researcher, HDR]
Prasad Naldurg [Inria, Advanced Research Position]
Exequiel Rivas Gadda [Inria, Starting Research Position, from Mar 2019]
Eric Tanter [Inria, Advanced Research Position]

External Collaborators
Victor Dumitrescu [Nomadic Labs]
Guido Martinez [CIFASIS-CONICET Rosario]
Jonathan Protzenko [Microsoft Research]
Eric Tanter [University of Chile, from Mar 2019 until Aug 2019]
Jean-Karim Zinzindohoué [Ministère de l'Intérieur, until Jun 2019]

Technical Staff
Cezar Constantin Andrici [Inria, Engineer, from Aug 2019 until Nov 2019]
Iness Ben Guirat [Inria, Engineer, until Apr 2019]
Adrien Durier [Inria, Engineer, from Sep 2019]
Florian Groult [Inria, Engineer]
Theo Laurent [Inria, Engineer]
Ramkumar Ramachandra [Inria, Engineer, from Aug 2019]
Antoine Van Muylder [Inria, Engineer, from Nov 2019]

PhD Students
Carmine Abate [Inria, PhD Student]
Benjamin Beurdouche [Inria, PhD Student, granted by FP7 ERC CIRCUS project]
Natalia Kulatova [Inria, PhD Student]
Benjamin Lipp [Inria, PhD Student]
Kenji Maillard [Ecole Normale Supérieure Paris, PhD Student, until Nov 2019]
Denis Merigoux [Inria, PhD Student]
Marina Polubelova [Inria, PhD Student]
Jérémy Thibault [Inria, PhD Student]

Post-Doctoral Fellow
Roberto Blanco Martinez [Inria, Post-Doctoral Fellow]

Administrative Assistants
Christelle Guiziou [Inria, Administrative Assistant]
Mathieu Mourey [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Programming securely with cryptography

In recent years, an increasing amount of sensitive data is being generated, manipulated, and accessed online, from bank accounts to health records. Both national security and individual privacy have come to rely on the security of web-based software applications. But even a single design flaw or implementation bug in an application may be exploited by a malicious criminal to steal, modify, or forge the private records of innocent users. Such *attacks* are becoming increasingly common and now affect millions of users every year.

The risks of deploying insecure software are too great to tolerate anything less than mathematical proof, but applications have become too large for security experts to examine by hand, and automated verification tools do not scale. Today, there is not a single widely-used web application for which we can give a proof of security, even against a small class of attacks. In fact, design and implementation flaws are still found in widely-distributed and thoroughly-vetted security libraries designed and implemented by experts.

Software security is in crisis. A focused research effort is needed if security programming and analysis techniques are to keep up with the rapid development and deployment of security-critical distributed applications based on new cryptographic protocols and secure hardware devices. The goal of our team PROSECCO is to draw upon our expertise in cryptographic protocols and program verification to make decisive contributions in this direction.

Our vision is that, over its lifetime, PROSECCO will contribute to making the use of formal techniques when programming with cryptography as natural as the use of a software debugger. To this end, our long-term goals are to design and implement programming language abstractions, cryptographic models, verification tools, and verified security libraries that developers can use to deploy provably secure distributed applications. Our target applications include cryptographic protocol implementations, hardware-based security APIs, smartphone- and browser-based web applications, and cloud-based web services. In particular, we aim to verify the full application: both the cryptographic core and the high-level application code. We aim to verify implementations, not just models. We aim to account for computational cryptography, not just its symbolic abstraction.

We identify five key focus areas for our research in the short- to medium term.

2.1.1. New programming languages for verified software

Building realistic verified applications requires new programming languages that enable the systematic development of efficient software hand-in-hand with their proofs of correctness. Our current focus is on designing and implementing the programming language F*, in collaboration with Microsoft Research. F* (pronounced F star) is a general-purpose functional programming language with effects aimed at program verification. Its type system includes polymorphism, dependent types, monadic effects, refinement types, and a weakest precondition calculus. Together, these features allow expressing precise and compact specifications for programs, including functional correctness and security properties. The F* type-checker aims to prove that programs meet their specifications using a combination of SMT solving and interactive proofs. Programs written in F* can be translated to efficient OCaml, F#, or C for execution. The main ongoing use case of F* is building a verified, drop-in replacement for the whole HTTPS stack in Project Everest (a larger collaboration with Microsoft Research). This includes verified implementations of TLS 1.2 and 1.3 and of the underlying cryptographic primitives.

2.1.2. Symbolic verification of cryptographic applications

We aim to develop our own security verification tools for models and implementations of cryptographic protocols and security APIs using symbolic cryptography. Our starting point is the tools we have previously developed: the specialized cryptographic prover ProVerif, the reverse engineering and formal test tool Tookan, and the F* verification system. These tools are already used to verify industrial-strength cryptographic protocol implementations and commercial cryptographic hardware. We plan to extend and combine these approaches to capture more sophisticated attacks on applications consisiting of protocols, software, and hardware, as well as to prove symbolic security properties for such composite systems.

2.1.3. Computational verification of cryptographic applications

We aim to develop our own cryptographic application verification tools that use the computational model of cryptography. The tools include the computational prover CryptoVerif, and the F* verification system. Working together, we plan to extend these tools to analyze, for the first time, cryptographic protocols, security APIs, and their implementations under fully precise cryptographic assumptions. We also plan to pursue links between symbolic and computational verification, such as computational soundness results that enable computational proofs by symbolic techniques.

2.1.4. Efficient formally secure compilers for tagged architectures

We aim to leverage emerging hardware capabilities for fine-grained protection to build the first, efficient secure compilation chains for realistic low-level programming languages (the C language, and Low* a safe subset of C embedded in F* for verification). These compilation chains will provide a secure semantics for all programs and will ensure that high-level abstractions cannot be violated even when interacting with untrusted low-level code. To achieve this level of security without sacrificing efficiency, our secure compilation chains target a tagged architecture, which associates a metadata tag to each word and efficiently propagates and checks tags according to software-defined rules.

2.1.5. Building provably secure web applications

We aim to develop analysis tools and verified libraries to help programmers build provably secure web applications. The tools will include static and dynamic verification tools for client- and server-side JavaScript web applications, their verified deployment within HTML5 websites and browser extensions, as well as type-preserving compilers from high-level applications written in F* to JavaScript. In addition, we plan to model new security APIs in browsers and smartphones and develop the first formal semantics for various HTML5 web standards. We plan to combine these tools and models to analyze the security of multi-party web applications, consisting of clients on browsers and smartphones, and servers in the cloud.

3. Research Program

3.1. Symbolic verification of cryptographic applications

Despite decades of experience, designing and implementing cryptographic applications remains dangerously error-prone, even for experts. This is partly because cryptographic security is an inherently hard problem, and partly because automated verification tools require carefully-crafted inputs and are not widely applicable. To take just the example of TLS, a widely-deployed and well-studied cryptographic protocol designed, implemented, and verified by security experts, the lack of a formal proof about all its details has regularly led to the discovery of major attacks (including several in PROSECCO) on both the protocol and its implementations, after many years of unsuspecting use.

As a result, the automated verification for cryptographic applications is an active area of research, with a wide variety of tools being employed for verifying different kinds of applications.

In previous work, we have developed the following three approaches:

- ProVerif: a symbolic prover for cryptographic protocol models
- Tookan: an attack-finder for PKCS#11 hardware security devices
- F*: a new language that enables the verification of cryptographic applications

3.1.1. Verifying cryptographic protocols with ProVerif

Given a model of a cryptographic protocol, the problem is to verify that an active attacker, possibly with access to some cryptographic keys but unable to guess other secrets, cannot thwart security goals such as authentication and secrecy [53]; it has motivated a serious research effort on the formal analysis of cryptographic protocols, starting with [49] and eventually leading to effective verification tools, such as our tool ProVerif.

To use ProVerif, one encodes a protocol model in a formal language, called the applied pi-calculus, and ProVerif abstracts it to a set of generalized Horn clauses. This abstraction is a small approximation: it just ignores the number of repetitions of each action, so ProVerif is still very precise, more precise than, say, tree automata-based techniques. The price to pay for this precision is that ProVerif does not always terminate; however, it terminates in most cases in practice, and it always terminates on the interesting class of *tagged protocols* [44]. ProVerif can handle a wide variety of cryptographic primitives, defined by rewrite rules or by some equations, and prove a wide variety of security properties: secrecy [42], [30], correspondences (including authentication) [43], and observational equivalences [41]. Observational equivalence means that an adversary cannot distinguish two processes (protocols); equivalences can be used to formalize a wide range of properties, but they are particularly difficult to prove. Even if the class of equivalences that ProVerif can prove is limited to equivalences between processes that differ only by the terms they contain, these equivalences are useful in practice and ProVerif has long been the only tool that proves equivalences for an unbounded number of sessions. (Maude-NPA in 2014 and Tamarin in 2015 adopted ProVerif's approach to proving equivalences.)

Using ProVerif, it is now possible to verify large parts of industrial-strength protocols, such as TLS [36], Signal [51], JFK [31], and Web Services Security [40], against powerful adversaries that can run an unlimited number of protocol sessions, for strong security properties expressed as correspondence queries or equivalence assertions. ProVerif is used by many teams at the international level, and has been used in more than 120 research papers (references available at http://proverif.inria.fr/proverif-users.html).

3.1.2. Verifying security APIs using Tookan

Security application programming interfaces (APIs) are interfaces that provide access to functionality while also enforcing a security policy, so that even if a malicious program makes calls to the interface, certain security properties will continue to hold. They are used, for example, by cryptographic devices such as smartcards and Hardware Security Modules (HSMs) to manage keys and provide access to cryptographic functions whilst keeping the keys secure. Like security protocols, their design is security critical and very difficult to get right. Hence formal techniques have been adapted from security protocols to security APIs.

The most widely used standard for cryptographic APIs is RSA PKCS#11, ubiquitous in devices from smartcards to HSMs. A 2003 paper highlighted possible flaws in PKCS#11 [46], results which were extended by formal analysis work using a Dolev-Yao style model of the standard [47]. However at this point it was not clear to what extent these flaws affected real commercial devices, since the standard is underspecified and can be implemented in many different ways. The Tookan tool, developed by Steel in collaboration with Bortolozzo, Centenaro and Focardi, was designed to address this problem. Tookan can reverse engineer the particular configuration of PKCS#11 used by a device under test by sending a carefully designed series of PKCS#11 commands and observing the return codes. These codes are used to instantiate a Dolev-Yao model of the device's API. This model can then be searched using a security protocol model checking tool to find attacks. If an attack is found, Tookan converts the trace from the model checker into the sequence of PKCS#11 queries needed to make the attack and executes the commands directly on the device. Results obtained by Tookan are remarkable: of 18 commercially available PKCS#11 devices tested, 10 were found to be susceptible to at least one attack.

3.1.3. Verifying cryptographic applications using F*

Verifying the implementation of a protocol has traditionally been considered much harder than verifying its model. This is mainly because implementations have to consider real-world details of the protocol, such as message formats [55], that models typically ignore. So even a protocol has been proved secure in theory, its implementation may be buggy and insecure. However, with recent advances in both program verification and symbolic protocol verification tools, it has become possible to verify fully functional protocol implementations in the symbolic model. One approach is to extract a symbolic protocol model from an implementation and then verify the model, say, using ProVerif. This approach has been quite successful, yielding a verified implementation of TLS in F# [39]. However, the generated models are typically quite large and whole-program symbolic verification does not scale very well.

An alternate approach is to develop a verification method directly for implementation code, using well-known program verification techniques. Our current focus is on designing and implementing the programming language F* [57], [34], [52], in collaboration with Microsoft Research. F* (pronounced F star) is an ML-like functional programming language aimed at program verification. Its type system includes polymorphism, dependent types, monadic effects, refinement types, and a weakest precondition calculus. Together, these features allow expressing precise and compact specifications for programs, including functional correctness and security properties. The F* type-checker aims to prove that programs meet their specifications using a combination of SMT solving and interactive proofs[23]. Programs written in F* can be translated to efficient OCaml, F#, or C for execution [54]. The main ongoing use case of F* is building a verified, drop-in replacement for the whole HTTPS stack in Project Everest [37] (a larger collaboration with Microsoft Research). This includes a verified implementation of TLS 1.2 and 1.3 [38] and of the underlying cryptographic primitives [58].

3.2. Computational verification of cryptographic applications

Proofs done by cryptographers in the computational model are mostly manual. Our goal is to provide computer support to build or verify these proofs. In order to reach this goal, we have designed the automatic tool CryptoVerif, which generates proofs by sequences of games. We already applied it to important protocols such as TLS [36] and Signal [51] but more work is still needed in order to develop this approach, so that it is easier to apply to more protocols. We also design and implement techniques for proving implementations of protocols secure in the computational model. In particular, CryptoVerif can generate implementations from CryptoVerif specifications that have been proved secure [45]. We plan to continue working on this approach.

A different approach is to directly verify cryptographic applications in the computational model by typing. A recent work [50] shows how to use refinement typechecking in F7 to prove computational security for protocol implementations. In this method, henceforth referred to as computational F7, typechecking is used as the main step to justify a classic game-hopping proof of computational security. The correctness of this method is based on a probabilistic semantics of F# programs and crucially relies on uses of type abstraction and parametricity to establish strong security properties, such as indistinguishability.

In principle, the two approaches, typechecking and game-based proofs, are complementary. Understanding how to combine these approaches remains an open and active topic of research.

An alternative to direct computation proofs is to identify the cryptographic assumptions under which symbolic proofs, which are typically easier to derive automatically, can be mapped to computational proofs. This line of research is sometimes called computational soundness and the extent of its applicability to real-world cryptographic protocols is an active area of investigation.

3.3. F*: A Higher-Order Effectful Language for Program Verification

F* [57], [34] is a verification system for effectful programs developed collaboratively by Inria and Microsoft Research. It puts together the automation of an SMT-backed deductive verification tool with the expressive power of a proof assistant based on dependent types. After verification, F* programs can be extracted to efficient OCaml, F#, or C code [54]. This enables verifying the functional correctness and security of realistic applications. F*'s type system includes dependent types, monadic effects, refinement types, and a weakest precondition calculus. Together, these features allow expressing precise and compact specifications for programs, including functional correctness and security properties. The F* type-checker aims to prove that programs meet their specifications using a combination of SMT solving and interactive proofs. The main ongoing use case of F* is building a verified, drop-in replacement for the whole HTTPS stack in Project Everest. This includes verified implementations of TLS 1.2 and 1.3 [38] and of the underlying cryptographic primitives [58].

3.4. Efficient Formally Secure Compilers to a Tagged Architecture

Severe low-level vulnerabilities abound in today's computer systems, allowing cyber-attackers to remotely gain full control. This happens in big part because our programming languages, compilers, and architectures were designed in an era of scarce hardware resources and too often trade off security for efficiency. The semantics of mainstream low-level languages like C is inherently insecure, and even for safer languages, establishing security with respect to a high-level semantics does not guarantee the absence of low-level attacks. Secure compilation using the coarse-grained protection mechanisms provided by mainstream hardware architectures would be too inefficient for most practical scenarios.

We aim to leverage emerging hardware capabilities for fine-grained protection to build the first, efficient secure compilation chains for realistic low-level programming languages (the C language, and Low* a safe subset of C embedded in F* for verification [54]). These compilation chains will provide a secure semantics for all programs and will ensure that high-level abstractions cannot be violated even when interacting with untrusted low-level code. To achieve this level of security without sacrificing efficiency, our secure compilation chains target a tagged architecture [35], which associates a metadata tag to each word and efficiently propagates and checks tags according to software-defined rules. We hope to experimentally evaluate and carefully optimize the efficiency of our secure compilation chains on realistic workloads and standard benchmark suites. We are also using property-based testing and formal verification to provide high confidence that our compilation chains are indeed secure. Formally, we are constructing machine-checked proofs of a new security criterion we call robustly safe compilation, which is defined as the preservation of safety properties even against an adversarial context [32], [33]. This strong criterion complements compiler correctness and ensures that no machine-code attacker can do more harm to securely compiled components than a component already could with respect to a secure source-level semantics.

3.5. Provably secure web applications

Web applications are fast becoming the dominant programming platform for new software, probably because they offer a quick and easy way for developers to deploy and sell their *app*s to a large number of customers. Third-party web-based apps for Facebook, Apple, and Google, already number in the hundreds of thousands and are likely to grow in number. Many of these applications store and manage private user data, such as health information, credit card data, and GPS locations. To protect this data, applications tend to use an ad hoc combination of cryptographic primitives and protocols. Since designing cryptographic applications is easy to get wrong even for experts, we believe this is an opportune moment to develop security libraries and verification techniques to help web application programmers.

As a typical example, consider commercial password managers, such as LastPass, RoboForm, and 1Password. They are implemented as browser-based web applications that, for a monthly fee, offer to store a user's passwords securely on the web and synchronize them across all of the user's computers and smartphones. The passwords are encrypted using a master password (known only to the user) and stored in the cloud. Hence, no-one except the user should ever be able to read her passwords. When the user visits a web page that has a login form, the password manager asks the user to decrypt her password for this website and automatically fills in the login form. Hence, the user no longer has to remember passwords (except her master password) and all her passwords are available on every computer she uses.

Password managers are available as browser extensions for mainstream browsers such as Firefox, Chrome, and Internet Explorer, and as downloadable apps for Android and Apple phones. So, seen as a distributed application, each password manager application consists of a web service (written in PHP or Java), some number of browser extensions (written in JavaScript), and some smartphone apps (written in Java or Objective C). Each of these components uses a different cryptographic library to encrypt and decrypt password data. How do we verify the correctness of all these components?

We propose three approaches. For client-side web applications and browser extensions written in JavaScript, we propose to build a static and dynamic program analysis framework to verify security invariants. To this end, we have developed two security-oriented type systems for JavaScript, Defensive JavaScript [48] and TS* [56], and used them to guarantee security properties for a number of JavaScript applications. For Android smartphone apps and web services written in Java, we propose to develop annotated JML cryptography libraries that can be used with static analysis tools like ESC/Java to verify the security of application code. For clients and web services written in F# for the .NET platform, we propose to use F* to verify their correctness. We also propose to translate verified F* web applications to JavaScript via a verified compiler that preserves the semantics of F* programs in JavaScript.

3.6. Design and Verification of next-generation protocols: identity, blockchains, and messaging

Building on our work on verifying and re-designing pre-existing protocols like TLS and Web Security in general, with the resources provided by the NEXTLEAP project, we are working on both designing and verifying new protocols in rapidly emerging areas like identity, blockchains, and secure messaging. These are all areas where existing protocols, such as the heavily used OAuth protocol, are in need of considerable re-design in order to maintain privacy and security properties. Other emerging areas, such as blockchains and secure messaging, can have modifications to existing pre-standard proposals or even a complete 'clean slate' design. As shown by Prosecco's work, newer standards, such as IETF OAuth, W3C Web Crypto, and

W3C Web Authentication API, can have vulnerabilities fixed before standardization is complete and heavily deployed. We hope that the tools used by Prosecco can shape the design of new protocols even before they are shipped to standards bodies. We have seen considerable progress in identity with the UnlimitID design and with messaging via the IETF MLS effort, with new work on blockchain technology underway.

4. Application Domains

4.1. Cryptographic Protocol Libraries

Cryptographic protocols such as TLS, SSH, IPSec, and Kerberos are the trusted base on which the security of modern distributed systems is built. Our work enables the analysis and verification of such protocols, both in their design and implementation. Hence, for example, we build and verify models and reference implementations for well-known protocols such as TLS and SSH, as well as analyze their popular implementations such as OpenSSL.

4.2. Hardware-based security APIs

Cryptographic devices such as Hardware Security Modules (HSMs) and smartcards are used to protect long-terms secrets in tamper-proof hardware, so that even attackers who gain physical access to the device cannot obtain its secrets. These devices are used in a variety of scenarios ranging from bank servers to transportation cards (e.g. Navigo). Our work investigates the security of commercial cryptographic hardware and evaluates the APIs they seek to implement.

4.3. Web application security

Web applications use a variety of cryptographic techniques to securely store and exchange sensitive data for their users. For example, a website may serve pages over HTTPS, authenticate users with a single sign-on protocol such as OAuth, encrypt user files on the server-side using XML encryption, and deploy client-side cryptographic mechanisms using a JavaScript cryptographic library. The security of these applications depends on the public key infrastructure (X.509 certificates), web browsers' implementation of HTTPS and the same origin policy (SOP), the semantics of JavaScript, HTML5, and their various associated security standards, as well as the correctness of the specific web application code of interest. We build analysis tools to find bugs in all these artifacts and verification tools that can analyze commercial web applications and evaluate their security against sophisticated web-based attacks.

5. Highlights of the Year

5.1. Highlights of the Year

- We published 12 papers at top-tier conferences and journals such as S&P (1), POPL (2), Euro S&P (2), ICFP (3), CSF (1), ESOP (1)
- Our cryptographic library HACL* was incorporated within the Linux kernel, Microsoft WinQuic, mbedTLS, and Concordium, in addition to the prior deployments in Mozilla Firefox and Tezos Blockchain
- Catalin Hritcu served as Program Chair of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP)

5.1.1. Awards

- EU Horizon Impact Award 2019 for Karthikeyan Bhargavan for his research on TLS 1.3.
- Distinguished Paper Award at CSF'19 for "Journey Beyond Full Abstraction"
- Distinguished Paper Award at POPL'19 for "Gradual Parametricity, Revisited"

BEST PAPERS AWARDS :

[17]
C. ABATE, R. BLANCO, D. GARG, C. HRIŢCU, M. PATRIGNANI, J. THIBAULT.*Journey Beyond Full Abstraction: Exploring Robust Property Preservation for Secure Compilation*, in "CSF 2019 - 32nd IEEE Computer Security Foundations Symposium", Hoboken, United States, IEEE, June 2019, p. 256-271, https://arxiv.org/abs/1807.04603 [*DOI :* 10.1109/CSF.2019.00025], https://hal.archives-ouvertes.fr/hal-02398915

[16]
M. TORO, E. LABRADA, É. TANTER.*Gradual Parametricity, Revisited*, in "Proceedings of the ACM on Programming Languages", 2019, vol. 3, n^o POPL, https://arxiv.org/abs/1807.04596 [*DOI :* 10.1145/3290330], https://hal.archives-ouvertes.fr/hal-01960553

6. New Software and Platforms

6.1. Cryptosense Analyzer

SCIENTIFIC DESCRIPTION: Cryptosense Analyzer (formerly known as Tookan) is a security analysis tool for cryptographic devices such as smartcards, security tokens and Hardware Security Modules that support the most widely-used industry standard interface, RSA PKCS#11. Each device implements PKCS#11 in a slightly different way since the standard is quite open, but finding a subset of the standard that results in a secure device, i.e. one where cryptographic keys cannot be revealed in clear, is actually rather tricky. Cryptosense Analyzer analyses a device by first reverse engineering the exact implementation of PKCS#11 in use, then building a logical model of this implementation for a model checker, calling a model checker to search for attacks, and in the case where an attack is found, executing it directly on the device. It has been used to find at least a dozen previously unknown flaws in commercially available devices.

FUNCTIONAL DESCRIPTION: Cryptosense Analyzer (formerly known as Tookan) is a security analysis tool for cryptographic devices such as smartcards,

- Participants: Graham Steel and Romain Bardou
- Contact: Graham Steel
- URL: https://cryptosense.com/

6.2. CryptoVerif

Cryptographic protocol verifier in the computational model

KEYWORDS: Security - Verification - Cryptographic protocol

FUNCTIONAL DESCRIPTION: CryptoVerif is an automatic protocol prover sound in the computational model. In this model, messages are bitstrings and the adversary is a polynomial-time probabilistic Turing machine. CryptoVerif can prove secrecy and correspondences, which include in particular authentication. It provides a generic mechanism for specifying the security assumptions on cryptographic primitives, which can handle in particular symmetric encryption, message authentication codes, public-key encryption, signatures, hash functions, and Diffie-Hellman key agreements. It also provides an explicit formula that gives the probability of breaking the protocol as a function of the probability of breaking each primitives, this is the exact security framework.

NEWS OF THE YEAR: We implemented the following features in CryptoVerif:

1) We added to the library of cryptographic primitives several variants of the PRF-ODH (pseudo-random function oracle Diffie-Hellman) assumption, pre-image resistant and second-preimage resistant hash functions, IND-CPA encryption with a nonce, IND-CPA and INT-CTXT encryption with a nonce, encryption schemes that satisfy IND$-CPA instead of IND-CPA.

2) To facilitate modular proofs, we allow querying indistinguishability properties with exactly the same syntax as the one used to specify indistinguishability assumptions on primitives.

3) To simplify declarations of assumptions on primitives, replications (which model any number of copies of processes or oracles) can be omitted at the root of indistinguishability assumptions. CryptoVerif adds them internally, thus inferring the assumption for N independent copies from the assumption for one copy. For instance, it infers the assumption for encryption with N keys from the assumption for encryption with a single key.

4) When we delay random number generations, we allow the user to specify expressions for which it is not necessary to generate the random value, so that the generation of the moved random value can be delayed further. In particular, we used this extension to prove that the OAEP scheme is IND-CCA2 assuming the underlying permutation is partial-domain one-way (a famous cryptographic result).

5) CryptoVerif can now remove parts of the code cannot be executed in case the adversary wins the game, by replacing them with event "AdversaryLoses". That is specially helpful in order to deal with complex cases of key compromise, e.g. for forward secrecy, by proving authentication by ignoring the compromise, showing that authentication is preserved in case the key is compromised (because the adversary never wins against the considered authentication property in case of compromise), and using the authentication to prove secrecy even in case of compromise. For instance, that allows us to show that the PSK-DHE handshake of TLS 1.3 preserves forward secrecy in case of compromise of the PSK.

6) After a cryptographic transformation, CryptoVerif expands terms into processes, which leads to duplicating code until the end of the protocol for each test that is expanded. The cryptographic transformation and the expansion were initially considered as a single transformation. There are now considered as separate transformations, so that other transformations can be performed in between, in particular to cut some branches of the code and reduce the code duplication.

These changes are included in CryptoVerif version 2.02 available at https://cryptoverif.inria.fr.

- Participants: Bruno Blanchet and David Cadé
- Contact: Bruno Blanchet
- Publications: Composition Theorems for CryptoVerif and Application to TLS 1.3 - Composition Theorems for CryptoVerif and Application to TLS 1.3 - A Mechanised Cryptographic Proof of the WireGuard Virtual Private Network Protocol - A Mechanised Cryptographic Proof of the WireGuard Virtual Private Network Protocol - Proved Implementations of Cryptographic Protocols in the Computational Model - Proved Generation of Implementations from Computationally Secure Protocol Specifications - Verified Models and Reference Implementations for the TLS 1.3 Standard Candidate - Verified Models and Reference Implementations for the TLS 1.3 Standard Candidate - Symbolic and Computational Mechanized Verification of the ARINC823 Avionic Protocols - Automated Verification for Secure Messaging Protocols and Their Implementations: A Symbolic and Computational Approach
- URL: http://cryptoverif.inria.fr/

6.3. F*

FStar

KEYWORDS: Programming language - Software Verification

FUNCTIONAL DESCRIPTION: F* is a new higher order, effectful programming language (like ML) designed with program verification in mind. Its type system is based on a core that resembles System Fw (hence the name), but is extended with dependent types, refined monadic effects, refinement types, and higher kinds. Together, these features allow expressing precise and compact specifications for programs, including functional correctness properties. The F* type-checker aims to prove that programs meet their specifications using an automated theorem prover (usually Z3) behind the scenes to discharge proof obligations. Programs written in F* can be translated to OCaml, F#, or JavaScript for execution.

- Participants: Antoine Delignat-Lavaud, Catalin Hritcu, Cedric Fournet, Chantal Keller, Karthikeyan Bhargavan and Pierre-Yves Strub

- Contact: Catalin Hritcu
- URL: https://www.fstar-lang.org/

6.4. miTLS

KEYWORDS: Cryptographic protocol - Software Verification

FUNCTIONAL DESCRIPTION: miTLS is a verified reference implementation of the TLS protocol. Our code fully supports its wire formats, ciphersuites, sessions and connections, re-handshakes and resumptions, alerts and errors, and data fragmentation, as prescribed in the RFCs, it interoperates with mainstream web browsers and servers. At the same time, our code is carefully structured to enable its modular, automated verification, from its main API down to computational assumptions on its cryptographic algorithms.

- Participants: Alfredo Pironti, Antoine Delignat-Lavaud, Cedric Fournet, Jean-Karim Zinzindohoué, Karthikeyan Bhargavan, Pierre-Yves Strub and Santiago Zanella
- Contact: Karthikeyan Bhargavan
- URL: https://github.com/mitls/mitls-fstar

6.5. ProVerif

KEYWORDS: Security - Verification - Cryptographic protocol

FUNCTIONAL DESCRIPTION: ProVerif is an automatic security protocol verifier in the symbolic model (so called Dolev-Yao model). In this model, cryptographic primitives are considered as black boxes. This protocol verifier is based on an abstract representation of the protocol by Horn clauses. Its main features are:

It can verify various security properties (secrecy, authentication, process equivalences).

It can handle many different cryptographic primitives, specified as rewrite rules or as equations.

It can handle an unbounded number of sessions of the protocol (even in parallel) and an unbounded message space.

NEWS OF THE YEAR: Vincent Cheval and Bruno Blanchet worked on several extensions of ProVerif: 1) support for integer counters, with incrementation and inequality tests, 2) lemmas and axioms to give intermediate results to ProVerif, which it exploits to help proving subsequent queries, by deriving additional information in the Horn clauses that it uses to perform the proofs, 3) proofs by induction on the length of the trace, by giving as lemma the property to prove, but obviously for strictly shorter traces. Detailed soundness proofs for these extensions are in progress. These features are not released yet.

- Participants: Bruno Blanchet, Marc Sylvestre and Vincent Cheval
- Contact: Bruno Blanchet
- Publications: Automated reasoning for equivalences in the applied pi calculus with barriers - Automated Reasoning for Equivalences in the Applied Pi Calculus with Barriers - Automated reasoning for equivalences in the applied pi calculus with barriers - Modeling and Verifying Security Protocols with the Applied Pi Calculus and ProVerif - Automatic Verification of Security Protocols in the Symbolic Model: The Verifier ProVerif - Verified Models and Reference Implementations for the TLS 1.3 Standard Candidate - Verified Models and Reference Implementations for the TLS 1.3 Standard Candidate - Automated Verification for Secure Messaging Protocols and Their Implementations: A Symbolic and Computational Approach - Symbolic and Computational Mechanized Verification of the ARINC823 Avionic Protocols - Symbolic and Computational Mechanized Verification of the ARINC823 Avionic Protocols
- URL: http://proverif.inria.fr/

6.6. HACL*

High Assurance Cryptography Library

KEYWORDS: Cryptography - Software Verification

FUNCTIONAL DESCRIPTION: HACL* is a formally verified cryptographic library in F*, developed by the Prosecco team at Inria Paris in collaboration with Microsoft Research, as part of Project Everest.

HACL stands for High-Assurance Cryptographic Library and its design is inspired by discussions at the HACS series of workshops. The goal of this library is to develop verified C reference implementations for popular cryptographic primitives and to verify them for memory safety, functional correctness, and secret independence.

- Contact: Karthikeyan Bhargavan
- URL: https://github.com/mitls/hacl-star

7. New Results

7.1. Verification of security protocols

Participants: Bruno Blanchet, Karthikeyan Bhargavan, Benjamin Lipp.

Our verification of the WireGuard open source Virtual Private Network (VPN) with CryptoVerif appears at EuroS&P 2019 [22], [27].

We continued the development of our protocol verification tools ProVerif and CryptoVerif. The new features of this year are detailed in the section on software.

In the setting of the ANR AnaStaSec project, we worked on the verification of avionic security protocols. More specifically, in 2015, we had verified the protocol of the Secure Dialog Service using ProVerif and CryptoVerif and recommended many changes to the specification. The ICAO started to take into account our remarks, and this year we analyzed a new version of the specification. Our analysis showed that many recommendations still need to be taken into account. Additionally, we also commented on the recent choice of using DTLS over UDP to secure the future ATN/IPS (Aeronautical Telecommunication Network / Internet Protocol Suite) network, which seems very positive. The details of these results are still confidential; they have been provided to ANR.

7.2. Verified Software for Cryptographic Web Applications

Participants: Karthikeyan Bhargavan, Benjamin Beurdouche, Denis Merigoux, Jonathan Protzenko.

WebAssembly in a new language runtime that is now supported by all major web browsers and web application frameworks. We developed a compiler from the Low* subset of the F* programming language to WebAssembly and used this compiler to translate our HACL* verified cryptographic library to WebAssembly, hence obtaining the first verified cryptographic library for the Web. We also used this framework to develop and verify an implementation of the Signal protocol in WebAssembly, and demonstrated how this implementation can be used as a drop-in replacement for the libsignal-protocol library used in mainstream messaging applications like Signal, WhatsApp, and Skype.

Our work was published at the IEEE Security and Privacy conference [24]. Our WebAssembly version of HACL* and our verified Signal implementation were publicly released as open source on GitHub.

7.3. Journey beyond full abstraction

Participants: Carmine Abate, Roberto Blanco, Deepak Garg [MPI-SWS], Catalin Hritcu, Marco Patrignani [Stanford and CISPA], Jérémy Thibault.

Even for safe languages, all guarantees are lost when interacting with low-level code, for instance when using low-level libraries. A compromised or malicious library that gets linked in can easily read and write data and code, jump to arbitrary instructions, or smash the stack, blatantly violating any source-level abstraction and breaking any guarantee obtained by source-level reasoning. Our goal is to build formally secure compartmentalizing compilation chains that defend against such attacks. We started by investigating what it means for a compilation chain to provide secure interoperability between a safe source language and linked target-level code that is adversarial. In this model, a secure compilation chain ensures that even linked adversarial target-level code cannot break the security properties of a compiled program any more than some linked source-level code could. However, the precise class of security properties one chooses to preserve crucially impacts not only the supported security goals and the strength of the attacker model, but also the kind of protections the compilation chain has to introduce and the kind of proof techniques one can use to make sure that the protections are watertight. We are the first to thoroughly explore a large space of secure compilation criteria based on the preservation against adversarial contexts of various classes of trace properties such as safety, of hyperproperties such as noninterference, and of relational hyperproperties such as trace equivalence [17], [10].

7.4. Principles of Program Verification for Arbitrary Monadic Effects

Participants: Kenji Maillard, Danel Ahman [University of Ljubljana], Robert Atkey [University of Strathclyde], Guido Martinez, Catalin Hritcu, Exequiel Rivas, Éric Tanter, Antoine Van Muylder, Cezar Andrici.

We devised a principled semantic framework for verifying programs with arbitrary monadic effects in a generic way with respect to expressive specifications. The starting point are Dijkstra monads, which are monad-like structures that classify effectful computations satisfying a specification drawn from a monad. Dijkstra monads have already proven valuable in practice for verifying effectful code, and in particular, they allow the F* program verifier to compute verification conditions.

We provide the first semantic investigation of the algebraic structure underlying Dijkstra monads [13], [11] and unveil a close relationship between Dijkstra monads and effect observations, i.e., mappings between a computational and a specification monad that respect their monadic structure. Effect observations are flexible enough to provide various interpretations of effects, for instance total vs partial correctness, or angelic vs demonic nondeterminism. Our semantic investigation relies on a general theory of specification monads and effect observations, using an enriched notion of relative monads and relative monad morphisms. We moreover show that a large variety of specification monads can be obtained by applying monad transformers to various base specification monads, including predicate transformers and Hoare-style pre- and postconditions. For defining correct monad transformers, we design a language inspired by the categorical analysis of the relationship between monad transformers and algebras for a monad.

We also adapt our framework to relational verification [14], [11], i.e., proving relational properties between multiple runs of one or more programs, such as noninterference or program equivalence. For this we extend specification monads and effect observations to the relational setting and use them to derive the semantics and core rules of a relational program logic generically for any monadic effect. Finally, we identify and overcome conceptual challenges that prevented previous relational program logics from properly dealing with effects such as exceptions, and are the first to provide a proper semantic foundation and a relational program logic for exceptions.

7.5. Meta-F*: Proof automation with SMT, Tactics, and Metaprograms

Participants: Guido Martinez, Danel Ahman, Victor Dumitrescu, Nick Giannarakis [Princeton University], Chris Hawblitzel [Microsoft Research], Catalin Hritcu, Monal Narasimhamurthy [University of Colorado Boulder], Zoe Paraskevopoulou [Princeton University], Clément Pit-Claudel [MIT], Jonathan Protzenko [Microsoft Research], Tahina Ramananandro [Microsoft Research], Aseem Rastogi [Microsoft Research], Nikhil Swamy [Microsoft Research].

We introduced Meta-F*[23], a tactics and metaprogramming framework for the F* program verifier. The main novelty of Meta-F* is allowing to use tactics and metaprogramming to discharge assertions not solvable by SMT, or to just simplify them into well-behaved SMT fragments. Plus, Meta-F* can be used to generate verified code automatically.

Meta-F* is implemented as an F* effect, which, given the powerful effect system of F*, heavily increases code reuse and even enables the lightweight verification of metaprograms. Metaprograms can be either interpreted, or compiled to efficient native code that can be dynamically loaded into the F* type-checker and can interoperate with interpreted code. Evaluation on realistic case studies shows that Meta-F* provides substantial gains in proof development, efficiency, and robustness.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

8.1.1. Evolution, Semantics, and Engineering of the F Verification System*

Grant from Nomadic Labs - Inria

PIs: Catalin Hritcu and Exequiel Rivas

Duration: March 2019 - April 2023

Abstract: While the F* verification system shows great promise in practice, many challenging conceptual problems remain to be solved, many of which can directly inform the further evolution and design of the language. Moreover, many engineering challenges remain in order to build a truly usable verification system. This proposal promises to help address this by focusing on the following 5 main topics: (1) *Generalizing Dijkstra monads*, i.e., a program verification technique for arbitrary monadic effects; (2) *Relational reasoning in F**: devising scalable verification techniques for properties of multiple program executions (e.g., confidentiality, noninterference) or of multiple programs (e.g., program equivalence); (3) *Making F*'s effect system more flexible*, by supporting tractable forms of effect polymorphism and allowing some of the effects of a computation to be hidden if they do not impact the observable behavior; (4) Working out more of the *F* semantics and metatheory*; (5) Solving the *engineering challenges* of building a usable verification system.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

9.1.1.1. AnaStaSec

Title: Static Analysis for Security Properties (ANR générique 2014.)

Other partners: Inria Paris/EPI Antique, Inria Rennes/EPI Celtique, Airbus Operations SAS, AMOSSYS, CEA-LIST, TrustInSoft

Duration: January 2015 - September 2019.

Coordinator: Jérôme Féret, EPI Antique, Inria Paris (France)

Participant: Bruno Blanchet

Abstract: The project aims at using automated static analysis techniques for verifying security and confidentiality properties of critical avionics software.

9.1.1.2. AJACS

Title: AJACS: Analyses of JavaScript Applications: Certification and Security

Other partners: Inria-Rennes/Celtique, Inria-Saclay/Toccata, Inria-Sophia Antipolis/INDES, Imperial College London

Duration: October 2014 - March 2019.

Coordinator: Alan Schmitt, Inria (France)

Participants: Karthikeyan Bhargavan, Bruno Blanchet, Nadim Kobeissi

Abstract: The goal of the AJACS project is to provide strong security and privacy guarantees for web application scripts. To this end, we propose to define a mechanized semantics of the full JavaScript language, the most widely used language for the Web, to develop and prove correct analyses for JavaScript programs, and to design and certify security and privacy enforcement mechanisms.

9.1.1.3. SafeTLS

Title: SafeTLS: La sécurisation de l'Internet du futur avec TLS 1.

Other partners: Université Rennes 1, IRMAR, Inria Sophia Antipolis, SGDSN/ANSSI

Duration: October 2016 - September 2020

Coordinator: Pierre-Alain Fouque, Université de Rennes 1 (France)

Participants: Karthikeyan Bhargavan

Abstract: Our project, SafeTLS, addresses the security of both TLS 1.3 and of TLS 1.2 as they are (expected to be) used, in three important ways: (1) A better understanding: We will provide a better understanding of how TLS 1.2 and 1.3 are used in real-world applications; (2) Empowering clients: By developing a tool that will show clients the quality of their TLS connection and inform them of potential security and privacy risks; (3) Analyzing implementations: We will analyze the soundness of current TLS 1.2 implementations and use automated verification to provide a backbone of a secure TLS 1.3 implementation.

9.1.1.4. TECAP

Title: TECAP: Protocol Analysis - Combining Existing Tools (ANR générique 2017.)

Other partners: Inria Nancy/EPI PESTO, Inria Sophia Antipolis/EPI MARELLE, IRISA, LIX, LSV - ENS Cachan.

Duration: January 2018 - December 2021

Coordinator: Vincent Cheval, EPI PESTO, Inria Nancy (France)

Participants: Bruno Blanchet, Benjamin Lipp

Abstract: A large variety of automated verification tools have been developed to prove or find attacks on security protocols. These tools differ in their scope, degree of automation, and attacker models. The aim of this project is to get the best of all these tools, meaning, on the one hand, to improve the theory and implementations of each individual tool towards the strengths of the others and, on the other hand, build bridges that allow the cooperations of the methods/tools. We will focus in this project on the tools CryptoVerif, EasyCrypt, Scary, ProVerif, Tamarin, AKiSs and APTE.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. ERC Consolidator Grant: CIRCUS

Title: CIRCUS: An end-to-end verification architecture for building Certified Implementations of Robust, Cryptographically Secure web applications

Duration: April 2016 - March 2021

Coordinator: Karthikeyan Bhargavan, Inria

The security of modern web applications depends on a variety of critical components including cryptographic libraries, Transport Layer Security (TLS), browser security mechanisms, and single sign-on protocols. Although these components are widely used, their security guarantees remain poorly understood, leading to subtle bugs and frequent attacks. Rather than fixing one attack at a time, we advocate the use of formal security verification to identify and eliminate entire classes of vulnerabilities in one go.

CIRCUS proposes to take on this challenge, by verifying the end-to-end security of web applications running in mainstream software. The key idea is to identify the core security components of web browsers and servers and replace them by rigorously verified components that offer the same functionality but with robust security guarantees.

9.2.1.2. ERC Starting Grant: SECOMP

Title: SECOMP: Efficient Formally Secure Compilers to a Tagged Architecture

Duration: Jan 2017 - December 2021

Coordinator: Catalin Hritcu, Inria

Abstract: The SECOMP project is aimed at leveraging emerging hardware capabilities for fine-grained protection to build the first, efficient secure compilation chains for realistic low-level programming languages (the C language, and Low* a safe subset of C embedded in F* for verification). These compilation chains will provide a secure semantics for all programs and will ensure that high-level abstractions cannot be violated even when interacting with untrusted low-level code. To achieve this level of security without sacrificing efficiency, our secure compilation chains target a tagged architecture, which associates a metadata tag to each word and efficiently propagates and checks tags according to software-defined rules. We will use property-based testing and formal verification to provide high confidence that our compilers are indeed secure.

9.2.1.3. NEXTLEAP (304)

Title: NEXTLEAP: NEXT generation Legal Encryption And Privacy

Programm: H2020

Duration: January 2016 - December 2018

Coordinator: Harry Halpin, Inria

Other partners: IMDEA, University College London, CNRS, IRI, and Merlinux

The objective of the NEXTLEAP project is to build the fundamental interdisciplinary internet science necessary to create decentralized, secure, and rights-preserving protocols for the next generation of collective awareness platforms. The long-term goal of NEXTLEAP is to have Europe take the "next leap ahead" of the rest of the world by solving the fundamental challenge of determining how both to scientifically build and how to help citizens and institutions adopt open-source decentralized and privacy-preserving digital social platforms in contrast to proprietary centralized cloud-based services and pervasive surveillance that function at the expense of rights and technological sovereignty.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

We have a range of long- and short-term collaborations with various universities and research labs. We summarize them by project:

- TLS analysis: Microsoft Research (Cambridge), Mozilla, University of Rennes
- F*: Microsoft Research (Redmond, Cambridge, Bangalore), MSR-Inria, CMU, MIT, University of Ljubljana, Nomadic Labs, Zen Protocol, Princeton University
- SECOMP: MPI-SWS, CISPA, Stanford University, CMU, University of Pennsylvania, Portland State University, University of Virginia, University of Iai
- Micro-Policies: University of Pennsylvania, Portland State University, MIT, Draper Labs, Dover Microsystems

9.3.2. Participation in Other International Programs

9.3.2.1. SSITH/HOPE

Title: Advanced New Hardware Optimized for Policy Enforcement, A New HOPE

Program: DARPA SSITH

Duration: December 2017 - February 2021

Coordinator: Charles Stark Draper Laboratory

Other Participants: Inria Paris, University of Pennsylvania, MIT, Portland State University, Dover Microsystems, DornerWorks

Participants from Inria Prosecco: Catalin Hritcu, Roberto Blanco, Jérémy Thibault

Abstract: A New HOPE builds on results from the Inherently Secure Processor (ISP) project that has been internally funded at Draper. Recent architectural improvements decouple the tagged architecture from the processor pipeline to improve performance and flexibility for new processors. HOPE securely maintains metadata for each word in application memory and checks every instruction against a set of installed security policies. The HOPE security architecture exposes tunable parameters that support Performance, Power, Area, Software compatibility and Security (PPASS) search space exploration. Flexible software-defined security policies cover all 7 SSITH CWE vulnerability classes, and policies can be tuned to meet PPASS requirements; for example, one can trade granularity of security checks against performance using different policy configurations. HOPE will design and formalize a new high-level domain-specific language (DSL) for defining security policies, based on previous research and on extensive experience with previous policy languages. HOPE will formally verify that installed security policies satisfy system-wide security requirements. A secure boot process enables policies to be securely updated on deployed HOPE systems. Security policies can adapt based on previously detected attacks. Over the multi-year, multi-million dollar Draper ISP project, the tagged security architecture approach has evolved from early prototypes based on results from the DARPA CRASH program towards easier integration with external designs, and is better able to scale from micro to server class implementations. A New HOPE team is led by Draper and includes faculty from University of Pennsylvania (Penn), Portland State University (PSU), Inria, and MIT, as well as industry collaborators from DornerWorks and Dover Microsystems. In addition to Draper's in-house expertise in hardware design, cyber-security (defensive and offensive, hardware and software) and formal methods, the HOPE team includes experts from all domains relevant to SSITH, including (a) computer architecture: DeHon (Penn), Shrobe (MIT); (b) formal methods including programming languages and security: Pierce (Penn), Tolmach (PSU), Hritcu (Inria); and (c) operating system integration (DornerWorks). Dover Microsystems is a spin-out from Draper that will commercialize concepts from the Draper ISP project.

9.3.2.2. Everest Expedition

Program: Microsoft Expedition and MSR-Inria Collaborative Research Project

Expedition Participants: Microsoft Research (Cambridge, Redmond, Bangalore), Inria, MSR-Inria, CMU, University of Edinburgh

Duration of current MSR-Inria Project: October 2017 – October 2020

Participants from Inria Prosecco: Karthikeyan Bhargavan, Catalin Hritcu, Danel Ahman, Benjamin Beurdouche, Victor Dumitrescu, Nadim Kobeissi, Théo Laurent, Guido Martínez, Denis Merigoux, Marina Polubelova, Jean-Karim Zinzindohoué

Participants from other Inria teams: David Pichardie (Celtique), Jean-Pierre Talpin (TEA)

Abstract: The HTTPS ecosystem (HTTPS and TLS protocols, X.509 public key infrastructure, crypto algorithms) is the foundation on which Internet security is built. Unfortunately, this ecosystem is brittle, with headline-grabbing attacks such as FREAK and LogJam and emergency patches many times a year.

Project Everest addresses this problem by constructing a high-performance, standards-compliant, formally verified implementation of components in HTTPS ecosystem, including TLS, the main protocol at the heart of HTTPS, as well as the main underlying cryptographic algorithms such as AES, SHA2 or X25519.

At the TLS level, for instance, we are developing new implementations of existing and forthcoming protocol standards and formally proving, by reduction to cryptographic assumptions on their core algorithms, that our implementations provide a secure-channel abstraction between the communicating endpoints. Implementations of the core algorithms themselves are also verified, producing performant portable C code or highly optimized assembly language.

We aim for our verified components to be drop-in replacements suitable for use in mainstream web browsers, servers, and other popular tools and are actively working with the community at large to improve the ecosystem.

https://project-everest.github.io

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Éric Tanter (University of Chile) joined Inria as a Visiting Professor from Jul 2018 to March 2019 and from August to December 2019; he gave various seminars at Inria including one entitled "Gradual Parametricity, Revisited";
- Li-yao Xia (University of Pennsylvania) visited Prosecco on 7 January and gave a talk entitled "From C to Interaction Trees";
- Matías Toro (University of Chile) visited Prosecco on 9 January and gave a talk entitled "Type-Driven Gradual Security with References";
- Deepak Garg (MPI-SWS) visited Prosecco on 29 January and 20 November;
- Gilles Barthe (MPI-SP) visited Prosecco on various occasions: 29 January, 3–6 June, 9–13 Sept, and 7–9 October 2019;
- Jeremy Siek (Indiana University) visited Prosecco on 21 February and gave a seminar entitled "Toward Efficient Gradual Typing";
- Andrew Tolmach (Portland State University) visted Prosecco on 8–12 April and gave a seminar on "Enforcing C-level security policies using machine-level tags";
- Guido Martinez (CIFASIS-CONICET Rosario) visited Prosecco on various occasions: April 15–19, ICFP, 30 September to 12 October
- Nikos Vasilakis (University of Pennsylvania) visited Prosecco on 15–19 July and gave a seminar on "Retrofitting Security, Module by Module";
- Clement Pit-Claudel (MPI) visited Prosecco on 14 August;
- Kevin Liao (MPI-SP) visited Prosecco on various occasions and gave a seminar on "ILC: A Calculus for Composable, Computational Cryptography";
- Tahina Ramananandro (Microsoft Research) visited Prosecco on 30 September to 15 October and gave a seminar on "EverParse";
- Nik Swamy (Microsoft Research) and Aymeric Fromherz (CMU) visited Prosecco from 7–11 October and gave a seminar on "Verifying a mixture of C and assembly code with Low* and Vale";
- Jonathan Protzenko (Microsoft Research) visited Prosecco on 30 September to 15 October and gave a seminar on "The EverCrypt verified cryptographic provider";
- Jakob von Raumer (University of Nottingham) visited Prosecco on 23 October and gave a seminar on "Indexed Inductive Types";

- Bas Spitters (COBRA, Aarhus University) visited Prosecco on 25–29 November and gave a seminar on “ConCert: A Smart Contract Certification Framework in Coq”;
- Adrien Koutsos (MPI-SP) visited Prosecco on 5 November and gave a talk on “5G-AKA authentication protocol privacy”;
- Akram El-Korashy (MPI-SWS) visited Prosecco on 20 November;
- Shin-ya Katsumata (NII, Tokyo, Japan) visited Prosecco on 25–28 November;
- Ian Miers (Johns Hopkins University) visited Prosecco on 29 November and gave a seminar on “Zcash, Blockchains, and the possibilities for formal verification with zero-knowledge”;

9.4.1.1. Internships

- Antoine Van Muylder (Paris 7): from April to September 2019 – advised by Catalin Hritcu, Exequiel Rivas, and Kenji Maillard
- Guillaume Gette: from April to September 2019 – advised by Karthikeyan Bhargavan
- Mikhail Volkhov: from April to August 2019 – advised by Karthikeyan Bhargavan and Prasad Naldurg

9.4.2. Visits to International Teams

- Catalin Hritcu visited EPFL Lausanne on 25–27 September;
- Catalin Hritcu, Carmine Abate, Roberto Blanco, and Jeremy Thibault visited MPI-SWS in Saarbrücken on 18–22 October and 1–3 December;
- Catalin Hritcu visited Chalmers University in Gothenburg on 4–6 December;

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- Catalin Hritcu is the Steering Committee Chair of the Workshop on Principles of Secure Compilation (PriSC)
- Catalin Hritcu is the main organizer a Dagstuhl Seminar on Secure Compilation (20201)
- Karthikeyan Bhargavan co-chaired the Workshop on Secure Messaging at EUROCRYPT 2019

10.1.1.2. Member of the Organizing Committees

- Catalin Hritcu is a Steering Committee Member of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP)

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- Catalin Hritcu served as Program Chair of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP)

10.1.2.2. Member of the Conference Program Committees

- Karthikeyan Bhargavan was PC member of CCS 2019
- Bruno Blanchet was PC member of CSF 2019
- Catalin Hritcu was PC member of RV 2019, SecDev 2019, CSF 2020

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Associate Editor

– of the *International Journal of Applied Cryptography (IJACT)* – Inderscience Publishers: Bruno Blanchet

10.1.4. Invited Talks

- Karthikeyan Bhargavan gave invited talks at ERCIM Rome, FSTTCS Mumbai, ICISS Hyderabad,
- Catalin Hritcu gave invited talks at EPFL Lausanne, Ruhr University Bochum, MPI-SWS, and Chalmers University;

10.1.5. Leadership within the Scientific Community

- Catalin Hritcu served as the Artifact Evaluation Co-Chair for POPL 2018 and POPL 2019

10.1.6. Scientific Expertise

- Bruno Blanchet is a member of the specialized temporary scientific committee of ANSM (*Agence nationale de sécurité du médicament et des produits de santé*), on the cybersecurity of software medical devices.
- Bruno Blanchet was a scientific consultant for Nomadic Labs, regarding the development of the blockchain Tezos.
- Karthikeyan Bhargavan was scientific consultant for Nomadic Labs, regarding verified cryptographic software.

10.1.7. Research Administration

- Bruno Blanchet was a member of the hiring scientific jury for Inria researchers (*chargé de recherche*) of the Inria Paris center.
- Bruno Blanchet was a representative of Inria Paris at the DIM RFSI (*Domaine d'Intérêt Majeur, Réseau Francilien en Sciences Informatiques*).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Master: Bruno Blanchet, Cryptographic protocols: formal and computational proofs, 18h equivalent TD, master M2 MPRI, université Paris VII
- Master: Karthikeyan Bhargavan, Cryptographic protocols: formal and computational proofs, 18h equivalent TD, master M2 MPRI, université Paris VII
- PhD: Karthikeyan Bhargavan, Verified Crypto for Verified Protocols, Sibenik Summer School on real-world crypto and privacy, 17-21 June, 2019
- PhD: Catalin Hritcu, Program Verification with F* course at Summer School on Verification Technology, Systems, and Applications, VSTA 2019, 1-5 July 2019 at University of Luxembourg
- PhD: Catalin Hritcu, Writing and Verifying Functional Programs in Coq course at Summer School on Cryptography, Blockchain, and Program Verification, Mathinfoly 2019, 24-31 August 2019 at INSA, Lyon

10.2.2. Supervision

- PhD in progress: Benjamin Lipp, On Mechanised Cryptographic Proofs of Protocols and their Link with Verified Implementations, ENS Paris, since October 2018, supervised by Bruno Blanchet and Karthikeyan Bhargavan.
- PhD in progress: Benjamin Beurdouche, Formal Verification for Real-World Cryptographic Protocols, PSL, since September 2016, supervised by Karthikeyan Bhargavan.

- PhD in progress: Natalia Kulatova, Formal Analysis of Security Devices, PSL, since September 2017, supervised by Karthikeyan Bhargavan and Graham Steel.
- PhD in progress: Marina Polubelova, Formal Verification of a Cryptographic Library, PSL, since September 2017, supervised by Karthikeyan Bhargavan.
- PhD in progress: Denis Merigoux, Verification framework for performance-oriented memory-safe programming languages, since September 2018, supervised by Karthikeyan Bhargavan and Jonathan Protzenko.
- PhD: Kenji Maillard, on Principles of Program Verification for Arbitrary Monadic Effects, started January 2017, supervised by Catalin Hritcu
- PhD in progress: Carmine Abate, The Formal Foundations of Secure Compilation, since June 2018, advised by Catalin Hritcu
- PhD in progress: Jérémy Thibault, Secure Compartmentalizing Compilation to a Tagged Architecture, from August 2018, advised by Catalin Hritcu
- PhD in progress: Guido Martínez (CIFASIS-CONICET Rosario), Metatheory for Semi-Automatic Verification of Effectful Programs, from April 2017, advised by Mauro Jaskelioff (CIFASIS-CONICET Rosario) and Catalin Hritcu

10.2.3. Juries

- Bruno Blanchet was member of the PhD jury of Adrien Koutsos (ENS Paris-Saclay).
- Karthikeyan Bhargavan was member of the PhD jury of Joseph Lallemand (Univ. Lorraine) and Guido Martinez (Univ Stuttgart).
- Catalin Hritcu was a discussion leader for the Licentiate defense of Maximilian Algehed (Chalmers University);

10.3. Popularization

10.3.1. Articles and contents

- Catalin Hritcu contributed an article on Secure Compilation to the SIGPLAN PL Perspectives blog
- Karthikeyan Bhargavan published a paper in Communications of the ACM

11. Bibliography

Major publications by the team in recent years

[1] M. ABADI, B. BLANCHET, C. FOURNET.*The Applied Pi Calculus: Mobile Values, New Names, and Secure Communication*, in "Journal of the ACM (JACM)", October 2017, vol. 65, n$^{\text{o}}$ 1, p. 1 - 103 [*DOI :* 10.1145/3127586], https://hal.inria.fr/hal-01636616

[2] C. ABATE, A. AZEVEDO DE AMORIM, R. BLANCO, A. N. EVANS, G. FACHINI, C. HRIŢCU, T. LAURENT, B. C. PIERCE, M. STRONATI, A. TOLMACH.*When Good Components Go Bad: Formally Secure Compilation Despite Dynamic Compromise*, in "25th ACM Conference on Computer and Communications Security (CCS)", Toronto, Canada, ACM, October 2018, p. 1351–1368, https://arxiv.org/abs/1802.00588 [*DOI :* 10.1145/3243734.3243745], https://hal.archives-ouvertes.fr/hal-01949202

[3] K. BHARGAVAN, B. BLANCHET, N. KOBEISSI.*Verified Models and Reference Implementations for the TLS 1.3 Standard Candidate*, in "38th IEEE Symposium on Security and Privacy", San Jose, United States, May 2017, p. 483 - 502 [*DOI :* 10.1109/SP.2017.26], https://hal.inria.fr/hal-01575920

[4] K. BHARGAVAN, A. DELIGNAT-LAVAUD, C. FOURNET, A. PIRONTI, P.-Y. STRUB.*Triple Handshakes and Cookie Cutters: Breaking and Fixing Authentication over TLS*, in "IEEE Symposium on Security and Privacy (Oakland)", 2014, p. 98–113, https://hal.inria.fr/hal-01102259

[5] B. BLANCHET.*Modeling and Verifying Security Protocols with the Applied Pi Calculus and ProVerif*, in "Foundations and Trends in Privacy and Security", October 2016, vol. 1, n^o 1–2, p. 1–135, https://hal.inria.fr/hal-01423760

[6] M. ISAAKIDIS, H. HALPIN, G. DANEZIS.*UnlimitID: Privacy-Preserving Federated Identity Management Using Algebraic MACs*, in "Proceedings of the 2016 ACM on Workshop on Privacy in the Electronic Society", New York, NY, USA, WPES '16, ACM, 2016, p. 139–142 [*DOI :* 10.1145/2994620.2994637], https://hal.inria.fr/hal-01426847

[7] N. KOBEISSI, K. BHARGAVAN, B. BLANCHET.*Automated Verification for Secure Messaging Protocols and Their Implementations: A Symbolic and Computational Approach*, in "2nd IEEE European Symposium on Security and Privacy", Paris, France, April 2017, p. 435 - 450 [*DOI :* 10.1109/EUROSP.2017.38], https://hal.inria.fr/hal-01575923

[8] N. SWAMY, C. HRIŢCU, C. KELLER, A. RASTOGI, A. DELIGNAT-LAVAUD, S. FOREST, K. BHARGAVAN, C. FOURNET, P.-Y. STRUB, M. KOHLWEISS, J. K. ZINZINDOHOUÉ, S. ZANELLA-BÉGUELIN.*Dependent Types and Multi-Monadic Effects in F**, in "43rd ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL)", ACM, January 2016, p. 256-270, https://hal.inria.fr/hal-01265793

[9] J. K. ZINZINDOHOUÉ, K. BHARGAVAN, J. PROTZENKO, B. BEURDOUCHE.*HACL*: A Verified Modern Cryptographic Library*, in "Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017, Dallas, TX, USA, October 30 - November 03, 2017", 2017, p. 1789–1806, https://hal.inria.fr/hal-01588421

Publications of the year

Doctoral Dissertations and Habilitation Theses

[10] C. HRIŢCU.*The Quest for Formally Secure Compartmentalizing Compilation*, ENS Paris ; PSL Research University, January 2019, Habilitation à diriger des recherches, https://tel.archives-ouvertes.fr/tel-01995823

[11] K. MAILLARD.*Principles of Program Verification for Arbitrary Monadic Effects*, ENS Paris - Ecole Normale Supérieure de Paris, November 2019, https://hal.archives-ouvertes.fr/tel-02416788

Articles in International Peer-Reviewed Journal

[12] J. EREMONDI, É. TANTER, R. GARCIA.*Approximate normalization for gradual dependent types*, in "Proceedings of the ACM on Programming Languages", July 2019, vol. 3, n^o ICFP, p. 1-30 [*DOI :* 10.1145/3341692], https://hal.archives-ouvertes.fr/hal-02399594

[13] K. MAILLARD, D. AHMAN, R. ATKEY, G. MARTÍNEZ, C. HRIŢCU, E. RIVAS, É. TANTER.*Dijkstra monads for all*, in "Proceedings of the ACM on Programming Languages", July 2019, vol. 3, n^o ICFP, p. 1-29, https://arxiv.org/abs/1903.01237 [*DOI :* 10.1145/3341708], https://hal.archives-ouvertes.fr/hal-02398919

[14] K. MAILLARD, C. HRIŢCU, E. RIVAS, A. VAN MUYLDER.*The Next 700 Relational Program Logics*, in "Proceedings of the ACM on Programming Languages", 2019, vol. 4, n° POPL, https://arxiv.org/abs/1907.05244 , forthcoming, https://hal.archives-ouvertes.fr/hal-02398927

[15] P.-M. PÉDROT, N. TABAREAU, H. J. FEHRMANN, É. TANTER.*A Reasonably Exceptional Type Theory*, in "Proceedings of the ACM on Programming Languages", August 2019, vol. 3, p. 1-29 [*DOI :* 10.1145/3341712], https://hal.inria.fr/hal-02189128

[16] *Best Paper*
M. TORO, E. LABRADA, É. TANTER.*Gradual Parametricity, Revisited*, in "Proceedings of the ACM on Programming Languages", 2019, vol. 3, n° POPL, https://arxiv.org/abs/1807.04596 [*DOI :* 10.1145/3290330], https://hal.archives-ouvertes.fr/hal-01960553.

International Conferences with Proceedings

[17] *Best Paper*
C. ABATE, R. BLANCO, D. GARG, C. HRIŢCU, M. PATRIGNANI, J. THIBAULT.*Journey Beyond Full Abstraction: Exploring Robust Property Preservation for Secure Compilation*, in "CSF 2019 - 32nd IEEE Computer Security Foundations Symposium", Hoboken, United States, IEEE, June 2019, p. 256-271, https://arxiv.org/abs/1807.04603 [*DOI :* 10.1109/CSF.2019.00025], https://hal.archives-ouvertes.fr/hal-02398915.

[18] R. BLANCO, D. MILLER, A. MOMIGLIANO.*Property-Based Testing via Proof Reconstruction*, in "PPDP 2019 - 21st International Symposium on Principles and Practice of Programming Languages", Porto, Portugal, ACM Press, October 2019, p. 1-13 [*DOI :* 10.1145/3354166.3354170], https://hal.inria.fr/hal-02368931

[19] R. CRUZ, É. TANTER.*Polymorphic Relaxed Noninterference*, in "SecDev 2019 : IEEE Secure Development Conference", McLean, VA, United States, IEEE, 2019, p. 101-113 [*DOI :* 10.1109/SECDEV.2019.00021], https://hal.archives-ouvertes.fr/hal-02399576

[20] T. DÍAZ, F. OLMEDO, É. TANTER.*A Mechanized Formalization of GraphQL*, in "CPP 2020 - 9th ACM SIGPLAN International Conference on Certified Programs and Proofs", New Orleans, United States, January 2020 [*DOI :* 10.1145/3372885.3373822], https://hal.archives-ouvertes.fr/hal-02422532

[21] N. KOBEISSI, G. NICOLAS, K. BHARGAVAN.*Noise Explorer: Fully Automated Modeling and Verification for Arbitrary Noise Protocols*, in "EuroS&P 2019 - 4th IEEE European Symposium on Security and Privacy", Stockholm, Sweden, June 2019, https://hal.inria.fr/hal-01948964

[22] B. LIPP, B. BLANCHET, K. BHARGAVAN.*A Mechanised Cryptographic Proof of the WireGuard Virtual Private Network Protocol*, in "4th IEEE European Symposium on Security and Privacy", Stockholm, Sweden, IEEE Computer Society, June 2019, p. 231-246, https://hal.inria.fr/hal-02396640

[23] G. MARTÍNEZ, D. AHMAN, V. DUMITRESCU, N. GIANNARAKIS, C. HAWBLITZEL, C. HRIŢCU, M. NARASIMHAMURTHY, Z. PARASKEVOPOULOU, C. PIT-CLAUDEL, J. PROTZENKO, T. RAMANANANDRO, A. RASTOGI, N. SWAMY.*Meta-F*: Proof automation with SMT, Tactics, and Metaprograms*, in "ESOP'19 - European Symposium on Programming", Prague, Czech Republic, April 2019, https://arxiv.org/abs/1803.06547 , https://hal.archives-ouvertes.fr/hal-01995376

[24] J. PROTZENKO, B. BEURDOUCHE, D. MERIGOUX, K. BHARGAVAN.*Formally Verified Cryptographic Web Applications in WebAssembly*, in "SP 2019 - 40th IEEE Symposium on Security and Privacy", San Francisco, United States, IEEE, May 2019, p. 1256-1274 [*DOI :* 10.1109/SP.2019.00064], https://hal.inria.fr/hal-02294935

National Conferences with Proceeding

[25] D. MERIGOUX, R. MONAT, C. GAIE.*Étude formelle de l'implémentation du code des impôts*, in "31ème Journées Francophones des Langages Applicatifs", Gruissan, France, January 2020, https://hal.inria.fr/hal-02320347

Research Reports

[26] K. BHARGAVAN, B. BEURDOUCHE, P. NALDURG.*Formal Models and Verified Protocols for Group Messaging: Attacks and Proofs for IETF MLS*, Inria Paris, December 2019, https://hal.inria.fr/hal-02425229

[27] B. LIPP, B. BLANCHET, K. BHARGAVAN.*A Mechanised Cryptographic Proof of the WireGuard Virtual Private Network Protocol*, Inria Paris, April 2019, n^{o} RR-9269, 49, https://hal.inria.fr/hal-02100345

Other Publications

[28] B. BEURDOUCHE.*MLS Architecture: analysis of the security, privacy and functional requirements*, January 2020, working paper or preprint, https://hal.inria.fr/hal-02439526

[29] E. RIVAS, M. JASKELIOFF.*Monads with merging*, June 2019, working paper or preprint, https://hal.inria.fr/hal-02150199

References in notes

[30] M. ABADI, B. BLANCHET.*Analyzing Security Protocols with Secrecy Types and Logic Programs*, in "Journal of the ACM", January 2005, vol. 52, n^{o} 1, p. 102–146, http://prosecco.gforge.inria.fr/personal/bblanche/publications/AbadiBlanchetJACM7037.pdf

[31] M. ABADI, B. BLANCHET, C. FOURNET.*Just Fast Keying in the Pi Calculus*, in "ACM Transactions on Information and System Security (TISSEC)", July 2007, vol. 10, n^{o} 3, p. 1–59, http://prosecco.gforge.inria.fr/personal/bblanche/publications/AbadiBlanchetFournetTISSEC07.pdf

[32] C. ABATE, R. BLANCO, D. GARG, C. HRIŢCU, M. PATRIGNANI, J. THIBAULT.*Journey Beyond Full Abstraction: Exploring Robust Property Preservation for Secure Compilation*, in "32nd IEEE Computer Security Foundations Symposium (CSF)", IEEE, June 2019, p. 256-271 [*DOI :* 10.1109/CSF.2019.00025], https://arxiv.org/abs/1807.04603

[33] C. ABATE, A. AZEVEDO DE AMORIM, R. BLANCO, A. N. EVANS, G. FACHINI, C. HRIŢCU, T. LAURENT, B. C. PIERCE, M. STRONATI, A. TOLMACH.*When Good Components Go Bad: Formally Secure Compilation Despite Dynamic Compromise*, in "25th ACM Conference on Computer and Communications Security (CCS)", ACM, October 2018, p. 1351–1368, https://arxiv.org/abs/1802.00588

[34] D. AHMAN, C. HRIŢCU, K. MAILLARD, G. MARTÍNEZ, G. PLOTKIN, J. PROTZENKO, A. RASTOGI, N. SWAMY.*Dijkstra Monads for Free*, in "44th ACM SIGPLAN Symposium on Principles of Programming

Languages (POPL)", ACM, January 2017, p. 515-529 [*DOI :* 10.1145/3009837.3009878], https://www.fstar-lang.org/papers/dm4free/

[35] A. AZEVEDO DE AMORIM, M. DÉNÈS, N. GIANNARAKIS, C. HRITCU, B. C. PIERCE, A. SPECTOR-ZABUSKY, A. TOLMACH.*Micro-Policies: Formally Verified, Tag-Based Security Monitors*, in "36th IEEE Symposium on Security and Privacy (Oakland S&P)", IEEE Computer Society, May 2015, p. 813–830 [*DOI :* 10.1109/SP.2015.55], http://prosecco.gforge.inria.fr/personal/hritcu/publications/micro-policies.pdf

[36] K. BHARGAVAN, B. BLANCHET, N. KOBEISSI.*Verified Models and Reference Implementations for the TLS 1.3 Standard Candidate*, in "38th IEEE Symposium on Security and Privacy", San Jose, United States, May 2017, p. 483 - 502 [*DOI :* 10.1109/SP.2017.26], https://hal.inria.fr/hal-01575920

[37] K. BHARGAVAN, B. BOND, A. DELIGNAT-LAVAUD, C. FOURNET, C. HAWBLITZEL, C. HRIŢCU, S. ISHTIAQ, M. KOHLWEISS, R. LEINO, J. LORCH, K. MAILLARD, J. PAN, B. PARNO, J. PROTZENKO, T. RAMANANANDRO, A. RANE, A. RASTOGI, N. SWAMY, L. THOMPSON, P. WANG, S. ZANELLA-BÉGUELIN, J. K. ZINZINDOHOUÉ.*Everest: Towards a Verified, Drop-in Replacement of HTTPS*, in "2nd Summit on Advances in Programming Languages (SNAPL)", May 2017, http://drops.dagstuhl.de/opus/volltexte/2017/7119/pdf/LIPIcs-SNAPL-2017-1.pdf

[38] K. BHARGAVAN, A. DELIGNAT-LAVAUD, C. FOURNET, M. KOHLWEISS, J. PAN, J. PROTZENKO, A. RASTOGI, N. SWAMY, S. ZANELLA-BÉGUELIN, J. K. ZINZINDOHOUÉ.*Implementing and Proving the TLS 1.3 Record Layer*, in "IEEE Symposium on Security and Privacy (Oakland)", 2017

[39] K. BHARGAVAN, C. FOURNET, R. CORIN, E. ZALINESCU.*Verified Cryptographic Implementations for TLS*, in "ACM Transactions Inf. Syst. Secur.", March 2012, vol. 15, n^{o} 1, p. 3:1–3:32, http://doi.acm.org/10.1145/2133375.2133378

[40] K. BHARGAVAN, C. FOURNET, A. D. GORDON, N. SWAMY.*Verified implementations of the information card federated identity-management protocol*, in "ACM Symposium on Information, Computer and Communications Security (ASIACCS)", 2008, p. 123-135

[41] B. BLANCHET, M. ABADI, C. FOURNET.*Automated Verification of Selected Equivalences for Security Protocols*, in "Journal of Logic and Algebraic Programming", February–March 2008, vol. 75, n^{o} 1, p. 3–51, http://prosecco.gforge.inria.fr/personal/bblanche/publications/BlanchetAbadiFournetJLAP07.pdf

[42] B. BLANCHET.*An Efficient Cryptographic Protocol Verifier Based on Prolog Rules*, in "14th IEEE Computer Security Foundations Workshop (CSFW'01)", 2001, p. 82–96

[43] B. BLANCHET.*Automatic Verification of Correspondences for Security Protocols*, in "Journal of Computer Security", July 2009, vol. 17, n^{o} 4, p. 363–434, http://prosecco.gforge.inria.fr/personal/bblanche/publications/BlanchetJCS08.pdf

[44] B. BLANCHET, A. PODELSKI.*Verification of Cryptographic Protocols: Tagging Enforces Termination*, in "Theoretical Computer Science", March 2005, vol. 333, n^{o} 1-2, p. 67–90, Special issue FoSSaCS'03., http://prosecco.gforge.inria.fr/personal/bblanche/publications/BlanchetPodelskiTCS04.html

[45] D. CADÉ, B. BLANCHET.*Proved Generation of Implementations from Computationally Secure Protocol Specifications*, in "Journal of Computer Security", 2015, vol. 23, n^{o} 3, p. 331–402

[46] J. CLULOW.*On the Security of PKCS#11*, in "CHES", 2003, p. 411-425

[47] S. DELAUNE, S. KREMER, G. STEEL.*Formal Analysis of PKCS#11 and Proprietary Extensions*, in "Journal of Computer Security", November 2010, vol. 18, n^o 6, p. 1211-1245 [*DOI :* 10.3233/JCS-2009-0394], http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/DKS-jcs09.pdf

[48] A. DELIGNAT-LAVAUD, K. BHARGAVAN, S. MAFFEIS.*Language-Based Defenses Against Untrusted Browser Origins*, in "Proceedings of the 22th USENIX Security Symposium", 2013, http://prosecco.inria.fr/personal/karthik/pubs/language-based-defenses-against-untrusted-origins-sec13.pdf

[49] D. DOLEV, A. YAO.*On the security of public key protocols*, in "IEEE Transactions on Information Theory", 1983, vol. IT–29, n^o 2, p. 198–208

[50] C. FOURNET, M. KOHLWEISS, P.-Y. STRUB.*Modular Code-Based Cryptographic Verification*, in "ACM Conference on Computer and Communications Security", 2011

[51] N. KOBEISSI, K. BHARGAVAN, B. BLANCHET.*Automated Verification for Secure Messaging Protocols and Their Implementations: A Symbolic and Computational Approach*, in "2nd IEEE European Symposium on Security and Privacy", Paris, France, April 2017, p. 435 - 450 [*DOI :* 10.1109/EUROSP.2017.38], https://hal.inria.fr/hal-01575923

[52] K. MAILLARD, D. AHMAN, R. ATKEY, G. MARTÍNEZ, C. HRIŢCU, E. RIVAS, É. TANTER.*Dijkstra Monads for All*, in "PACMPL", 2019, vol. 3, n^o ICFP, p. 104:1–104:29 [*DOI :* 10.1145/3341708], https://arxiv.org/abs/1903.01237

[53] R. NEEDHAM, M. SCHROEDER.*Using encryption for authentication in large networks of computers*, in "Communications of the ACM", 1978, vol. 21, n^o 12, p. 993–999

[54] J. PROTZENKO, J. K. ZINZINDOHOUÉ, A. RASTOGI, T. RAMANANANDRO, P. WANG, S. ZANELLA-BÉGUELIN, A. DELIGNAT-LAVAUD, C. HRIŢCU, K. BHARGAVAN, C. FOURNET, N. SWAMY.*Verified Low-Level Programming Embedded in F**, in "PACMPL", September 2017, vol. 1, n^o ICFP, p. 17:1–17:29 [*DOI :* 10.1145/3110261], http://arxiv.org/abs/1703.00053

[55] T. RAMANANANDRO, A. DELIGNAT-LAVAUD, C. FOURNET, N. SWAMY, T. CHAJED, N. KOBEISSI, J. PROTZENKO.*EverParse: Verified Secure Zero-Copy Parsers for Authenticated Message Formats*, in "28th USENIX Security Symposium, USENIX Security 2019, Santa Clara, CA, USA, August 14-16, 2019", N. HENINGER, P. TRAYNOR (editors), USENIX Association, 2019, p. 1465–1482, https://www.usenix.org/conference/usenixsecurity19/presentation/delignat-lavaud

[56] N. SWAMY, C. FOURNET, A. RASTOGI, K. BHARGAVAN, J. CHEN, P.-Y. STRUB, G. M. BIERMAN.*Gradual typing embedded securely in JavaScript*, in "41st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL)", 2014, p. 425-438, http://prosecco.inria.fr/personal/karthik/pubs/tsstar-popl14.pdf

[57] N. SWAMY, C. HRIŢCU, C. KELLER, A. RASTOGI, A. DELIGNAT-LAVAUD, S. FOREST, K. BHARGAVAN, C. FOURNET, P.-Y. STRUB, M. KOHLWEISS, J. K. ZINZINDOHOUÉ, S. ZANELLA-BÉGUELIN.*Dependent Types and Multi-Monadic Effects in F**, in "43rd ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL)", ACM, January 2016, p. 256-270, https://www.fstar-lang.org/papers/mumon/

[58] J. K. ZINZINDOHOUÉ, K. BHARGAVAN, J. PROTZENKO, B. BEURDOUCHE.*HACL*: A Verified Modern Cryptographic Library*, in "Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017, Dallas, TX, USA, October 30 - November 03, 2017", 2017, p. 1789–1806, http://doi.acm.org/10.1145/3133956.3134043

Project-Team QUANTIC

QUANTum Information Circuits

IN COLLABORATION WITH: Centre Automatique et Systèmes, Laboratoire Pierre Aigrain

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Paris

Mines ParisTech

Sorbonne Université (UPMC)

RESEARCH CENTER
Paris

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team QUANTIC

Creation of the Team: 2013 September 12, updated into Project-Team: 2015 April 01

Keywords:

Computer Science and Digital Science:

A1.1.11. - Quantum architectures
A4.2. - Correcting codes
A6. - Modeling, simulation and control
A6.1. - Methods in mathematical modeling
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling
A6.1.3. - Discrete Modeling (multi-agent, people centered)
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
A6.3.1. - Inverse problems
A6.3.2. - Data assimilation
A6.3.3. - Data processing
A6.3.4. - Model reduction
A6.4. - Automatic control
A6.4.1. - Deterministic control
A6.4.2. - Stochastic control
A6.4.3. - Observability and Controlability
A6.4.4. - Stability and Stabilization

Other Research Topics and Application Domains:

B5.3. - Nanotechnology
B5.4. - Microelectronics
B6.5. - Information systems
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

Mazyar Mirrahimi [Team leader, Inria, Senior Researcher]
Philippe Campagne Ibarcq [Inria, Researcher, from Oct 2019]
Alain Sarlette [Inria, Researcher]

Faculty Members

Zaki Leghtas [Ecole Nationale Supérieure des Mines de Paris , Associate Professor]
Pierre Rouchon [Ecole Nationale Supérieure des Mines de Paris , Professor, HDR]

External Collaborator

Michel Sorine [Inria, HDR]

PhD Students

Michiel Burgelman [Inria, PhD Student]
Gerardo Cardona Sanchez [Ecole Nationale Supérieure des Mines de Paris , PhD Student, until Nov 2019]
Jeremie Guillaud [Inria, PhD Student]
Vincent Martin [Inria, PhD Student]
Christian Siegele [Inria, PhD Student, from Sep 2019]
Lucas Verney [Ecole Normale Supérieure Paris, PhD Student, until Aug 2019]

Post-Doctoral Fellows

Paolo Forni [Ecole Nationale Supérieure des Mines de Paris , Post-Doctoral Fellow, until May 2019]
William Smith [Inria, Post-Doctoral Fellow]

Administrative Assistant

Derya Gok [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Overall objectives

The research activities of QUANTIC team lie at the border between theoretical and experimental efforts in the emerging field of quantum systems engineering. Our research topics are in direct continuation of a historic research theme of Inria, classical automatic control, while opening completely new perspectives toward quantum control: by developing a new mathematical system theory for quantum circuits, we will realize the components of a future quantum information processing unit.

One of the unique features of our team concerns the large spectrum of our subjects going from the mathematical analysis of the physical systems (development of systematic mathematical methods for control and estimation of quantum systems), and the numerical analysis of the proposed solutions, to the experimental implementation of the quantum circuits based on these solutions. This is made possible by the constant and profound interaction between the applied mathematicians and the physicists in the group. Indeed, this close collaboration has already brought a significant acceleration in our research efforts. In a long run, this synergy should lead to a deeper understanding of the physical phenomena behind these emerging technologies and the development of new research directions within the field of quantum information processing.

Towards this ultimate task of practical quantum digital systems, the approach of the QUANTIC team is complementary to the one taken by teams with expertise in quantum algorithms. Indeed, we start from the specific controls that can be realistically applied on physical systems, to propose designs which combine them into *hardware shortcuts* implementing *robust* behaviors useful for quantum information processing. Whenever a significant new element of quantum engineering architecture is developed, the initial motivation is to prove an enabling technology with major impact for the groups working one abstraction layer higher: on quantum algorithms but also on e.g. secure communication and metrology applications.

3. Research Program

3.1. Hardware-efficient quantum information processing

In this scientific program, we will explore various theoretical and experimental issues concerning protection and manipulation of quantum information. Indeed, the next, critical stage in the development of Quantum Information Processing (QIP) is most certainly the active quantum error correction (QEC). Through this stage one designs, possibly using many physical qubits, an encoded logical qubit which is protected against major decoherence channels and hence admits a significantly longer effective coherence time than a physical qubit. Reliable (fault-tolerant) computation with protected logical qubits usually comes at the expense of a significant overhead in the hardware (up to thousands of physical qubits per logical qubit). Each of the involved physical qubits still needs to satisfy the best achievable properties (coherence times, coupling strengths and tunability). More remarkably, one needs to avoid undesired interactions between various subsystems. This is going to be a major difficulty for qubits on a single chip.

The usual approach for the realization of QEC is to use many qubits to obtain a larger Hilbert space of the qubit register [88], [92]. By redundantly encoding quantum information in this Hilbert space of larger dimension one make the QEC tractable: different error channels lead to distinguishable error syndromes. There are two major drawbacks in using multi-qubit registers. The first, fundamental, drawback is that with each added physical qubit, several new decoherence channels are added. Because of the exponential increase of the Hilbert's space dimension versus the linear increase in the number of decay channels, using enough qubits, one is able to eventually protect quantum information against decoherence. However, multiplying the number of possible errors, this requires measuring more error syndromes. Note furthermore that, in general, some of these new decoherence channels can lead to correlated action on many qubits and this needs to be taken into account with extra care: in particular, such kind of non-local error channels are problematic for surface codes. The second, more practical, drawback is that it is still extremely challenging to build a register of more than on the order of 10 qubits where each of the qubits is required to satisfy near the best achieved properties: these properties include the coherence time, the coupling strengths and the tunability. Indeed, building such a register is not merely only a fabrication task but rather, one requirers to look for architectures such that, each individual qubit can be addressed and controlled independently from the others. One is also required to make sure that all the noise channels are well-controlled and uncorrelated for the QEC to be effective.

We have recently introduced a new paradigm for encoding and protecting quantum information in a quantum harmonic oscillator (e.g. a high-Q mode of a 3D superconducting cavity) instead of a multi-qubit register [64]. The infinite dimensional Hilbert space of such a system can be used to redundantly encode quantum information. The power of this idea lies in the fact that the dominant decoherence channel in a cavity is photon damping, and no more decay channels are added if we increase the number of photons we insert in the cavity. Hence, only a single error syndrome needs to be measured to identify if an error has occurred or not. Indeed, we are convinced that most early proposals on continuous variable QIP [61], [55] could be revisited taking into account the design flexibilities of Quantum Superconducting Circuits (QSC) and the new coupling regimes that are provided by these systems. In particular, we have illustrated that coupling a qubit to the cavity mode in the strong dispersive regime provides an important controllability over the Hilbert space of the cavity mode [63]. Through a recent experimental work [97], we benefit from this controllability to prepare superpositions of quasi-orthogonal coherent states, also known as Schrödinger cat states.

In this Scheme, the logical qubit is encoded in a four-component Schrödinger cat state. Continuous quantum non-demolition (QND) monitoring of a single physical observable, consisting of photon number parity, enables then the tractability of single photon jumps. We obtain therefore a first-order quantum error correcting code using only a single high-Q cavity mode (for the storage of quantum information), a single qubit (providing the non-linearity needed for controllability) and a single low-Q cavity mode (for reading out the error syndrome). An earlier experiment on such QND photon-number parity measurements [93] has recently led to a first experimental realization of a full quantum error correcting code improving the coherence time of quantum information [8]. As shown in Figure 1, this leads to a significant hardware economy for realization of a protected logical qubit. Our goal here is to push these ideas towards a reliable and hardware-efficient paradigm for universal quantum computation.

3.2. Reservoir (dissipation) engineering and autonomous stabilization of quantum systems

Being at the heart of any QEC protocol, the concept of feedback is central for the protection of quantum information, enabling many-qubit quantum computation or long-distance quantum communication. However, such a closed-loop control which requires a real-time and continuous measurement of the quantum system has been for long considered as counter-intuitive or even impossible. This thought was mainly caused by properties of quantum measurements: any measurement implies an instantaneous strong perturbation to the system's state. The concept of *quantum non-demolition* (QND) measurement has played a crucial role in understanding and resolving this difficulty [37]. In the context of cavity quantum electro-dynamics (cavity QED) with Rydberg atoms [57], a first experiment on continuous QND measurements of the number of microwave photons was performed by the group at Laboratoire Kastler-Brossel (ENS) [56]. Later on, this

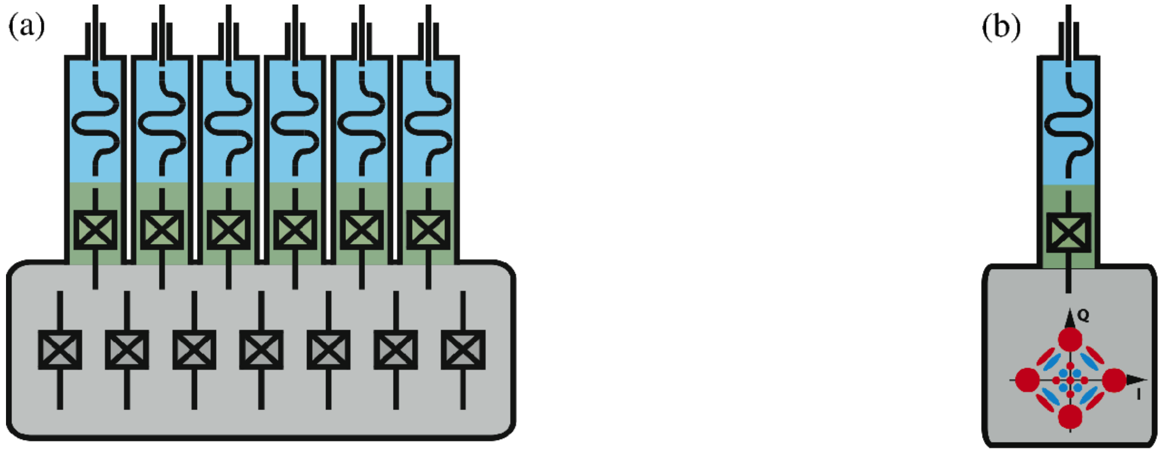

Figure 1. (a) A protected logical qubit consisting of a register of many qubits: here, we see a possible architecture for the Steane code [92] consisting of 7 qubits requiring the measurement of 6 error syndromes. In this sketch, 7 transmon qubits in a high-Q resonator and the measurement of the 6 error syndromes is ensured through 6 additional ancillary qubits with the possibility of individual readout of the ancillary qubits via independent low-Q resonators. (b) Minimal architecture for a protected logical qubit, adapted to circuit quantum electrodynamics experiments. Quantum information is encoded in a Schrödinger cat state of a single high-Q resonator mode and a single error syndrome is measured, using a single ancillary transmon qubit and the associated readout low-Q resonator.

ability of performing continuous measurements allowed the same group to realize the first continuous quantum feedback protocol stabilizing highly non-classical states of the microwave field in the cavity, the so-called photon number states [10] (this ground-breaking work was mentioned in the Nobel prize attributed to Serge Haroche). The QUANTIC team contributed to the theoretical work behind this experiment [47], [29], [91] [1]. These contributions include the development and optimization of the quantum filters taking into account the quantum measurement back-action and various measurement noises and uncertainties, the development of a feedback law based on control Lyapunov techniques, and the compensation of the feedback delay.

In the context of circuit quantum electrodynamics (circuit QED) [45], recent advances in quantum-limited amplifiers [81], [95] have opened doors to high-fidelity non-demolition measurements and real-time feedback for superconducting qubits [58]. This ability to perform high-fidelity non-demolition measurements of a quantum signal has very recently led to quantum feedback experiments with quantum superconducting circuits [95], [80], [39]. Here again, the QUANTIC team has participated to one of the first experiments in the field where the control objective is to track a dynamical trajectory of a single qubit rather than stabilizing a stationary state. Such quantum trajectory tracking could be further explored to achieve metrological goals such as the stabilization of the amplitude of a microwave drive [71].

While all this progress has led to a strong optimism about the possibility to perform active protection of quantum information against decoherence, the rather short dynamical time scales of these systems limit, to a great amount, the complexity of the feedback strategies that could be employed. Indeed, in such measurement-based feedback protocols, the time-consuming data acquisition and post-treatment of the output signal leads to an important latency in the feedback procedure.

The reservoir (dissipation) engineering [78] and the closely related coherent feedback [69] are considered as alternative approaches circumventing the necessity of a real-time data acquisition, signal processing and feedback calculations. In the context of quantum information, the decoherence, caused by the coupling of a system to uncontrolled external degrees of freedom, is generally considered as the main obstacle to synthesize quantum states and to observe quantum effects. Paradoxically, it is possible to intentionally engineer a particular coupling to a reservoir in the aim of maintaining the coherence of some particular quantum states. In a general viewpoint, these approaches could be understood in the following manner: by coupling the quantum system to be stabilized to a strongly dissipative ancillary quantum system, one evacuates the entropy of the main system through the dissipation of the ancillary one. By building the feedback loop into the Hamiltonian, this type of autonomous feedback obviates the need for a complicated external control loop to correct errors. On the experimental side, such autonomous feedback techniques have been used for qubit reset [54], single-qubit state stabilization [73], and the creation [32] and stabilization [62], [68], [87] of states of multipartite quantum systems.

Such reservoir engineering techniques could be widely revisited exploring the flexibility in the Hamiltonian design for QSC. We have recently developed theoretical proposals leading to extremely efficient, and simple to implement, stabilization schemes for systems consisting of a single, two or three qubits [54], [66], [43], [46]. The experimental results based on these protocols have illustrated the efficiency of the approach [54], [87]. Through these experiments, we exploit the strong dispersive interaction [85] between superconducting qubits and a single low-Q cavity mode playing the role of a dissipative reservoir. Applying continuous-wave (cw) microwave drives with well-chosen fixed frequencies, amplitudes, and phases, we engineer an effective interaction Hamiltonian which evacuates the entropy of the system interacting with a noisy environment: by driving the qubits and cavity with continuous-wave drives, we induce an autonomous feedback loop which corrects the state of the qubits every time it decays out of the desired target state. The schemes are robust against small variations of the control parameters (drives amplitudes and phase) and require only some basic calibration. Finally, by avoiding resonant interactions between the qubits and the low-Q cavity mode, the qubits remain protected against the Purcell effect, which would reduce the coherence times. We have also investigated both theoretically and experimentally the autonomous stabilization of non-classical states (such as Schrodinger cat states and Fock states) of microwave field confined in a high-Q cavity mode [83], [59][7], [6].

3.3. System theory for quantum information processing

In parallel and in strong interactions with the above experimental goals, we develop systematic mathematical methods for dynamical analysis, control and estimation of composite and open quantum systems. These systems are built with several quantum subsystems whose irreversible dynamics results from measurements and/or decoherence. A special attention is given to spin/spring systems made with qubits and harmonic oscillators. These developments are done in the spirit of our recent contributions [82], [29], [90], [84], [91][9], [1] resulting from collaborations with the cavity quantum electrodynamics group of Laboratoire Kastler Brossel.

3.3.1. Stabilization by measurement-based feedback

The protection of quantum information via efficient QEC is a combination of (i) tailored dynamics of a quantum system in order to protect an informational qubit from certain decoherence channels, and (ii) controlled reaction to measurements that efficiently detect and correct the dominating disturbances that are not rejected by the tailored quantum dynamics.

In such feedback scheme, the system and its measurement are quantum objects whereas the controller and the control input are classical. The stabilizing control law is based on the past values of the measurement outcomes. During our work on the LKB photon box, we have developed, for single input systems subject to quantum non-demolition measurement, a systematic stabilization method [1]: it is based on a discrete-time formulation of the dynamics, on the construction of a strict control Lyapunov function and on an explicit compensation of the feedback-loop delay. Keeping the QND measurement assumptions, extensions of such stabilization schemes will be investigated in the following directions: finite set of values for the control input with application to the convergence analysis of the atomic feedback scheme experimentally tested in [98]; multi-input case where the construction by inversion of a Metzler matrix of the strict Lyapunov function is not straightforward; continuous-time systems governed by diffusive master equations; stabilization towards a set of density operators included in a target subspace; adaptive measurement by feedback to accelerate the convergence towards a stationary state as experimentally tested in [76]. Without the QND measurement assumptions, we will also address the stabilization of non-stationary states and trajectory tracking, with applications to systems similar to those considered in [58], [39].

3.3.2. Filtering, quantum state and parameter estimations

The performance of every feedback controller crucially depends on its online estimation of the current situation. This becomes even more important for quantum systems, where full state measurements are physically impossible. Therefore the ultimate performance of feedback correction depends on fast, efficient and optimally accurate state and parameter estimations.

A quantum filter takes into account imperfection and decoherence and provides the quantum state at time $t \geq 0$ from an initial value at $t = 0$ and the measurement outcomes between 0 and t. Quantum filtering goes back to the work of Belavkin [33] and is related to quantum trajectories [41], [44]. A modern and mathematical exposure of the diffusive models is given in [31]. In [99] a first convergence analysis of diffusive filters is proposed. Nevertheless the convergence characterization and estimation of convergence rate remain open and difficult problems. For discrete time filters, a general stability result based on fidelity is proven in [82], [90]. This stability result is extended to a large class of continuous-time filters in [30]. Further efforts are required to characterize asymptotic and exponential stability. Estimations of convergence rates are available only for quantum non-demolition measurements [34]. Parameter estimations based on measurement data of quantum trajectories can be formulated within such quantum filtering framework [49], [74].

We will continue to investigate stability and convergence of quantum filtering. We will also exploit our fidelity-based stability result to justify maximum likelihood estimation and to propose, for open quantum system, parameter estimation algorithms inspired of existing estimation algorithms for classical systems. We will also investigate a more specific quantum approach: it is noticed in [38] that post-selection statistics and "past quantum" state analysis [50] enhance sensitivity to parameters and could be interesting towards increasing the precision of an estimation.

3.3.3. Stabilization by interconnections

In such stabilization schemes, the controller is also a quantum object: it is coupled to the system of interest and is subject to decoherence and thus admits an irreversible evolution. These stabilization schemes are closely related to reservoir engineering and coherent feedback [78], [69]. The closed-loop system is then a composite system built with the original system and its controller. In fact, and given our particular recent expertise in this domain [9] [87], [54], this subsection is dedicated to further developing such stabilization techniques, both experimentally and theoretically.

The main analysis issues are to prove the closed-loop convergence and to estimate the convergence rates. Since these systems are governed by Lindblad differential equations (continuous-time case) or Kraus maps (discrete-time case), their stability is automatically guaranteed: such dynamics are contractions for a large set of metrics (see [77]). Convergence and asymptotic stability is less well understood. In particular most of the convergence results consider the case where the target steady-state is a density operator of maximum rank (see, e.g., [28][chapter 4, section 6]). When the goal steady-state is not full rank very few convergence results are available.

We will focus on this geometric situation where the goal steady-state is on the boundary of the cone of positive Hermitian operators of finite trace. A specific attention will be given to adapt standard tools (Lyapunov function, passivity, contraction and Lasalle's invariance principle) for infinite dimensional systems to spin/spring structures inspired of [9], [7] [87], [54] and their associated Fokker-Planck equations for the Wigner functions.

We will also explore the Heisenberg point of view in connection with recent results of the Inria project-team MAXPLUS (algorithms and applications of algebras of max-plus type) relative to Perron-Frobenius theory [53], [52]. We will start with [86] and [79] where, based on a theorem due to Birkhoff [35], dual

Lindblad equations and dual Kraus maps governing the Heisenberg evolution of any operator are shown to be contractions on the cone of Hermitian operators equipped with Hilbert's projective metric. As the Heisenberg picture is characterized by convergence of all operators to a multiple of the identity, it might provide a mean to circumvent the rank issues. We hope that such contraction tools will be especially well adapted to analyzing quantum systems composed of multiple components, motivated by the facts that the same geometry describes the contraction of classical systems undergoing synchronizing interactions [94] and by our recent generalized extension of the latter synchronizing interactions to quantum systems [70].

Besides these analysis tasks, the major challenge in stabilization by interconnections is to provide systematic methods for the design, from typical building blocks, of control systems that stabilize a specific quantum goal (state, set of states, operation) when coupled to the target system. While constructions exist for so-called linear quantum systems [75], this does not cover the states that are more interesting for quantum applications. Various strategies have been proposed that concatenate iterative control steps for open-loop steering [96], [67] with experimental limitations. The characterization of Kraus maps to stabilize any types of states has also been established [36], but without considering experimental implementations. A viable stabilization by interaction has to combine the capabilities of these various approaches, and this is a missing piece that we want to address.

3.3.3.1. Perturbation methods

With this subsection we turn towards more fundamental developments that are necessary in order to address the complexity of quantum networks with efficient reduction techniques. This should yield both efficient mathematical methods, as well as insights towards unravelling dominant physical phenomena/mechanisms in multipartite quantum dynamical systems.

In the Schrödinger point of view, the dynamics of open quantum systems are governed by master equations, either deterministic or stochastic [57], [51]. Dynamical models of composite systems are based on tensor products of Hilbert spaces and operators attached to the constitutive subsystems. Generally, a hierarchy of different timescales is present. Perturbation techniques can be very useful to construct reliable models adapted to the timescale of interest.

To eliminate high frequency oscillations possibly induced by quasi-resonant classical drives, averaging techniques are used (rotating wave approximation). These techniques are well established for closed systems without any dissipation nor irreversible effect due to measurement or decoherence. We will consider in a first step the adaptation of these averaging techniques to deterministic Lindblad master equations governing the quantum state, i.e. the system density operator. Emphasis will be put on first order and higher order corrections based on non-commutative computations with the different operators appearing in the Lindblad equations. Higher order terms could be of some interest for the protected logical qubit of figure 1b. In future steps, we intend to explore the possibility to explicitly exploit averaging or singular perturbation properties in the design of coherent quantum feedback systems; this should be an open-systems counterpart of works like [65].

To eliminate subsystems subject to fast convergence induced by decoherence, singular perturbation techniques can be used. They provide reduced models of smaller dimension via the adiabatic elimination of the rapidly converging subsystems. The derivation of the slow dynamics is far from being obvious (see, e.g., the computations of page 142 in [40] for the adiabatic elimination of low-Q cavity). Conversely to the classical composite systems where we have to eliminate one component in a Cartesian product, we here have to eliminate one component in a tensor product. We will adapt geometric singular perturbations [48] and invariant manifold techniques [42] to such tensor product computations to derive reduced slow approximations of any order. Such adaptations will be very useful in the context of quantum Zeno dynamics to obtain approximations of the slow dynamics on the decoherence-free subspace corresponding to the slow attractive manifold.

Perturbation methods are also precious to analyze convergence rates. Deriving the spectrum attached to the Lindblad differential equation is not obvious. We will focus on the situation where the decoherence terms of the form $L\rho L^\dagger - (L^\dagger L\rho + \rho L^\dagger L)/2$ are small compared to the conservative terms $-i[H/\hbar, \rho]$. The difficulty to overcome here is the degeneracy of the unperturbed spectrum attached to the conservative evolution $\frac{d}{dt}\rho = -i[H/\hbar, \rho]$. The degree of degeneracy of the zero eigenvalue always exceeds the dimension of the Hilbert space. Adaptations of usual perturbation techniques [60] will be investigated. They will provide estimates of convergence rates for slightly open quantum systems. We expect that such estimates will help to understand the dependence on the experimental parameters of the convergence rates observed in [54], [87], [66].

As particular outcomes for the other subsections, we expect that these developments towards simpler dominant dynamics will guide the search for optimal control strategies, both in open-loop microwave networks and in autonomous stabilization schemes such as reservoir engineering. It will further help to efficiently compute explicit convergence rates and quantitative performances for all the intended experiments.

4. Application Domains

4.1. Quantum engineering

A new field of quantum systems engineering has emerged during the last few decades. This field englobes a wide range of applications including nano-electromechanical devices, nuclear magnetic resonance applications, quantum chemical synthesis, high resolution measurement devices and finally quantum information processing devices for implementing quantum computation and quantum communication. Recent theoretical and experimental achievements have shown that the quantum dynamics can be studied within the framework of estimation and control theory, but give rise to new models that have not been fully explored yet.

The QUANTIC team's activities are defined at the border between theoretical and experimental efforts of this emerging field with an emphasis on the applications in quantum information, computation and communication. The main objective of this interdisciplinary team is to develop quantum devices ensuring a robust processing of quantum information.

On the theory side, this is done by following a system theory approach: we develop estimation and control tools adapted to particular features of quantum systems. The most important features, requiring the development of new engineering methods, are related to the concept of measurement and feedback for composite quantum systems. The destructive and partial [0] nature of measurements for quantum systems lead to major difficulties in extending classical control theory tools. Indeed, design of appropriate measurement protocols and, in the sequel, the corresponding quantum filters estimating the state of the system from the partial measurement record, are themselves building blocks of the quantum system theory to be developed.

On the experimental side, we develop new quantum information processing devices based on quantum superconducting circuits. Indeed, by realizing superconducting circuits at low temperatures and using microwave measurement techniques, the macroscopic and collective degrees of freedom such as the voltage and the current are forced to behave according to the laws of quantum mechanics. Our quantum devices are aimed to protect and process quantum information through these integrated circuits.

5. Highlights of the Year

5.1. Highlights of the Year

- Zaki Leghtas has obtained an ERC starting grant in pannel PE3 entitled ECLIPSE (Exotic superconducting CIrcuits to probe and protect quantum States of light and mattEr).
- Our team (Zaki Leghtas and Mazyar Mirrahimi) has obtained a european QUANTERA grant entitled QuCOS (Quantum Computation with Schrödinger cat states).
- Philippe Campagne-Ibarcq was hired as a CRCN Inria in the QUANTIC team.
- Zaki Leghtas was an invited speaker of American Physical Society March Meeting in Boston, USA.

[0] Here the partiality means that no single quantum measurement is capable of providing the complete information on the state of the system.

- Pierre Rouchon was a semi-plenary speaker at the IFAC Mechatronics and NOLCOS conference, September 4-6, Vienna, Austria.
- Successful PhD defense of Gerardo Cardona, under the direction of P. Rouchon and A. Sarlette.
- Successful PhD defense of Lucas Verney, under the direction of M. Mirrahimi and Z. Leghtas.
- Successful PhD defense of Zhifei Zhang, under the direction of A. Sarlette at Ghent University.

6. New Results

6.1. Highly coherent spin states in carbon nanotubes coupled to cavity photons

Participants: Zaki Leghtas

Spins confined in quantum dots are considered as a promising platform for quantum information processing. While many advanced quantum operations have been demonstrated, experimental as well as theoretical efforts are now focusing on the development of scalable spin quantum bit architectures. One particularly promising method relies on the coupling of spin quantum bits to microwave cavity photons. This would enable the coupling of distant spins via the exchange of virtual photons for two qubit gate applications, which still remains to be demonstrated with spin qubits. Here, we use a circuit QED spin-photon interface to drive a single electronic spin in a carbon nanotube based double quantum dot using cavity photons. The microwave spectroscopy allows us to identify an electrically controlled spin transition with a decoherence rate which can be tuned to be as low as 250kHz. We show that this value is consistent with the expected hyperfine coupling in carbon nanotubes. These coherence properties, which can be attributed to the use of pristine carbon nanotubes stapled inside the cavity, should enable coherent spin-spin interaction via cavity photons and compare favourably to the ones recently demonstrated in Si-based circuit QED experiments. This experimental result is a collaboration between Zaki Leghtas (QUANTIC) and the group of Takis Kontos at ENS and was published in [13].

6.2. Escape of a Driven Quantum Josephson Circuit into Unconfined States

Participants: Raphaël Lescanne, Zaki Leghtas, Mazyar Mirrahimi and Lucas Verney

Josephson circuits have been ideal systems to study complex nonlinear dynamics that can lead to chaotic behavior and instabilities. More recently, Josephson circuits in the quantum regime, particularly in the presence of microwave drives, have demonstrated their ability to emulate a variety of Hamiltonians that are useful for the processing of quantum information. In this work, we show that these drives lead to an instability that results in the escape of the circuit mode into states that are not confined by the Josephson cosine potential. We observe this escape in a ubiquitous circuit: a transmon embedded in a 3D cavity. When the transmon occupies these free-particle-like states, the circuit behaves as though the junction had been removed and all nonlinearities are lost. This work deepens our understanding of strongly driven Josephson circuits, which is important for fundamental and application perspectives, such as the engineering of Hamiltonians by parametric pumping. This experimental work published in [17] demonstrates elements of the theory derived by [22].

6.3. Structural Instability of Driven Josephson Circuits Prevented by an Inductive Shunt

Participants: Raphaël Lescanne, Zaki Leghtas, Mazyar Mirrahimi and Lucas Verney

Superconducting circuits are a versatile platform to implement a multitude of Hamiltonians that perform quantum computation, simulation, and sensing tasks. A key ingredient for realizing a desired Hamiltonian is the irradiation of the circuit by a strong drive. These strong drives provide an in situ control of couplings, which cannot be obtained by near-equilibrium Hamiltonians. However, as shown in this theoretical study, out-of-equilibrium systems are easily plagued by complex dynamics, leading to instabilities. The prediction and prevention of these instabilities is crucial, both from a fundamental and application perspective. We propose an inductively shunted transmon as the elementary circuit optimized for strong parametric drives. Developing a numerical approach that avoids the built-in limitations of perturbative analysis, we demonstrate that adding the inductive shunt significantly extends the range of pump powers over which the circuit behaves in a stable manner. This theoretical result was published in [22] and analyzes the experiment [17].

6.4. Fast and virtually exact quantum gate generation in $U(n)$ via Iterative Lyapunov Methods

Participants: Pierre Rouchon

This work presents an iterative algorithm published in [20] and named RIGA for Reference Input Generation Algorithm. This algorithm constructs smooth control pulses for quantum gate preparations of closed quantum systems. It combines right translation invariance and Lyapunov trajectory tracking. It exhibits exponential convergence when the system is controllable. It can be seen as a closed-loop version of the widely used GRAPE algorithm. Two numerical case-studies borrowed from the recent literature are addressed. The first one is relative to a system of 10 coupled qubits with local controls. The second one considers a C-NOT gate generation involving the lower levels of two coupled nonlinear cavities (transmon-qubits).

6.5. Benchmarking maximum-likelihood state estimation with an entangled two-cavity state

Participants: Pierre Rouchon

The efficient quantum state reconstruction algorithm described in the PhD of Pierre Six, a former student of the Quantic team, (see [89]) is experimentally implemented on the non-local state of two microwave cavities entangled by a circular Rydberg atom. In [19], we use information provided by long sequences of measurements performed by resonant and dispersive probe atoms over time scales involving the system decoherence. Moreover, we benefit from the consolidation, in the same reconstruction, of different measurement protocols providing complementary information. Finally, we obtain realistic error bars for the matrix elements of the reconstructed density operator. These results demonstrate the pertinence and precision of the method, directly applicable to any complex quantum system.

6.6. Towards tight impossibility and possibility results for string stability

Participants: Alain Sarlette

This is the last step of the PhD thesis of Arash Farnam under the direction of A. Sarlette at Ghent University. The aim was to study which elements are really essential in so-called “string instability” results, which exist in several variants and with several assumptions. In [15], we have significantly extended the often frequency-based linear approach, by showing how the assumptions lead to string instability also in any nonlinear systems with reasonable bandwidth. This should allow to clarify that how the problem setting must be adapted in order to obtain more positive results. In [14], we show how other elements do not help, and we clarify how the knowledge of individual vehicles’ absolute velocity is key to enable strong versions of string stability, in conjunction with PID control. Previous studies had only considered weaker versions, with PD type control. A more detailed study of string stability with absolute velocity control has also been published in [25].

6.7. Stabilization of quantum systems under continuous non-demolition measurements

Participants: Gerardo Cardona, Alain Sarlette and Pierre Rouchon

The stabilization of quantum states or quantum subspaces using feedback signals from quantum non-demolition measurements is a basic control task; in discrete-time, this is the fundamental control property shown in the first quantum feedback experiment by Serge Haroche. In continuous-time, the problem is harder. So-called Markovian feedback can stabilize some states, but in particular the quantum non-demolition eigenstates which would be marginally stable under measurements, cannot be stabilized asymptotically with this technique. Stochastic control techniques, based on feedback from a full state estimator, have been proposed and analyzed to stabilize such eigenstates, proving convergence but not much more. In [12], we prove how a relatively simple controller, feeding back Wiener noise with a gain that depends on eigenstate populations, allows to exponentially stabilize the target eigenstate. This generalizes our previous results about the qubit. In [24], we provide a similar scheme and convergence proof for stabilizing an invariant subspace of the measurement, namely the codespace of a repetition code for quantum error correction. To the best of our knowledge there was no convergence proof so far for stabilizing such subspaces on the basis of continuous measurements.

6.8. Modified Integral Control Globally Counters Symmetry-Breaking Biases

Participants: Alain Sarlette

This is the end of the PhD thesis of Zhifei Zhang, under joint supervision of A.Sarlette at Ghent University and Zhihao Ling at ECUST Shanghai. The work [23] builds on our earlier proposal of formulating integral control on nonlinear groups as the integral of proportional correcting feedback actions: since these actions belong to a Lie algebra, they can be integrated in this vector space and applied at the current point. We here show how this controller can be modified in order to recover coordinated motion among steering-controlled vehicles, in a situation where biases would make the standard controller fail. We prove how the simple addition of the integral controller allows to recover global convergence towards the coordinated motion, restoring symmetry exactly.

6.9. Quantum Fast-Forwarding: Markov Chains and graph property testing

Participants: Alain Sarlette

This is the end of the PhD thesis of S.Apers, under supervision of A.Sarlette at Ghent University. In [11], we propose a quantum algorithmic routine, called Quantum Fast-Forwarding, which allows to simulate a Markov chain quadratically faster on a quantum computer than on a classical one. The key novelty, from an application point of view, is that we can achieve this acceleration not only for reaching the asymptotic distribution, but also for any intermediate time that one would be interested in. Such transient behaviors of Markov chains are important in algorithmic context, for instance to distinguish clusters in graphs. We explicitly work out those applications on graph properties.

6.10. Quantum Adiabatic Elimination: extension to rotating systems

Participants: Paolo Forni, Timothée Launay, Alain Sarlette and Pierre Rouchon

Adiabatic elimination is a technique to eliminate fast converging variables of a large system, while retaining their impact on slower dynamics of interest. Its most extreme form is a standard procedure when neglecting the dynamics of e.g. actuators or measurement devices in dynamical systems. In quantum systems it is particularly relevant to eliminate subsystems in tensor product structure. However, a major constraint is to obtain a reduced system in quantum form (Lindblad equations), preserving positivity and the unit trace. After having set up the framework for quantum adiabatic elimination to arbitrary order as a series expansion during the thesis of Rémi Azouit, we had worked out first- and second-order Lindblad equations only. With Paolo Forni, we have been pursuing the development of explicit formulas for higher-order cases. In [26], we present an extension of the technique for the case where the slowly decaying subsystem of interest, is subject to fast Hamiltonian dynamics. This appears e.g. in systems with significant detunings, where a description in rotating frame would lead to time-dependent equations if one does not want to neglect fast oscillating terms.

6.11. Minimizing decoherence on target in bipartite open quantum systems

Participants: Paolo Forni and Alain Sarlette

We consider a target quantum system, coupled to an auxiliary quantum system which dissipates rapidly at somewhat adjustable rates. The goal is to minimize the dissipation induced on the target system by this coupling. In [27], we use explicit model reduction formulas to express this as a quadratic optimization problem. We prove that maybe counterintuitively, when the auxiliary system dissipates along Hermitian (entropy-increasing) channels, the minimum induced dissipation is reached by maximizing the dissipation rate of the auxiliary system. This may be interpreted as a dynamical decoupling among the target system and the auxiliary one, induced not by standard Hamiltonian control acting on the target, but by noise acting on the environment. This link has been pursued with PhD student Michiel Burgelman and should lead to further results next year.

6.12. Repetition Cat Qubits for Fault-Tolerant Quantum Computation

Participants: Jérémie Guillaud and Mazyar Mirrahimi

We present a 1D repetition code based on the so-called cat qubits as a viable approach toward hardware-efficient universal and fault-tolerant quantum computation. The cat qubits that are stabilized by a two-photon driven-dissipative process exhibit a tunable noise bias where the effective bit-flip errors are exponentially suppressed with the average number of photons. We propose a realization of a set of gates on the cat qubits that preserve such a noise bias. Combining these base qubit operations, we build, at the level of the repetition cat qubit, a universal set of fully protected logical gates. This set includes single-qubit preparations and measurements, not, controlled-not, and controlled-controlled-not (Toffoli) gates. Remarkably, this construction avoids the costly magic state preparation, distillation, and injection. Finally, all required operations on the cat qubits could be performed with slight modifications of existing experimental setups.

This result was recently published in Physical Review X [16].

6.13. Experimental Implementation of a Raman-Assisted Eight-Wave Mixing Process

Participants: Mazyar Mirrahimi

Nonlinear processes in the quantum regime are essential for many applications, such as quantum-limited amplification, measurement, and control of quantum systems. In particular, the field of quantum error correction relies heavily on high-order nonlinear interactions between various modes of a quantum system. However, the required order of nonlinearity is often not directly available or weak compared to dissipation present in the system. Here, following our earlier theoretical proposal [72] we experimentally demonstrate a route to obtain higher-order nonlinearity by combining more easily available lower-order nonlinear processes, using a generalization of the Raman transition. In particular, we show a transformation of four photons of a high-Q superconducting resonator into two excitations of a superconducting transmon mode and two pump photons, and vice versa. The resulting eight-wave mixing process is obtained by cascading two fourth-order nonlinear processes through a virtual state. We expect this type of process to become a key component of hardware-efficient quantum error correction using continuous-variable error-correction codes. This work in collaboration with the group of Michel Devoret at Yale university was published in [18].

6.14. Stabilized Cat in a Driven Nonlinear Cavity: A Fault-Tolerant Error Syndrome Detector

Participants: Philippe Campagne-Ibarcq and Mazyar Mirrahimi

In quantum error correction, information is encoded in a high-dimensional system to protect it from the environment. A crucial step is to use natural, two-body operations with an ancilla to extract information about errors without causing backaction on the encoded information. Essentially, ancilla errors must not propagate to the encoded system and induce errors beyond those which can be corrected. The current schemes for achieving this fault tolerance to ancilla errors come at the cost of increased overhead requirements. An efficient way to extract error syndromes in a fault-tolerant manner is by using a single ancilla with a strongly biased noise channel. Typically, however, required elementary operations can become challenging when the noise is extremely biased. In this collaborative work with the groups of Steven Girvin and Michel Devoret at Yale University, we propose to overcome this shortcoming by using a bosonic-cat ancilla in a parametrically driven nonlinear oscillator. Such a cat qubit experiences only bit-flip noise, while the phase flips are exponentially suppressed. To highlight the flexibility of this approach, we illustrate the syndrome extraction process in a variety of codes such as qubit-based toric, bosonic-cat, and Gottesman-Kitaev-Preskill codes. Our results open a path for realizing hardware-efficient, fault-tolerant error syndrome extraction. This work was published in [21].

7. Partnerships and Cooperations

7.1. Regional Initiatives

- **Paris EMERGENCE project ENDURANCE**: In the framework of the Paris Ile de France program “EMERGENCE”, Zaki Leghtas has received a funding for his research program "Multi-photon processes in superconducting circuits for quantum error correction". This grant of 230k euros has allowed us to purchase the experimental equipment to complement the experiment based at ENS.
- **DIM SIRTEQ PhD fellowship**: We have received funding from DIM SIRTEQ to cover half of the PhD of Jérémie Guillaud under supervision of Mazyar Mirrahimi.
- **DIM SIRTEQ project SCOOP**:Half a PhD grant for Marius Villiers, supervised by Zaki Leghtas and Audrey Cottet (ENS Paris). The project is to use quantum circuits to detect the entanglement of a single Cooper pair. University.
- **EDPIF PhD fellowship**: Ecole Doctorale de Physique en Ile de France has funded half a PhD grant for Marius Villiers.
- **DGA PhD fellowship**: Direction Générale de l’Armement has funded half a PhD grant for Camille Berdou supervised by Zaki Leghtas. The project is to build a repetition code of cat-qubits.
- **Mines Paristech PhD Fellowship**: Ecole des Mines Paristech has funded half a PhD grant for Camille Berdou.
- **PSL working group on “structural stability and chaos in open quantum systems”**: This is a Groupe de Travail with researchers from CEREMADE (Paris Dauphine) and Observatoire de Paris under the direction of Jacques Fejoz. In the framework of the PhD thesis of Michiel Burgelman, we study the dynamics of superconducting Josephson circuits driven by strong microwave drives.

7.2. National Initiatives

- **ANR project ENDURANCE**: In the framework of the ANR program “Accueil de chercheur de haut niveau”, Zaki Leghtas has received a funding for his research program "Multi-photon processes in superconducting circuits for quantum error correction". This grant of 400k euros has allowed us to purchase the experimental equipment to build a new experiment based at ENS. The project started in March 2016 for 42 months.
- **ANR project HAMROQS**: In the framework of the ANR program JCJC, Alain Sarlette has received a funding for his research program "High-accuracy model reduction for open quantum systems". This grant of 212k euros started on april 2019 and will run for 4 years.

7.3. European Initiatives

7.3.1. FP7 & H2020 Projects

Program: H2020

Type: ERC

Project acronym: ECLIPSE

Project title: Exotic superconducting CIrcuits to Probe and protect quantum States of light and mattEr

Duration: 2019-

Coordinator: Zaki Leghtas, Mines Paristech

Program: H2020

Type: Quantera

Project acronym: QuCos

Project title: Quantum Computation with Schrödinger cat states

Duration: 2019-

Coordinator: Gerhard Kirchmair, University of Innsbruck, Austria.

Inria contacts: Zaki Leghtas and Mazyar Mirrahimi

Other partners: ENS Lyon (France), Karlsruhe Institut of Technology (Germany), Quantum Machines (Israel), National Institute for Research and Development of Isotopic and Molecular Technologies, Romania.

Abstract: This project seeks to establish a radically new, alternative approach to realizing the fundamental building blocks of quantum computers with superconducting qubits. In the next 3 years, we plan to employ only a handful of realistic components to realize robust error-corrected logical quantum bits. We aim to demonstrate the same level of protection provided by a few hundreds of qubits (with properties beyond the state of the art) in today's mainstream approach of the so-called surface code architecture. Our alternative approach is known as cat codes, because it employs multiple interconnected high coherence cavity modes with non-linear dissipation, to encode a qubit in superpositions of Schrödinger cat states. Our project combines realizing the quantum processor architecture as well as the control system and the protocols that drive it, building towards a full-stack error-corrected quantum computer. The partners in our collaboration form a strong synergetic group that has the full range of expertise needed to design and realize these systems, and to obtain these challenging goals. Furthermore, all partners of our project, including both industry and academia, have worked together and published works in the fields of quantum computing and quantum information processing. We aim to implement error protected qubits, fault tolerant operations, and demonstrate the scalability of this approach by realizing a repetition code. Our project will enable quantum experiments towards the ambitious and well-defined goal of constructing a logical qubit, on which we can perform gates, and most importantly, quantum error-correction (QEC).

7.3.2. Collaborations with Major European Organizations

Partner 1: ENS Lyon

We are pursuing our interdisciplinary work about quantum control from theoretical aspects in direct collaboration with existing experiments (ENS Lyon) with the group of Benjamin Huard, former member of the QUANTIC team. Joint papers are published and underway. The ANR-JCJC project HAMROQS by Alain Sarlette has Benjamin Huard as external supporting collaborator.

Partner 2: Laboratoire Kastler Brossel

We have been continuing collaborations with the teams of Samuel Deleglise and Igor Dotsenko from Laboratoire Kastler Brossel on the theoretical analysis of their experiments.

Partner 3: Ghent University.

Alain Sarlette has been collaborating with applied mathematicians interested in quantum control at UGent in the framework of thesis co-supervisions. One PhD student has successfully defended his thesis this year (Zhifei Zhang).

7.4. International Initiatives

7.4.1. Inria International Labs

Inria@EastCoast

Associate Team involved in the International Lab:

7.4.1.1. TAQUILLA

Title: TAilored QUantum Information protocoLs for quAntum superconducting circuits

International Partner (Institution - Laboratory - Researcher):

Université Yale (United States) -Department of Applied Physics - Michel Devoret

Start year: 2019

See also: https://team.inria.fr/quantic/Taquilla.html

We seek to establish an alternative approach to quantum error correction (QEC) for superconducting qubits. This approach, developed through the Inria-Yale collaboration, is known under the name of cat codes, because it employs multiple interconnected high coherence cavity modes with non-linear dissipation to encode a qubit in superpositions of Schrödinger cat states. We aim to implement error protected qubits, fault tolerant operations, and demonstrate the scalability of this approach. Our project will enable quantum experiments towards the ambitious and well-defined goal of constructing a logical qubit, on which we can perform gates, and most importantly, QEC.

7.4.2. Participation in Other International Programs

- **Yale-ARO subaward:** In the framework of the collaborations with Yale university, Quantic team has received a sub-award of 500k dollars over 4 years starting in 2018 from Yale university. This sub-award is part of an ARO (Army Research Office) grant received by our collaborators at Yale and covers the expenses related to our collaborations (hiring of new PhD students and postdocs at Inria and travels between Inria and Yale).
- **DARPA:** Alain Sarlette is international key personnel on the DARPA project "The Quantum Computing Revolution and Optimization: Challenges and Opportunities" led by optimization researchers at Lehigh Universty. This project of about 2M dollars can fund some exchanges during the coming years.
- **Berkeley exchange initiative:** P. Rouchon and A. Sarlette have set up an exchange initiative with Birgitta Whaley about quantum control and error correction based on continuous measurements. This initiative has funded a research visit of Gerardo Cardona at Berkeley; a student from Berkeley is bound to visit us soon in return.

7.5. International Research Visitors

7.5.1. Visits of International Scientists

P.S. Pereira da Silva (Escola Politecnica, PTC, University of Sao Paulo, Brazil) made two visits (September 16 to 27 and December 2 to 6) to investigate with Pierre Rouchon motion planning issues based on Lyapunov tracking for quantum gate generations for open quantum systems governed by Lindblad master equations.

7.5.2. Visits to International Teams

7.5.2.1. Research Stays Abroad

- In the framework of our collaborations with Yale (Taquilla associated team), Mazyar Mirrahimi and Michiel Burgelman have spent 3 months at Yale. In the same framework Philippe Campagne-Ibarcq and Christian Siegele also made a visit of 5 days during the same period.
- In the framework of the Berkeley exchange initiative, Gerardo Cardona has spent a month (October 2019) in the research group of Birgitta Whaley at Berkeley University.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Journal

8.1.1.1. Member of the Editorial Boards

Pierre Rouchon is member of the editorial board of Annual Reviews in Control.

8.1.1.2. Reviewer - Reviewing Activities

- Philippe Campagne-Ibarcq was a reviewer for Physical Review Letters.
- Zaki Leghtas was a reviewer for Physical Review Letters, Physical Review X and Nature Physics.
- Alain Sarlette was a reviewer for several automatic control and dynamical systems journals and conferences, as well as for Physical Review Journals

8.1.2. Invited Talks

- Philippe Campagne-Ibarcq : GDR-IQFA (Groupement de Recherche sur l'Ingénierie Quantique, des Aspects Fondamentaux aux Applications).
- Jeremie Guillaud: CEA Leti innovation day, Grenoble.
- Jeremie Guillaud: Byron Bay Quantum Workshop on Bosonic Error-Correcting Codes, Australia.
- Zaki Leghtas: GDR Physique Mésoscopique, Aussois.
- Zaki Leghtas: CIFAR Workshop on Quantum Cavities, Jouvence, Canada.
- Zaki Leghtas: American Physical Society march meeting, Boston, USA.
- Zaki Leghtas: Les Houches summer school on Quantum Information Machines.
- Zaki Leghtas: CEA Leti innovation day, Grenoble.
- Zaki Leghtas: Oulu university, Finland.
- Zaki Leghtas: Quantum and neuromorphic computing, C2N Saclay.
- Mazyar Mirrahimi: Lecture series at Les Houches summer school on Quantum Information Machines.
- Mazyar Mirrahimi: Conference on "Marching towards quantum supremacy", Princeton Univ., USA.
- Mazyar Mirrahimi: Conference for "20'th anniversary of superconducting qubits", Tsukuba, Japan.
- Pierre Rouchon: First Quantum Science, Engineering and Technology (qSET) conference, UNSW Canberra, Australia.
- Pierre Rouchon: IFAC Mechatronics and Nolcos Conference, Semi-plenary speaker, Vienna, Austria.
- Pierre Rouchon: 21-hour course on quantum control in the International Graduate School on Mathematical Control, Mathematical College Sichuan University, Chengdu, China.

- Pierre Rouchon: 6-hour course on flat system in the 8eme École d'été de mécanique théorique à destination des doctorants et chercheurs en Mécanique, Quiberon.
- Alain Sarlette: GANIL (Grand Accélérateur National d'Ions Lourd), Caen.
- Alain Sarlette: CIGREF (Club Informatique des Grandes Entreprises Françaises), Paris.
- Alain Sarlette has given an overview of the mathematics involved in quantum dynamics at Ghent University at several occasions.

8.1.3. Scientific Expertise

- Pierre Rouchon is a member of the scientific committee of LAGEP (Laboratoire d'Automatique et de Génie des Procédés) since 2017.
- Pierre Rouchon is a member of the "Conseil Scientifique du DIM Math Innov" since 2017.
- Pierre Rouchon is a member of the "Conseil de la recherche de PSL " since 2016.
- Mazyar Mirrahimi is the co-president of Inria's comité des emplois scientifiques.
- Mazyar Mirrahimi was the vice-president of ANR Comité d'Evaluation Scientifique on Quantum Technologies.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Cycle Ingénieur : Zaki Leghtas, Quantum Mechanics and Statistical Physics, Mines ParisTech, 12 hours, France.

Cycle Ingénieur : Zaki Leghtas, Quantum Computing, Mines ParisTech, 20 hours, France.

Cycle ingénieur: Alain Sarlette, Probabilities and Stochastic Processes, 24 hours TD, Mines Paristech, France.

Master: Alain Sarlette, Robotics, 24 hours, Ghent University, Belgium.

Cycle Ingénieur : Mazyar Mirrahimi, Automatic control with Applications in Robotics and in Quantum engineering, 8 hour amphi and 16 hours TD, 3rd year, Ecole Polytechnique, France.

Cycle Ingénieur : Mazyar Mirrahimi, Control of dynamical models, 30 hours TD, 2nd year, Ecole Polytechnique, France.

Cycle Ingénieur : Mazyar Mirrahimi, Module algorithmique quantum control, 24 hours TD, 2nd year, Ecole Polytechnique, France.

Master: Mazyar Mirrahimi and Pierre Rouchon, Dynamics and control of quantum systems, 18 hours amphi, M2, Sorbonne Université, France.

8.2.2. Supervision

PhD in progress : Michiel Burgelman, A systematic study of strongly driven and dissipative quantum systems towards high-accuracy quantum control designs, advisors: Pierre Rouchon and Alain Sarlette, starting date: Nov 2018.

PhD in progress : Jérémie Guillaud, Fault-tolerant quantum computation with cat-qubits, advisor: Mazyar Mirrahimi, starting date: Nov 2017.

PhD in progress : Vincent Martin, Fault-tolerance of quantum systems under continuous-time feedback stabilization, advisor: Mazyar Mirrahimi and Alain Sarlette, starting date: Oct 2018.

PhD in progress: Raphaël Lescanne, Engineering Multi-Photon Dissipation In Superconducting Circuits For Quantum Error Correction, advisors: Zaki Leghtas and Takis Kontos, starting date Sept 2016.

PhD in progress: Marius Villiers, Probing the spin entanglement of single Cooper pair, advisors: Zaki Leghtas and Takis Kontos, starting date: September 2019.

PhD in progress: Camille Berdou, A cat-qubit repetition code, advisors: Zaki Leghtas and Pierre Rouchon, starting date Sept 2019.

PhD in progress: Christian Siegele, Quantum error correction with grid states of light, advisors: Philippe Campagne-Ibarcq and Mazyar Mirrahimi.

PhD: Lucas Verney, Strongly driven quantum Josephson circuits, advisors: Zaki Leghtas and Mazyar Mirrahimi, Defended on July 11th 2019.

PhD: Gerardo Cardona Sanchez, Exponential stabilization of quantum systems subject to non-demolition measurements in continuous time, advisors: Pierre Rouchon and Alain Sarlette, Defended on October 30th 2019.

8.2.3. Juries

- Zaki Leghtas was a PhD opponent of Iivari Pietikainen, Oulu Finland.
- Zaki Leghtas was a PhD jury member of Romain Albert, CEA Grenoble.
- Mazyar Mirrahimi was an examiner in the PhD defense of Filippo Vicentini, Paris Diderot University.
- Mazyar Mirrahimi was an examiner in the HDR defense of Igor Dotsenko, Collège de France.
- Pierre Rouchon was reviewer for the Habiliation thesis of Hector Ramirez Estay, Université de Franche-Comté.
- Pierre Rouchon was a reviewer of the PhD thesis of Andreas Deutschmann (TU Vienna), Amira Amraoui (Université de Nice), Yuanlong Wang (University of New South Wales in Canberra).
- Pierre Rouchon was a committee member for the PhD thesis of Christophe Zhan (Sorbonne Université), Vincent Metillon (ENS Paris), Yahao Chen (Universté de Rouen), Weichao Liang (Université Paris-Saclay).
- Alain Sarlette has been member of the PhD jury of Nicolas Augier (Ecole Polytechnique) and of Estelle Massart (Université Catholique de Louvain-la-Neuve).

8.3. Popularization

8.3.1. Articles and contents

- Zaki Leghtas article in Usbek and Rica (https://usbeketrica.com/article/aucun-etat-veut-rater-coche-ordinateur-quantique)

8.3.2. Interventions

- Zaki Leghtas: Pint of Science on quantum computing

9. Bibliography

Major publications by the team in recent years

[1] H. Amini, A. Somaraju, I. Dotsenko, C. Sayrin, M. Mirrahimi, P. Rouchon.*Feedback stabilization of discrete-time quantum systems subject to non-demolition measurements with imperfections and delays*, in "Automatica", 2013, vol. 49, n^{o} 9, p. 2683–2692

[2] P. Campagne-Ibarcq, P. Six, L. Bretheau, A. Sarlette, M. Mirrahimi, P. Rouchon, B. Huard.*Observing Quantum State Diffusion by Heterodyne Detection of Fluorescence* , in "Physical Review X", January 2016, vol. 6, 011002 [*DOI :* 10.1103/PhysRevX.6.011002], https://hal-mines-paristech.archives-ouvertes.fr/hal-01264326

[3] G. CARDONA, A. SARLETTE, P. ROUCHON.*Exponential stabilization of quantum systems under continuous non-demolition measurements*, in "Automatica", December 2019, https://hal.inria.fr/hal-02394692

[4] J. GUILLAUD, M. MIRRAHIMI.*Repetition Cat Qubits for Fault-Tolerant Quantum Computation*, in "Physical Review X", December 2019, https://arxiv.org/abs/1904.09474 - 22 pages, 11 figures [*DOI :* 10.1103/PHYSREVX.9.041053], https://hal.inria.fr/hal-02413978

[5] Z. LEGHTAS, G. KIRCHMAIR, B. VLASTAKIS, R. J. SCHOELKOPF, M. H. DEVORET, M. MIRRAHIMI.*Hardware-Efficient Autonomous Quantum Memory Protection*, in "Phys. Rev. Lett.", Sep 2013, vol. 111, 120501, https://link.aps.org/doi/10.1103/PhysRevLett.111.120501

[6] Z. LEGHTAS, S. TOUZARD, I. M. POP, A. KOU, B. VLASTAKIS, A. PETRENKO, K. M. SLIWA, A. NARLA, S. SHANKAR, M. J. HATRIDGE, M. REAGOR, L. FRUNZIO, R. J. SCHOELKOPF, M. MIRRAHIMI, M. H. DEVORET.*Confining the state of light to a quantum manifold by engineered two-photon loss*, in "Science", February 2015, vol. 347, n^{o} 6224, p. 853-857 [*DOI :* 10.1126/SCIENCE.AAA2085], https://hal.inria.fr/hal-01240210

[7] M. MIRRAHIMI, Z. LEGHTAS, V. V. ALBERT, S. TOUZARD, R. J. SCHOELKOPF, L. JIANG, M. H. DEVORET.*Dynamically protected cat-qubits: a new paradigm for universal quantum computation*, in "New Journal of Physics", apr 2014, vol. 16, n^{o} 4, 045014

[8] N. OFEK, A. PETRENKO, R. HEERES, P. REINHOLD, Z. LEGHTAS, B. VLASTAKIS, Y. LIU, L. FRUNZIO, S. GIRVIN, L. JIANG, M. MIRRAHIMI, M. H. DEVORET, R. J. SCHOELKOPF.*Extending the lifetime of a quantum bit with error correction in superconducting circuits*, in "Nature", 2016, vol. 536, 5

[9] A. SARLETTE, J.-M. RAIMOND, M. BRUNE, P. ROUCHON.*Stabilization of nonclassical states of the radiation field in a cavity by reservoir engineering*, in "Phys. Rev. Lett.", 2011, vol. 107, 010402

[10] C. SAYRIN, I. DOTSENKO, X. ZHOU, B. PEAUDECERF, T. RYBARCZYK, S. GLEYZES, P. ROUCHON, M. MIRRAHIMI, H. AMINI, M. BRUNE, J.-M. RAIMOND, S. HAROCHE.*Real-time quantum feedback prepares and stabilizes photon number states*, in "Nature", 2011, vol. 477, p. 73–77

Publications of the year

Articles in International Peer-Reviewed Journal

[11] S. APERS, A. SARLETTE.*Quantum Fast-Forwarding: Markov chains and graph property testing*, in "Quantum Information & Computation", April 2019, https://hal.inria.fr/hal-02394399

[12] G. CARDONA, A. SARLETTE, P. ROUCHON.*Exponential stabilization of quantum systems under continuous non-demolition measurements*, in "Automatica", December 2019 [*DOI :* 10.1016/J.AUTOMATICA.2019.108719], https://hal.inria.fr/hal-02394692

[13] T. CUBAYNES, M. R. DELBECQ, M. DARTIAILH, R. ASSOULY, M. DESJARDINS, L. CONTAMIN, L. BRUHAT, Z. LEGHTAS, F. MALLET, A. COTTET, T. KONTOS.*Highly coherent spin states in carbon nanotubes coupled to cavity photons*, in "npj Quantum Information", December 2019, vol. 5, 47, https://arxiv.org/abs/1903.05229 [*DOI :* 10.1038/S41534-019-0169-4], https://hal.archives-ouvertes.fr/hal-02318737

[14] A. FARNAM, A. SARLETTE.*About string stability of a vehicle chain with unidirectional controller*, in "European Journal of Control", December 2019 [*DOI :* 10.1016/J.EJCON.2019.03.002], https://hal.inria.fr/hal-02394705

[15] A. FARNAM, A. SARLETTE.*Towards a comprehensive impossibility result for string stability*, in "IEEE Transactions on Automatic Control", December 2019 [*DOI :* 10.1109/TAC.2019.2929967], https://hal.inria.fr/hal-02394703

[16] J. GUILLAUD, M. MIRRAHIMI.*Repetition Cat Qubits for Fault-Tolerant Quantum Computation*, in "Physical Review X", December 2019, https://arxiv.org/abs/1904.09474 - 22 pages, 11 figures [*DOI :* 10.1103/PHYSREVX.9.041053], https://hal.inria.fr/hal-02413978

[17] R. LESCANNE, L. VERNEY, Q. FICHEUX, M. H. DEVORET, B. HUARD, M. MIRRAHIMI, Z. LEGHTAS.*Escape of a Driven Quantum Josephson Circuit into Unconfined States*, in "Physical Review Applied", January 2019, vol. 11, n^{o} 1, 014030, https://arxiv.org/abs/1805.05198 [*DOI :* 10.1103/PHYSREVAPPLIED.11.014030], https://hal.archives-ouvertes.fr/hal-02318762

[18] S. O. MUNDHADA, A. GRIMM, J. VENKATRAMAN, Z. K. MINEV, S. TOUZARD, N. E. FRATTINI, V. V. SIVAK, K. M. SLIWA, P. REINHOLD, S. SHANKAR, M. MIRRAHIMI, M. H. DEVORET.*Experimental implementation of a Raman-assisted eight-wave mixing process*, in "Physical Review Applied", November 2019, https://arxiv.org/abs/1811.06589 - 13 pages, 5 figures [*DOI :* 10.1103/PHYSREVAPPLIED.12.054051], https://hal.inria.fr/hal-01936696

[19] V. MÉTILLON, S. GERLICH, M. BRUNE, J.-M. RAIMOND, P. ROUCHON, I. DOTSENKO.*Benchmarking Maximum-Likelihood State Estimation with an Entangled Two-Cavity State*, in "Physical Review Letters", August 2019, vol. 123, n^{o} 6, https://arxiv.org/abs/1904.04681 - 3 figures, Supplementary information [*DOI :* 10.1103/PHYSREVLETT.123.060404], https://hal.archives-ouvertes.fr/hal-02095250

[20] P. S. PEREIRA DA SILVA, H. B. SILVEIRA, P. ROUCHON.*Fast and virtually exact quantum gate generation in U (n) via iterative Lyapunov methods*, in "International Journal of Control", June 2019, vol. 00, p. 1-15 [*DOI :* 10.1080/00207179.2019.1626023], https://hal-mines-paristech.archives-ouvertes.fr/hal-02289258

[21] S. PURI, A. GRIMM, P. CAMPAGNE-IBARCQ, A. EICKBUSCH, K. NOH, G. ROBERTS, L. JIANG, M. MIRRAHIMI, M. H. DEVORET, S. M. GIRVIN.*Stabilized Cat in Driven Nonlinear Cavity: A Fault-Tolerant Error Syndrome Detector*, in "Physical Review X", October 2019, https://arxiv.org/abs/1807.09334 [*DOI :* 10.1103/PHYSREVX.9.041009], https://hal.inria.fr/hal-01936700

[22] L. VERNEY, R. LESCANNE, M. H. DEVORET, Z. LEGHTAS, M. MIRRAHIMI.*Structural Instability of Driven Josephson Circuits Prevented by an Inductive Shunt*, in "Physical Review Applied", February 2019, vol. 11, n^{o} 2, 024003, https://arxiv.org/abs/1805.07542 [*DOI :* 10.1103/PHYSREVAPPLIED.11.024003], https://hal.archives-ouvertes.fr/hal-02318777

[23] Z. ZHANG, Z. LING, A. SARLETTE.*Modified Integral Control Globally Counters Symmetry-Breaking Biases*, in "Symmetry", May 2019 [*DOI :* 10.3390/SYM11050639], https://hal.inria.fr/hal-02394405

International Conferences with Proceedings

[24] G. CARDONA, A. SARLETTE, P. ROUCHON.*Continuous-time Quantum Error Correction with Noise-assisted Quantum Feedback*, in "NOLCOS 2019 - IFAC Symposium on Nonlinear Control Systems", Vienna, Austria, September 2019, https://hal.inria.fr/hal-02394694

[25] A. FARNAM, A. SARLETTE.*About strong string stability of a vehicle chain with time-headway control*, in "CDC 2018 - IEEE Conference on Decision and Control 2018", Miami, United States, December 2019, https://hal.inria.fr/hal-02394701

[26] P. FORNI, T. LAUNAY, A. SARLETTE, P. ROUCHON.*A palette of approaches for adiabatic elimination in bipartite open quantum systems with Hamiltonian dynamics on target*, in "IEEE Conference on Decision and Control 2019", Nice, France, December 2019, https://hal.inria.fr/hal-02393960

[27] P. FORNI, A. SARLETTE.*Minimizing decoherence on target in bipartite open quantum systems*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://hal.inria.fr/hal-02393976

References in notes

[28] S. ATTAL, A. JOYE, C.-A. PILLET (editors). *Open Quantum Systems III: Recent Developments*, Springer, Lecture notes in Mathematics 1880, 2006

[29] H. AMINI, M. MIRRAHIMI, P. ROUCHON.*Stabilization of a delayed quantum system: the Photon Box case-study*, in "IEEE Trans. Automatic Control", 2012, vol. 57, n^{o} 8, p. 1918–1930

[30] H. AMINI, C. PELLEGRINI, P. ROUCHON.*Stability of continuous-time quantum filters with measurement imperfections*, in "Russian Journal of Mathematical Physics", 2014, vol. 21, p. 297–315

[31] A. BARCHIELLI, M. GREGORATTI.*Quantum Trajectories and Measurements in Continuous Time: the Diffusive Case*, Springer Verlag, 2009

[32] J. BARREIRO, M. MULLER, P. SCHINDLER, D. NIGG, T. MONZ, M. CHWALLA, M. HENNRICH, C. ROOS, P. ZOLLER, R. BLATT.*An open-system quantum simulator with trapped ions*, in "Nature", 2011, vol. 470, 486

[33] V. BELAVKIN.*Quantum stochastic calculus and quantum nonlinear filtering*, in "Journal of Multivariate Analysis", 1992, vol. 42, n^{o} 2, p. 171–201

[34] T. BENOIST, C. PELLEGRINI.*Large Time Behavior and Convergence Rate for Quantum Filters Under Standard Non Demolition Conditions*, in "Communications in Mathematical Physics", 2014, p. 1-21, http://dx.doi.org/10.1007/s00220-014-2029-6

[35] G. BIRKHOFF.*Extensions of Jentzch's theorem*, in "Trans. Amer. Math. Soc.", 1957, vol. 85, p. 219–227

[36] S. BOLOGNANI, F. TICOZZI.*Engineering stable discrete-time quantum dynamics via a canonical QR decomposition*, in "IEEE Trans. Autom. Control", 2010, vol. 55

[37] V. BRAGINSKI, F. KHALILI.*Quantum Measurements*, Cambridge University Press, 1992

[38] P. CAMPAGNE-IBARCQ, L. BRETHEAU, E. FLURIN, A. AUFFÈVES, F. MALLET, B. HUARD.*Observing Interferences between Past and Future Quantum States in Resonance Fluorescence*, in "Phys. Rev. Lett.", May 2014, vol. 112, 180402, http://link.aps.org/doi/10.1103/PhysRevLett.112.180402

[39] P. CAMPAGNE-IBARCQ, E. FLURIN, N. ROCH, D. DARSON, P. MORFIN, M. MIRRAHIMI, M. H. DEVORET, F. MALLET, B. HUARD.*Persistent Control of a Superconducting Qubit by Stroboscopic Measurement Feedback*, in "Phys. Rev. X", 2013, vol. 3, 021008

[40] H. CARMICHAEL.*Statistical Methods in Quantum Optics 2: Non-Classical Fields*, Spinger, 2007

[41] H. CARMICHAEL.*An Open Systems Approach to Quantum Optics*, Springer-Verlag, 1993

[42] J. CARR.*Application of Center Manifold Theory*, Springer, 1981

[43] J. COHEN, M. MIRRAHIMI.*Dissipation-induced continuous quantum error correction for superconducting circuits*, in "Phys. Rev. A", 2014, vol. 90, 062344

[44] J. DALIBARD, Y. CASTIN, K. MÖLMER.*Wave-function approach to dissipative processes in quantum optics*, in "Phys. Rev. Lett.", 1992, vol. 68, n^{o} 5, p. 580–583

[45] M. H. DEVORET, A. WALLRAFF, J. MARTINIS.*Superconducting Qubits: A Short Review*, 2004, arXiv:cond-mat/0411174

[46] N. DIDIER, J. GUILLAUD, S. SHANKAR, M. MIRRAHIMI.*Remote entanglement stabilization and concentration by quantum reservoir engineering*, in "Physical Review A", July 2018, https://arxiv.org/abs/1703.03379 - 5 pages, 4 figures [*DOI :* 10.1103/PHYSREVA.98.012329], https://hal.inria.fr/hal-01652766

[47] I. DOTSENKO, M. MIRRAHIMI, M. BRUNE, S. HAROCHE, J.-M. RAIMOND, P. ROUCHON.*Quantum feedback by discrete quantum non-demolition measurements: towards on-demand generation of photon-number states*, in "Physical Review A", 2009, vol. 80: 013805-013813

[48] N. FENICHEL.*Geometric singular perturbation theory for ordinary differential equations*, in "J. Diff. Equations", 1979, vol. 31, p. 53–98

[49] J. GAMBETTA, H. M. WISEMAN.*State and dynamical parameter estimation for open quantum systems*, in "Phys. Rev. A", September 2001, vol. 64, n^{o} 4, 042105, http://link.aps.org/doi/10.1103/PhysRevA.64.042105

[50] S. GAMMELMARK, B. JULSGAARD, K. MÖLMER.*Past Quantum States of a Monitored System*, in "Phys. Rev. Lett.", October 2013, vol. 111, n^{o} 16, 160401, http://link.aps.org/doi/10.1103/PhysRevLett.111.160401

[51] C. GARDINER, P. ZOLLER.*Quantum Noise*, third, Springer, 2010

[52] S. GAUBERT, Z. QU.*Checking the strict positivity of Kraus maps is NP-hard*, in "arXiv:1402.1429", 2014

[53] S. GAUBERT, Z. QU.*The contraction rate in Thompson's part metric of order-preserving flows on a cone - Application to generalized Riccati equations*, in "Journal of Differential Equations", April 2014, vol. 256, n^{o} 8, p. 2902–2948, http://www.sciencedirect.com/science/article/pii/S0022039614000424

[54] K. GEERLINGS, Z. LEGHTAS, I. POP, S. SHANKAR, L. FRUNZIO, R. SCHOELKOPF, M. MIRRAHIMI, M. H. DEVORET.*Demonstrating a Driven Reset Protocol of a Superconducting Qubit*, in "Phys. Rev. Lett.", 2013, vol. 110, 120501

[55] D. GOTTESMAN, A. KITAEV, J. PRESKILL.*Encoding a qubit in an oscillator*, in "Phys. Rev. A", 2001, vol. 64, 012310

[56] C. GUERLIN, J. BERNU, S. DELÉGLISE, C. SAYRIN, S. GLEYZES, S. KUHR, M. BRUNE, J.-M. RAIMOND, S. HAROCHE.*Progressive field-state collapse and quantum non-demolition photon counting*, in "Nature", 2007, vol. 448, p. 889-893

[57] S. HAROCHE, J.-M. RAIMOND.*Exploring the Quantum: Atoms, Cavities and Photons*, Oxford University Press, 2006

[58] M. HATRIDGE, S. SHANKAR, M. MIRRAHIMI, F. SCHACKERT, K. GEERLINGS, T. BRECHT, K. SLIWA, B. ABDO, L. FRUNZIO, S. GIRVIN, R. SCHOELKOPF, M. H. DEVORET.*Quantum back-action of an individual variable-strength measurement*, in "Science", 2013, vol. 339, p. 178–181

[59] E. HOLLAND, B. VLASTAKIS, R. HEERES, M. REAGOR, U. VOOL, Z. LEGHTAS, L. FRUNZIO, G. KIRCHMAIR, M. DEVORET, M. MIRRAHIMI, R. SCHOELKOPF.*Single-photon-resolved cross-Kerr interaction for autonomous stabilization of photon-number states*, in "Phys. Rev. Lett.", 2015, vol. 115, 180501

[60] T. KATO.*Perturbation Theory for Linear Operators*, Springer, 1966

[61] E. KNILL, R. LAFLAMME, G. MILBURN.*A scheme for efficient quantum computation with linear optics*, in "Nature", 2001, vol. 409, 46

[62] H. KRAUTER, C. MUSCHIK, K. JENSEN, W. WASILEWSKI, J. PETERSEN, J. CIRAC, E. POLZIK.*Entanglement Generated by Dissipation and Steady State Entanglement of Two Macroscopic Objects*, in "Phys. Rev. Lett.", 2011, vol. 107, 080503

[63] Z. LEGHTAS, G. KIRCHMAIR, B. VLASTAKIS, M. H. DEVORET, R. J. SCHOELKOPF, M. MIRRAHIMI.*Deterministic protocol for mapping a qubit to coherent state superpositions in a cavity*, in "Phys. Rev. A", 2013, vol. 87, 042315

[64] Z. LEGHTAS, G. KIRCHMAIR, B. VLASTAKIS, R. J. SCHOELKOPF, M. H. DEVORET, M. MIRRAHIMI.*Hardware-efficient autonomous quantum memory protection*, in "Phys. Rev. Lett.", 2013, vol. 111, 120501

[65] Z. LEGHTAS, A. SARLETTE, P. ROUCHON.*Adiabatic passage and ensemble control of quantum systems*, in "J. Phys. B", 2011, vol. 44, 154017

[66] Z. LEGHTAS, U. VOOL, S. SHANKAR, M. HATRIDGE, S. GIRVIN, M. H. DEVORET, M. MIRRAHIMI.*Stabilizing a Bell state of two superconducting qubits by dissipation engineering*, in "Phys. Rev. A", 2013, vol. 88, 023849

[67] J.-S. LI, N. KHANEJA.*Ensemble control of Bloch equations*, in "IEEE Trans. Autom. Control", 2009, vol. 54, p. 528–536

[68] Y. LIN, J. GAEBLER, F. REITER, T. TAN, R. BOWLER, A. SORENSEN, D. LEIBFRIED, D. WINELAND.*Dissipative production of a maximally entangled steady state of two quantum bits*, in "Nature", 2013, vol. 504, p. 415–418

[69] S. LLOYD.*Coherent quantum feedback*, in "Phys. Rev. A", 2000, vol. 62, 022108

[70] L. MAZZARELLA, A. SARLETTE, F. TICOZZI.*Consensus for quantum networks: from symmetry to gossip iterations*, in "IEEE Trans. Automat. Control", 2014, in press

[71] M. MIRRAHIMI, B. HUARD, M. H. DEVORET.*Strong measurement and quantum feedback for persistent Rabi oscillations in circuit QED experiments*, in "IEEE Conference on Decision and Control", IEEE Conference on Decision and Control, 2012

[72] S. O. MUNDHADA, A. GRIMM, S. TOUZARD, U. VOOL, S. SHANKAR, M. H. DEVORET, M. MIRRAHIMI.*Generating higher-order quantum dissipation from lower-order parametric processes*, in "Quantum Science and Technology", may 2017, vol. 2, n^o 2, 024005

[73] K. MURCH, U. VOOL, D. ZHOU, S. WEBER, S. GIRVIN, I. SIDDIQI.*Cavity-assisted quantum bath engineering*, in "Phys. Rev. Lett.", 2012, vol. 109, 183602

[74] A. NEGRETTI, K. MÖLMER.*Estimation of classical parameters via continuous probing of complementary quantum observables*, in "New Journal of Physics", 2013, vol. 15, n^o 12, 125002, http://stacks.iop.org/1367-2630/15/i=12/a=125002

[75] H. NURDIN, M. JAMES, I. PETERSEN.*Coherent quantum LQG control*, in "Automatica", 2009, vol. 45, p. 1837–1846

[76] B. PEAUDECERF, T. RYBARCZYK, S. GERLICH, S. GLEYZES, J.-M. RAIMOND, S. HAROCHE, I. DOTSENKO, M. BRUNE.*Adaptive Quantum Nondemolition Measurement of a Photon Number*, in "Phys. Rev. Lett.", Feb 2014, vol. 112, n^o 8, 080401, http://link.aps.org/doi/10.1103/PhysRevLett.112.080401

[77] D. PETZ.*Monotone Metrics on matrix spaces*, in "Linear Algebra and its Applications", 1996, vol. 244, p. 81–96

[78] J. POYATOS, J. CIRAC, P. ZOLLER.*Quantum Reservoir Engineering with Laser Cooled Trapped Ions*, in "Phys. Rev. Lett.", 1996, vol. 77, n^o 23, p. 4728–4731

[79] D. REEB, M. J. KASTORYANO, M. M. WOLF.*Hilbert's projective metric in quantum information theory*, in "Journal of Mathematical Physics", August 2011, vol. 52, n^o 8, 082201, http://dx.doi.org/10.1063/1.3615729

[80] D. RISTÈ, J. LEEUWEN, H.-S. KU, K. LEHNERT, L. DICARLO.*Initialization by measurement of a superconducting quantum bit circuit*, in "Phys. Rev. Lett.", 2012, vol. 109, 050507

[81] N. ROCH, E. FLURIN, F. NGUYEN, P. MORFIN, P. CAMPAGNE-IBARCQ, M. H. DEVORET, B. HUARD.*Widely tunable, non-degenerate three-wave mixing microwave device operating near the quantum limit*, in "Phys. Rev. Lett.", 2012, vol. 108, 147701

[82] P. ROUCHON.*Fidelity is a Sub-Martingale for Discrete-Time Quantum Filters*, in "IEEE Transactions on Automatic Control", 2011, vol. 56, n^o 11, p. 2743–2747

[83] A. ROY, Z. LEGHTAS, A. STONE, M. DEVORET, M. MIRRAHIMI.*Continuous generation and stabilization of mesoscopic field superposition states in a quantum circuit*, in "Phys. Rev. A", 2015, vol. 91, 013810

[84] A. SARLETTE, Z. LEGHTAS, M. BRUNE, J.-M. RAIMOND, P. ROUCHON.*Stabilization of nonclassical states of one- and two-mode radiation fields by reservoir engineering*, in "Phys. Rev. A", 2012, vol. 86, 012114

[85] D. SCHUSTER, A. HOUCK, J. SCHREIER, A. WALLRAFF, J. GAMBETTA, A. BLAIS, L. FRUNZIO, J. MAJER, B. JOHNSON, M. H. DEVORET, S. GIRVIN, R. J. SCHOELKOPF.*Resolving photon number states in a superconducting circuit*, in "Nature", 2007, vol. 445, p. 515–518

[86] R. SEPULCHRE, A. SARLETTE, P. ROUCHON.*Consensus in non-commutative spaces*, in "Decision and Control (CDC), 2010 49th IEEE Conference on", 2010, p. 6596–6601

[87] S. SHANKAR, M. HATRIDGE, Z. LEGHTAS, K. SLIWA, A. NARLA, U. VOOL, S. GIRVIN, L. FRUNZIO, M. MIRRAHIMI, M. H. DEVORET.*Autonomously stabilized entanglement between two superconducting quantum bits*, in "Nature", 2013, vol. 504, p. 419–422

[88] P. SHOR.*Scheme for reducing decoherence in quantum memory*, in "Phys. Rev. A", 1995, vol. 52, p. 2493–2496

[89] P. SIX, P. CAMPAGNE-IBARCQ, I. DOTSENKO, A. SARLETTE, B. HUARD, P. ROUCHON.*Quantum state tomography with noninstantaneous measurements, imperfections, and decoherence*, in "Phys. Rev. A", 2016, vol. 93, 012109

[90] A. SOMARAJU, I. DOTSENKO, C. SAYRIN, P. ROUCHON.*Design and Stability of Discrete-Time Quantum Filters with Measurement Imperfections*, in "American Control Conference", 2012, p. 5084–5089

[91] A. SOMARAJU, M. MIRRAHIMI, P. ROUCHON.*Approximate stabilization of infinite dimensional quantum stochastic system*, in "Reviews in Mathematical Physics", 2013, vol. 25, 1350001

[92] A. STEANE.*Error Correcting Codes in Quantum Theory*, in "Phys. Rev. Lett", 1996, vol. 77, n^o 5

[93] L. SUN, A. PETRENKO, Z. LEGHTAS, B. VLASTAKIS, G. KIRCHMAIR, K. SLIWA, A. NARLA, M. HATRIDGE, S. SHANKAR, J. BLUMOFF, L. FRUNZIO, M. MIRRAHIMI, M. H. DEVORET, R. J. SCHOELKOPF.*Tracking photon jumps with repeated quantum non-demolition parity measurements*, in "Nature", 2014, vol. 511, p. 444–448

[94] J. TSITSIKLIS.*Problems in decentralized decision making and computation*, in "PhD Thesis, MIT", 1984

[95] R. VIJAY, C. MACKLIN, D. SLICHTER, S. WEBER, K. MURCH, R. NAIK, A. KOROTKOV, I. SIDDIQI.*Stabilizing Rabi oscillations in a superconducting qubit using quantum feedback*, in "Nature", 2012, vol. 490, p. 77–80

[96] L. VIOLA, E. KNILL, S. LLOYD.*Dynamical decoupling of open quantum system*, in "Phys. Rev. Lett.", 1999, vol. 82, p. 2417-2421

[97] B. VLASTAKIS, G. KIRCHMAIR, Z. LEGHTAS, S. NIGG, L. FRUNZIO, S. GIRVIN, M. MIRRAHIMI, M. H. DEVORET, R. J. SCHOELKOPF.*Deterministically encoding quantum information using 100-photon Schrödinger cat states*, in "Science", 2013, vol. 342, p. 607–610

[98] X. ZHOU, I. DOTSENKO, B. PEAUDECERF, T. RYBARCZYK, C. SAYRIN, S. GLEYZES, J.-M. RAIMOND, M. BRUNE, S. HAROCHE.*Field locked to Fock state by quantum feedback with single photon corrections*, in "Physical Review Letter", 2012, vol. 108, 243602

[99] R. VAN HANDEL.*The stability of quantum Markov filters*, in "Infin. Dimens. Anal. Quantum Probab. Relat. Top.", 2009, vol. 12, p. 153–172

Team REO

Numerical simulation of biological flows

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
Paris

THEME
Modeling and Control for Life Sciences

Table of contents

Team REO

Creation of the Team: 2019 January 01

Keywords:

Computer Science and Digital Science:

A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.5. - Multiphysics modeling
A6.3.2. - Data assimilation
A6.5.2. - Fluid mechanics

Other Research Topics and Application Domains:

B2.2. - Physiology and diseases
B2.4.3. - Surgery

1. Team, Visitors, External Collaborators

Research Scientists

Irene Vignon-Clementel [Team leader, Inria, Senior Researcher, HDR]
Miguel Angel Fernandez Varela [Team leader, Inria, Senior Researcher, until May 2019, HDR]
Céline Grandmont [Inria, Senior Researcher, until May 2019, HDR]
Damiano Lombardi [Inria, Researcher, until May 2019]
Marina Vidrascu [Inria, Emeritus, until May 2019, HDR]

Faculty Members

Laurent Boudin [Univ Pierre et Marie Curie, Associate Professor, until May 2019]
Muriel Boulakia [Univ Pierre et Marie Curie, Associate Professor, until May 2019]

Technical Staff

Daniele Carlo Corti [Inria, Engineer, until May 2019]
Fabien Raphel [Inria, Engineer, until Mar 2019]

PhD Students

Nicolas Golse [APHP-Inria, PhD Student]
Ludovic Boilevin-Kayl [Inria, PhD Student, until May 2019]
Chen-Yu Chiang [Univ Pierre et Marie Curie, PhD Student, until Jan 2019]
Felipe Galarce Marin [Inria, PhD Student, until May 2019]
Fannie Gerosa [Inria, PhD Student, until May 2019]
Fabien Raphel [Inria, PhD Student, from Apr 2019 until May 2019]

Post-Doctoral Fellow

Jean Jerome Casanova [Inria, Post-Doctoral Fellow, until May 2019]

Visiting Scientists

Damien Dousse [Inserm, until Nov 2019]
David Sampedro Puente [from Feb 2019 until Apr 2019]

Administrative Assistants

Laurence Bourcier [Inria, Administrative Assistant]
Julien Guieu [Inria, Administrative Assistant, from Apr 2019]
Maryse Desnous [Inria, Administrative Assistant, until Apr 2019]

2. Overall Objectives

2.1. Overall Objectives

REO is a joint project-team of the Inria Research Center of Paris and the Jacques-Louis Lions Laboratory (LJLL) of the Pierre and Marie Curie University (Sorbonne Universié, UPMC Paris 6) and CNRS (UMR7598). Its main objectives are:

- the modeling of blood flow in large vessels, air flow in the respiratory tract, and the cardiac electrophysiology;
- the design and the analysis of efficient and robust numerical methods for these problems;
- the development of numerical software to assist medical decisions and to contribute to the design of medical devices.

REO put a strong effort in working with real data, coming either from clinicians or industrial partners. The development of methods for the interaction of data and simulation is therefore an important aspect of the activity of the team.
Most of the REO members are now in the project-team COMMEDIA since June 2019. The team description (research program and application domains) remains the original REO one. The other sections, and in particular the new results only reflect the remaining members. The articles of the year 2019 are the ones of REO members that did not move to COMMEDIA.

3. Research Program

3.1. Multiphysics modeling

In large vessels and in large bronchi, blood and air flows are generally supposed to be governed by the incompressible Navier-Stokes equations. Indeed in large arteries, blood can be supposed to be Newtonian, and at rest air can be modeled as an incompressible fluid. The cornerstone of the simulations is therefore a Navier-Stokes solver. But other physical features have also to be taken into account in simulations of biological flows, in particular fluid-structure interaction in large vessels and transport of sprays, particles or chemical species.

3.1.1. Fluid-structure interaction

Fluid-structure coupling occurs both in the respiratory and in the circulatory systems. We focus mainly on blood flows since our work is more advanced in this field. But the methods developed for blood flows could be also applied to the respiratory system.

Here "fluid-structure interaction" means a coupling between the 3D Navier-Stokes equations and a 3D (possibly thin) structure in large displacements.

The numerical simulations of the interaction between the artery wall and the blood flows raise many issues: (1) the displacement of the wall cannot be supposed to be infinitesimal, geometrical nonlinearities are therefore present in the structure and the fluid problem have to be solved on a moving domain (2) the densities of the artery walls and the blood being close, the coupling is strong and has to be tackled very carefully to avoid numerical instabilities, (3) "naive" boundary conditions on the artificial boundaries induce spurious reflection phenomena.

Simulation of valves, either at the outflow of the cardiac chambers or in veins, is another example of difficult fluid-structure problems arising in blood flows. In addition, very large displacements and changes of topology (contact problems) have to be handled in those cases.

Due to stability reasons, it seems impossible to successfully apply in hemodynamics the explicit coupling schemes used in other fluid-structure problems, like aeroelasticity. As a result, fluid-structure interaction in biological flows raise new challenging issues in scientific computing and numerical analysis : new schemes have to be developed and analyzed.

We have proposed and analyzed over the last few years several efficient fluid-structure interaction algorithms. This topic remains very active. We are now using these algorithms to address inverse problems in blood flows to make patient specific simulations (for example, estimation of artery wall stiffness from medical imaging).

3.1.2. Aerosol

Complex two-phase fluids can be modeled in many different ways. Eulerian models describe both phases by physical quantities such as the density, velocity or energy of each phase. In the mixed fluid-kinetic models, the biphasic fluid has one dispersed phase, which is constituted by a spray of droplets, with a possibly variable size, and a continuous classical fluid.

This type of model was first introduced by Williams [36] in the frame of combustion. It was later used to develop the Kiva code [26] at the Los Alamos National Laboratory, or the Hesione code [31], for example. It has a wide range of applications, besides the nuclear setting: diesel engines, rocket engines [29], therapeutic sprays, *etc.* One of the interests of such a model is that various phenomena on the droplets can be taken into account with an accurate precision: collision, breakups, coagulation, vaporization, chemical reactions, *etc.*, at the level of the droplets.

The model usually consists in coupling a kinetic equation, that describes the spray through a probability density function, and classical fluid equations (typically Navier-Stokes). The numerical solution of this system relies on the coupling of a method for the fluid equations (for instance, a finite volume method) with a method fitted to the spray (particle method, Monte Carlo).

We are mainly interested in modeling therapeutic sprays either for local or general treatments. The study of the underlying kinetic equations should lead us to a global model of the ambient fluid and the droplets, with some mathematical significance. Well-chosen numerical methods can give some tracks on the solutions behavior and help to fit the physical parameters which appear in the models.

3.2. Multiscale modeling

Multiscale modeling is a necessary step for blood and respiratory flows. In this section, we focus on blood flows. Nevertheless, similar investigations are currently carried out on respiratory flows.

3.2.1. Arterial tree modeling

Problems arising in the numerical modeling of the human cardiovascular system often require an accurate description of the flow in a specific sensible subregion (carotid bifurcation, stented artery, *etc.*). The description of such local phenomena is better addressed by means of three-dimensional (3D) simulations, based on the numerical approximation of the incompressible Navier-Stokes equations, possibly accounting for compliant (moving) boundaries. These simulations require the specification of boundary data on artificial boundaries that have to be introduced to delimit the vascular district under study. The definition of such boundary conditions is critical and, in fact, influenced by the global systemic dynamics. Whenever the boundary data is not available from accurate measurements, a proper boundary condition requires a mathematical description of the action of the reminder of the circulatory system on the local district. From the computational point of view, it is not affordable to describe the whole circulatory system keeping the same level of detail. Therefore, this mathematical description relies on simpler models, leading to the concept of *geometrical multiscale* modeling of the circulation [32]. The underlying idea consists in coupling different models (3D, 1D or 0D) with a decreasing level of accuracy, which is compensated by their decreasing level of computational complexity.

The research on this topic aims at providing a correct methodology and a mathematical and numerical framework for the simulation of blood flow in the whole cardiovascular system by means of a geometric multiscale approach. In particular, one of the main issues will be the definition of stable coupling strategies between 3D and reduced order models.

To model the arterial tree, a standard way consists of imposing a pressure or a flow rate at the inlet of the aorta, *i.e.* at the network entry. This strategy does not allow to describe important features as the overload in the heart caused by backward traveling waves. Indeed imposing a boundary condition at the beginning of the aorta artificially disturbs physiological pressure waves going from the arterial tree to the heart. The only way to catch this physiological behavior is to couple the arteries with a model of heart, or at least a model of left ventricle.

A constitutive law for the myocardium, controlled by an electrical command, has been developed in the CardioSense3D project [0]. One of our objectives is to couple artery models with this heart model.

A long term goal is to achieve 3D simulations of a system including heart and arteries. One of the difficulties of this very challenging task is to model the cardiac valves. To this purpose, we investigate a mix of arbitrary Lagrangian Eulerian and fictitious domain approaches or x-fem strategies, or simplified valve models based on an immersed surface strategy.

3.2.2. Heart perfusion modeling

The heart is the organ that regulates, through its periodical contraction, the distribution of oxygenated blood in human vessels in order to nourish the different parts of the body. The heart needs its own supply of blood to work. The coronary arteries are the vessels that accomplish this task. The phenomenon by which blood reaches myocardial heart tissue starting from the blood vessels is called in medicine perfusion. The analysis of heart perfusion is an interesting and challenging problem. Our aim is to perform a three-dimensional dynamical numerical simulation of perfusion in the beating heart, in order to better understand the phenomena linked to perfusion. In particular the role of the ventricle contraction on the perfusion of the heart is investigated as well as the influence of blood on the solid mechanics of the ventricle. Heart perfusion in fact implies the interaction between heart muscle and blood vessels, in a sponge-like material that contracts at every heartbeat via the myocardium fibers.

Despite recent advances on the anatomical description and measurements of the coronary tree and on the corresponding physiological, physical and numerical modeling aspects, the complete modeling and simulation of blood flows inside the large and the many small vessels feeding the heart is still out of reach. Therefore, in order to model blood perfusion in the cardiac tissue, we must limit the description of the detailed flows at a given space scale, and simplify the modeling of the smaller scale flows by aggregating these phenomena into macroscopic quantities, by some kind of "homogenization" procedure. To that purpose, the modeling of the fluid-solid coupling within the framework of porous media appears appropriate.

Poromechanics is a simplified mixture theory where a complex fluid-structure interaction problem is replaced by a superposition of both components, each of them representing a fraction of the complete material at every point. It originally emerged in soils mechanics with the work of Terzaghi [35], and Biot [27] later gave a description of the mechanical behavior of a porous medium using an elastic formulation for the solid matrix, and Darcy's law for the fluid flow through the matrix. Finite strain poroelastic models have been proposed (see references in [28]), albeit with *ad hoc* formulations for which compatibility with thermodynamics laws and incompressibility conditions is not established.

3.2.3. Tumor and vascularization

The same way the myocardium needs to be perfused for the heart to beat, when it has reached a certain size, tumor tissue needs to be perfused by enough blood to grow. It thus triggers the creation of new blood vessels (angiogenesis) to continue to grow. The interaction of tumor and its micro-environment is an active field of research. One of the challenges is that phenomena (tumor cell proliferation and death, blood vessel adaptation, nutrient transport and diffusion, etc) occur at different scales. A multi-scale approach is thus being developed to tackle this issue. The long term objective is to predict the efficiency of drugs and optimize therapy of cancer.

[0] http://www-sop.inria.fr/CardioSense3D/

3.2.4. Respiratory tract modeling

We aim at developing a multiscale model of the respiratory tract. Intraprenchymal airways distal from generation 7 of the tracheabronchial tree (TBT), which cannot be visualized by common medical imaging techniques, are modeled either by a single simple model or by a model set according to their order in TBT. The single model is based on straight pipe fully developed flow (Poiseuille flow in steady regimes) with given alveolar pressure at the end of each compartment. It will provide boundary conditions at the bronchial ends of 3D TBT reconstructed from imaging data. The model set includes three serial models. The generation down to the pulmonary lobule will be modeled by reduced basis elements. The lobular airways will be represented by a fractal homogenization approach. The alveoli, which are the gas exchange loci between blood and inhaled air, inflating during inspiration and deflating during expiration, will be described by multiphysics homogenization.

4. Application Domains

4.1. Blood flows

Cardiovascular diseases like atherosclerosis or aneurysms are a major cause of mortality. It is generally admitted that a better knowledge of local flow patterns could improve the treatment of these pathologies (although many other biophysical phenomena obviously take place in the development of such diseases). In particular, it has been known for years that the association of low wall shear stress and high oscillatory shear index give relevant indications to localize possible zones of atherosclerosis. It is also known that medical devices (graft or stent) perturb blood flows and may create local stresses favorable with atherogenesis. Numerical simulations of blood flows can give access to this local quantities and may therefore help to design new medical devices with less negative impacts. In the case of aneurysms, numerical simulations may help to predict possible zones of rupture and could therefore give a guide for treatment planning.

In clinical routine, many indices are used for diagnosis. For example, the size of a stenosis is estimated by a few measures of flow rate around the stenosis and by application of simple fluid mechanics rules. In some situations, for example in the case a sub-valvular stenosis, it is known that such indices often give false estimations. Numerical simulations may give indications to define new indices, simple enough to be used in clinical exams, but more precise than those currently used.

It is well-known that the arterial circulation and the heart (or more specifically the left ventricle) are strongly coupled. Modifications of arterial walls or blood flows may indeed affect the mechanical properties of the left ventricle. Numerical simulations of the arterial tree coupled to the heart model could shed light on this complex relationship.

One of the goals of the REO team is to provide various models and simulation tools of the cardiovascular system. The scaling of these models will be adapted to the application in mind: low resolution for modeling the global circulation, high resolution for modeling a small portion of vessel.

4.2. Respiratory tracts

Breathing, or "external" respiration ("internal" respiration corresponds to cellular respiration) involves gas transport though the respiratory tract with its visible ends, nose and mouth. Air streams then from the pharynx down to the trachea. Food and drink entry into the trachea is usually prevented by the larynx structure (epiglottis). The trachea extends from the neck into the thorax, where it divides into right and left main bronchi, which enter the corresponding lungs (the left being smaller to accommodate the heart). Inhaled air is then convected in the bronchus tree which ends in alveoli, where gaseous exchange occurs. Surfactant reduces the surface tension on the alveolus wall, allowing them to expand. Gaseous exchange relies on simple diffusion on a large surface area over a short path between the alveolus and the blood capillary under concentration gradients between alveolar air and blood. The lungs are divided into lobes (three on the right, two on the left) supplied by lobar bronchi. Each lobe of the lung is further divided into segments (ten segments of the right lung and eight of the left). Inhaled air contains dust and debris, which must be filtered, if possible, before they reach the alveoli. The tracheobronchial tree is lined by a layer of sticky mucus, secreted by the epithelium. Particles which hit the side wall of the tract are trapped in this mucus. Cilia on the epithelial cells move the mucous continually towards the nose and mouth.

Each lung is enclosed in a space bounded below by the diaphragm and laterally by the chest wall and the mediastinum. The air movement is achieved by alternately increasing and decreasing the chest pressure (and volume). When the airspace transmural pressure rises, air is sucked in. When it decreases, airspaces collapse and air is expelled. Each lung is surrounded by a pleural cavity, except at its hilum where the inner pleura give birth to the outer pleura. The pleural layers slide over each other. The tidal volume is nearly equal to 500 ml.

The lungs may fail to maintain an adequate supply of air. In premature infants surfactant is not yet active. Accidental inhalation of liquid or solid and airway infection may occur. Chronic obstructive lung diseases and lung cancers are frequent pathologies and among the three first death causes in France.

One of the goals of REO team in the ventilation field is to visualize the airways (virtual endoscopy) and simulate flow in image-based 3D models of the upper airways (nose, pharynx, larynx) and the first generations of the tracheobronchial tree (trachea is generation 0), whereas simple models of the small bronchi and alveoli are used (reduced-basis element method, fractal homogenization, multiphysics homogenization, lumped parameter models), in order to provide the flow distribution within the lung segments.

4.3. Cardiac electrophysiology

The purpose is to simulate the propagation of the action potential in the heart. A lot of works has already been devoted to this topic in the literature (see *e.g.* [30], [34], [33] and the references therein), nevertheless there are only very few studies showing realistic electrocardiograms obtained from partial differential equations models. Our goal is to find a compromise between two opposite requirements: on the one hand, we want to use predictive models, and therefore models based on physiology, on the other hand, we want to use models simple enough to be parametrized (in view of patient-specific simulations). One of the goal is to use our ECG simulator to address the inverse problem of electrocardiology. In collaboration with the Macs/M3disym project-team, we are interested in the electromechanical coupling in the myocardium. We are also interested in various clinical and industrial issues related to cardiac electrophysiology, in particular the simulation of experimental measurement of the field potential of cardiac stem cells in multi-electrode arrays.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

ERC consolidator grant MoDeLLiver (I Vignon-Clementel).

6. New Results

6.1. Numerical methods for fluid mechanics and application to blood flows

Participants: Irene Vignon-Clementel

If abdominal aortic aneurysms (AAA) are known to be associated with altered morphology and blood flow, intraluminal thrombus deposit and clinical symptoms, the growth mechanisms are yet to be fully understood. In this retrospective longitudinal study of 138 scans, morphological analysis and blood flow simulations for 32 patients with clinically diagnosed AAAs and several follow-up CT-scans, are performed and compared to 9 control subjects [21]. Local correlations between hemodynamic metrics and AAA growth are also explored. Finally, high-risk predictors trained with successively clinical, morphological, hemodynamic and all data, and their link to the AAA evolution are built from supervise learning.

In this paper [19], we perform a verification study of the Coupled-Momentum Method (CMM), a 3D fluid-structure interaction (FSI) model which uses a thin linear elastic membrane and linear kinematics to describe the mechanical behavior of the vessel wall. The verification of this model is done using Womersley's deformable wall analytical solution for pulsatile flow in a semi-infinite cylindrical vessel. This solution is, under certain premises, the analytical solution of the CMM and can thus be used for model verification. For the numerical solution, we employ an impedance boundary condition to define a reflection-free outflow boundary condition and thus mimic the physics of the analytical solution, which is defined on a semi-infinite domain. We first provide a rigorous derivation of Womersley's deformable wall theory via scale analysis. We then illustrate different characteristics of the analytical solution and verification tests comparing the CMM with Womersley's theory.

Superior cavopulmonary circulation can be achieved by either the Hemi-Fontan or Bidirectional Glenn connection. Debate remains as to which results in best hemodynamic results. In [22], adopting patient-specific multiscale computational modeling, we examined both the local dynamics and global physiology to determine if surgical choice can lead to different hemodynamic outcomes.

6.2. Liver biomedical research

Participants: Irene Vignon-Clementel, Nicolas Golse

Nicolas Golse, as part of his medical activity has published 7 articles in 2019 that are not reported here.

The hepatic volume gain following resection is essential for clinical recovery. Previous studies have focused on cellular regeneration. In [13], the study aims to explore the rate of hepatic regeneration of the porcine liver following major resection, highlighting estimates of the early microarchitectural changes that occur during the cellular regeneration. Nineteen large white pigs had 75% resection with serial measurements of the hepatic volume, density, blood flow, and architectural changes that are analyzed at different days to highlight differences pre-resection and in the days following resection.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

Participants: Lazaros Papamanolis, Irene Vignon-Clementel [local coordinator].

Contract with ESIEE (H. Talbot, L. Najman) for collaboration with the Heartflow company.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

Irene Vignon Clementel is a member of the project iLite (09/16-10/21), RHU-santé grant, a large French hospital-medical research consortium that aims at developing innovations for liver and tissue engineering (Inria PI: Dirk Drasdo).

8.1.2. APHP-Inria collaboration

Participants: Nour Bou Saleh, Quentin Nicolas, Nicolas Golse, Irene Vignon-Clementel [local coordinator].

Collaboration with Eric Vibert (APHP - Inserm U1193) for cosupervision of surgery interns (N. Bousaleh, D. Dousse) and engineering intern (Q Nicolas) in the context of the APHP-Inria PhD of N. Golse, on liver modeling and ICG fluorescence.

8.2. European Initiatives

8.2.1. Collaborations in European Programs, Except FP7 & H2020

SimInhale COST Action MP1404, a pan-European network of experts in the field of inhaled medicine, coordinated by Prof. Stavros Kassinos, end: 2019 (http://www.siminhale-cost.eu).

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners

Collaboration with :

- Prof. Pal Dag Line from U. of Oslo, Oslo hospital U. with E. Vibert (APHP, Inserm) - N. Golse, I. Vignon-Clementel
- CHUM Centre Hospitalier de l'Université de Montreal (G Soulez and colleagues) - F. Joly (Inria), I. Vignon-Clementel

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. General Chair, Scientific Chair

I. Vignon-Clementel: Co-chair of the international conference VPH2020

9.1.1.2. Member of the Organizing Committees

Irene Vignon-Clementel: session to foster collaboration between scientists and medical doctors at the Inria/CentraleSupelec/Ap-Hp meeting, March 28th, Palaiseau, France

9.1.2. Scientific Events: Selection

9.1.2.1. Member of the Conference Program Committees

- I. Vignon-Clementel
  - Programme committee member, Computational and Mathematical Biomedical Engineering Conference
  - Conference steering committee, International Conference on Engineering Frontiers in Pediatric and Congenital Heart Disease

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Irene Vignon-Clementel: Associate Editor of the International Journal for Numerical Methods in Biomedical Engineering

9.1.3.2. Reviewer - Reviewing Activities

Irene Vignon-Clementel for several journals such as Annals Biomed Eng, BMMB, Med. engineering &Physics.

9.1.4. Conferences and seminars

- Nicolas Golse
  - talk, 15th French congress of hepato-biliary surgery (ACHBT-SFCD), Hotel Newport Bay, Paris, 27-29 Novembre 2019
  - As part of his clinical research, N. Golse presented 7 conference talks and 5 posters
- Nour Bou Saleh
  - poster: 15th French congress of hepato-biliary surgery (ACHBT-SFCD), Hotel Newport Bay, Paris, 27-29 Novembre 2019
- Florian Joly (supported by Reo)
  - talk, ESB2019, European Society of Biomechanics Conference, July 7-10 2019, Vienna, Austria
  - talk, SB2019, Congress of the French society of biomechanics October 28-30 2019, Poitiers, France
- Irene Vignon-Clementel
  - Invited talk, Day on AI and health, School of medicine, U. Paris-Saclay, Nov 25th, Kremlin-Bicetre, France
  - Invited talk, National conference of Biomechanics, Oct 28-30, Poitiers, France
  - Round table on Healthcare 4.0, IMT, Oct 15th, St Etienne, France
  - Seminar, U. of Tokyo, June 13th, Tokyo, Japan
  - Keynote speaker, CMBE conference, June 10th-12th, Tohoku University, Sendai City, Japan
  - Keynote speaker, FIMH, June 6-8th, Bordeaux, France
  - Invited talk, International Conference on Engineering Frontiers in Pediatric and Congenital Heart Disease, May 9th-11th, Philadelphia, USA

9.1.5. Leadership within the Scientific Community

Irene Vignon-Clementel: VPHi board meeting, Oct 4th, Paris, France

9.1.6. Scientific Expertise

Irene Vignon-Clementel: Reviewer for several institutions: Helmholtz Institute (Germany), Israeli Science Foundation, Medical research council (UK), Heart research (UK)

9.1.7. Research Administration

- I. Vignon-Clementel
  - Working group to foster innovation for the Inserm strategic plan
  - Technology grant committee (Commission de développement technologique), Inria Paris center
  - Committee member for PhD students at Inria - Commission consultative des doctorants
  - Mediator between PhD students and their supervisors for Inria Paris

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence:
Irene Vignon-Clementel

- Numerical Methods for Ordinary Differential Equations, 24h ETD, L3, UPMC (2nd semester 2018-2019)
- Invited lecture as part of the undergraduate continuum mechanics class at AgroParisTech (engineering school), France

Master:

- Irene Vignon-Clementel
  - Fall school on Ventilation and Music, Cargèse, France (2h30 EDT)
  - MEC 550 - Biofluid Mechanics and Mass Transport, Ecole Polytechnique (engineering school), France (1h30 EDT)
  - Doctoral school cours Innovations thérapeutiques: du fondamental à l'appliqué, bioengineering module at ENS Cachan, France (1h30 EDT)
  - Master 2 Sciences Chirurgicales de l'Université Paris Sud, France (1h30 EDT)
- Nicolas Golse
  - Master 1 Sciences, Techniques et Santé, Faculté Médecine Kremlin Bicêtre (4.5h EDT)

University diploma (DU)
Nicolas Golse:

- DU of hepatobiliary and pancreatic surgery, 6 sessions (9h EDT)

9.2.2. Supervision

PhD in progress : Nicolas Golse, Apports de la modélisation anatomique et hémodynamique du foie dans l'anticipation, la réalisation et l'enseignement de la chirurgie hépatique, since 10/2016, co-directed by Eric Vibert, Irene Vignon-Clementel, Stephane Cotin

Defended PhD: Chen-Yu Chiang, Transport in biological systems. Monolithic method for fluid-structure interaction, Sorbonne U., 11 January 2019, co-directed by Marc Thiriet, Olivier Pironneau and Toni Sheu (National Taiwan U)

9.2.3. Juries

Irene Vignon-Clementel

- Referee for the PhD defense of Miguel Veira, KCL, London, UK, June 18th
- Jury for research projects (undergraduates), CentraleSupelec, June 4th & Nov 4th & 6th
- Hiring committee for engineer position, U. of Bordeaux, France, May 5th
- Jury member (invited), PhD defense of F. Joly, CHUM, Canada, April 15th
- Referee for the PhD defense of Mathias Braun, U. of Bordeaux, France, April 4th
- President of the jury for the HDR of Paul VanLiedekerke, Sorbonne U., France, March 26th

9.3. Popularization

9.3.1. Interventions

- N Golse (MD)
  - conference at Ecole des Mines, 14th October 2019, "Healthcare simulation"
- Irene Vignon-Clementel
  - High school forum on scientific career, 7th December, Lycee Corneille, La Celle St Cloud
  - Junior high school intervention 'women in engineering', 22nd February, Ville d'Avray
  - High school forum on scientific career, 19th January, Lycee Corneille, La Celle St Cloud
  - High school intervention, 15th January, Blanche de Castille, Le Chesnay

9.3.2. Internal action

Interviews of Irene Vignon-Clementel by journalist for Inria, 13th March and 25th Nov (see Inria website articles)

10. Bibliography

Major publications by the team in recent years

[1] L. BOUDIN, L. DESVILLETTES, C. GRANDMONT, A. MOUSSA.*Global existence of solutions for the coupled Vlasov and Navier-Stokes equations*, in "Differential and integral equations", November 2009, vol. 22, n° 11-12, p. 1247-1271, https://hal.archives-ouvertes.fr/hal-00331895

[2] L. BOUDIN, B. GREC, F. SALVARANI.*A mathematical and numerical analysis of the Maxwell-Stefan diffusion equations*, in "Discrete and Continuous Dynamical Systems - Series B", 2012, vol. 17, n° 5, p. 1427-1440 [*DOI :* 10.3934/DCDSB.2012.17.1427], https://hal.archives-ouvertes.fr/hal-00490511

[3] M. BOULAKIA, S. CAZEAU, M. A. FERNÁNDEZ, J.-F. GERBEAU, N. ZEMZEMI.*Mathematical Modeling of Electrocardiograms: A Numerical Study*, in "Annals of Biomedical Engineering", 2010, vol. 38, n° 3, p. 1071-1097 [*DOI :* 10.1007/S10439-009-9873-0], https://hal.inria.fr/inria-00400490

[4] M. BOULAKIA, S. GUERRERO.*Regular solutions of a problem coupling a compressible fluid and an elastic structure*, in "Journal de Mathématiques Pures et Appliquées", 2010, vol. 94, n° 4, p. 341-365 [*DOI :* 10.1016/J.MATPUR.2010.04.002], https://hal.inria.fr/hal-00648710

[5] J. CHRISTOPHE, T. ISHIKAWA, N. MATSUKI, Y. IMAI, K. TAKASE, M. THIRIET, T. YAMAGUCHI.*Patient-specific morphological and blood flow analysis of pulmonary artery in the case of severe deformations of the lung due to pneumothorax*, in "Journal of Biomechanical Science and Engineering", 2010, vol. 5, n° 5, p. 485-498, https://hal.inria.fr/inria-00543090

[6] M. A. FERNÁNDEZ, J. MULLAERT, M. VIDRASCU.*Explicit Robin-Neumann schemes for the coupling of incompressible fluids with thin-walled structures*, in "Computer Methods in Applied Mechanics and Engineering", 2013, vol. 267, p. 566-593 [*DOI :* 10.1016/J.CMA.2013.09.020], https://hal.inria.fr/hal-00784903

[7] J.-F. GERBEAU, D. LOMBARDI.*Approximated Lax Pairs for the Reduced Order Integration of Nonlinear Evolution Equations*, in "Journal of Computational Physics", May 2014, vol. 265, p. 246-269 [*DOI :* 10.1016/J.JCP.2014.01.047], https://hal.inria.fr/hal-00933172

[8] C. GRANDMONT, M. HILLAIRET.*Existence of global strong solutions to a beam-fluid interaction system*, in "Archive for Rational Mechanics and Analysis", 2016 [*DOI :* 10.1007/S00205-015-0954-Y], https://hal.inria.fr/hal-01138736

[9] P. MOIREAU, C. BERTOGLIO, N. XIAO, C. A. FIGUEROA, C. TAYLOR, D. CHAPELLE, J.-F. GERBEAU.*Sequential identification of boundary support parameters in a fluid-structure vascular model using patient image data*, in "Biomechanics and Modeling in Mechanobiology", July 2012, vol. 12, n° 3, p. 475-496 [*DOI :* 10.1007/S10237-012-0418-3], https://hal.inria.fr/hal-00760703

[10] S. PANT, B. FABRÈGES, J.-F. GERBEAU, I. VIGNON-CLEMENTEL.*A methodological paradigm for patient-specific multi-scale CFD simulations: from clinical measurements to parameter estimates for individual analysis*, in "International Journal for Numerical Methods in Biomedical Engineering", December 2014, vol. 30, n° 12, p. 1614–1648 [*DOI :* 10.1002/CNM.2692], https://hal.inria.fr/hal-01093879

[11] I. Vignon-Clementel, A. Marsden, J. Feinstein. *A Primer on Computational Simulation in Congenital Heart Disease for the Clinician*, in "Progress in Pediatric Cardiology", 2010, vol. 30, n° 1-2, p. 3-13, Fondation Leducq [*DOI :* 10.1016/j.ppedcard.2010.09.002], https://hal.inria.fr/inria-00542957

Publications of the year

Doctoral Dissertations and Habilitation Theses

[12] L. Boilevin-Kayl. *Modeling and numerical simulation of implantable cardiovascular devices*, Sorbonne Université, July 2019, https://hal.inria.fr/tel-02217259

Articles in International Peer-Reviewed Journal

[13] M. Bekheit, P. O. Bucur, C. Audebert, E. Miquelestorena-Standley, I. Vignon-Clementel, E. Vibert. *Kinetics of Hepatic Volume Evolution and Architectural Changes after Major Resection in a Porcine Model*, in "European Surgical Research", August 2019, vol. 60, n° 1-2, p. 31-44 [*DOI :* 10.1159/000491691], https://hal.inria.fr/hal-02413756

[14] L. Boilevin-Kayl, M. A. Fernández, J.-F. Gerbeau. *A loosely coupled scheme for fictitious domain approximations of fluid-structure interaction problems with immersed thin-walled structures*, in "SIAM Journal on Scientific Computing", February 2019, vol. 41, n° 2, p. 351-374 [*DOI :* 10.1137/18M1192779], https://hal.inria.fr/hal-01811290

[15] L. Boilevin-Kayl, M. A. Fernández, J.-F. Gerbeau. *Numerical methods for immersed FSI with thin-walled structures*, in "Computers and Fluids", January 2019, vol. 179, p. 744-763 [*DOI :* 10.1016/j.compfluid.2018.05.024], https://hal.inria.fr/hal-01704575

[16] A. Bondesan, L. Boudin, B. Grec. *A numerical scheme for a kinetic model for mixtures in the diffusive limit using the moment method*, in "Numerical Methods for Partial Differential Equations", May 2019, vol. 35, n° 3, p. 1184-1205, forthcoming [*DOI :* 10.1002/num.22345], https://hal.archives-ouvertes.fr/hal-01727725

[17] L. Boudin, B. Grec, V. Pavan. *Diffusion models for mixtures using a stiff dissipative hyperbolic formalism*, in "Journal of Hyperbolic Differential Equations", August 2019, vol. 16, n° 2, p. 293–312 [*DOI :* 10.1142/S0219891619500115], https://hal.archives-ouvertes.fr/hal-01786824

[18] M. A. Fernández, M. Landajuela. *Splitting schemes and unfitted mesh methods for the coupling of an incompressible fluid with a thin-walled structure*, in "IMA Journal of Numerical Analysis", January 2019 [*DOI :* 10.1093/imanum/dry098], https://hal.inria.fr/hal-01309462

[19] V. Filonova, C. J. Arthurs, I. Vignon-Clementel, C. A. Figueroa. *Verification of the Coupled-Momentum Method with Womersley's Deformable Wall Analytical Solution*, in "International Journal for Numerical Methods in Biomedical Engineering", October 2019 [*DOI :* 10.1002/cnm.3266], https://hal.inria.fr/hal-02414672

[20] 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", April 2019, vol. 66, n° 4, p. 946-955 [*DOI :* 10.1109/TBME.2018.2865667], https://hal.archives-ouvertes.fr/hal-01869264

[21] F. JOLY, G. SOULEZ, S. LESSARD, C. KAUFFMANN, I. VIGNON-CLEMENTEL.*A cohort longitudinal study identifies morphology and hemodynamics predictors of abdominal aortic aneurysm growth*, in "Annals of Biomedical Engineering", October 2019, forthcoming [*DOI :* 10.1007/S10439-019-02375-1], https://hal.archives-ouvertes.fr/hal-02302464

[22] E. KUNG, C. CORSINI, A. L. MARSDEN, I. VIGNON-CLEMENTEL, G. PENNATI, R. S. FIGLIOLA, T.-Y. HSIA, A. TAYLOR, S. KHAMBADKONE, S. SCHIEVANO, M. DE LEVAL, E. BOVE, A. DORFMAN, G. H. BAKER, A. HLAVACEK, F. MIGLIAVACCA, G. DUBINI, J. D. MCGREGOR.*Multiscale Modeling of Superior Cavopulmonary Circulation: Hemi-Fontan and Bidirectional Glenn Are Equivalent*, in "Seminars in Thoracic and Cardiovascular Surgery", September 2019 [*DOI :* 10.1053/J.SEMTCVS.2019.09.007], https://hal.inria.fr/hal-02414734

[23] A. THIS, L. BOILEVIN-KAYL, M. A. FERNÁNDEZ, J.-F. GERBEAU.*Augmented Resistive Immersed Surfaces valve model for the simulation of cardiac hemodynamics with isovolumetric phases*, in "International Journal for Numerical Methods in Biomedical Engineering", May 2019, forthcoming [*DOI :* 10.1002/CNM.3223], https://hal.inria.fr/hal-01944798

[24] A. THIS, H. G. MORALES, O. BONNEFOUS, M. A. FERNÁNDEZ, J.-F. GERBEAU.*A pipeline for image based intracardiac CFD modeling and application to the evaluation of the PISA method*, in "Computer Methods in Applied Mechanics and Engineering", 2019, vol. 358, n^o 1 [*DOI :* 10.1016/J.CMA.2019.112627], https://hal.archives-ouvertes.fr/hal-02142416

Other Publications

[25] L. BOUDIN, C. GRANDMONT, B. GREC, S. MARTIN, A. MECHERBET, F. NOËL.*Fluid-kinetic modelling for respiratory aerosols with variable size and temperature*, April 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02092574

References in notes

[26] A. AMSDEN, P. O'ROURKE, T. BUTLER.*A computer program for chemically reactive flows with sprays*, Los Alamos National Laboratory, 1989, n^o LA-11560-MS

[27] M. A. BIOT.*Theory of propagation of elastic waves in a fluid-saturated porous solid. II higher frequency range*, in "J. Acoust. Soc. Am.", 1956, vol. 28, p. 179–191

[28] D. CHAPELLE, J. SAINTE-MARIE, J.-F. GERBEAU, I. VIGNON-CLEMENTEL.*A poroelastic model valid in large strains with applications to perfusion in cardiac modeling*, in "Computational Mechanics", 2010, vol. 46, n^o 1, p. 91-101 [*DOI :* 10.1007/S00466-009-0452-X]

[29] J. DUPAYS, Y. FABIGNON, P. VILLEDIEU, G. LAVERGNE, G. ESTIVALEZES.*Some aspects of two phase flows in solid propellant rocket motors, in Solid propellant chemistry, combustion and interior ballistics*, in "Progress in Astronautics and Aeronautics", V. YANG, T. BRILL, W. PEN (editors), Academic Press, 2000, vol. 185

[30] G. LINES, P. GROTTUM, A. TVEITO.*Modeling the electrical activity of the heart. A bidomain model of the ventricles embedded in a torso*, in "Comput. Visual. Sci.", 2003, vol. 5, p. 195-213

[31] R. MOTTE.*A numerical method for solving particle-fluid equations*, in "Trends in numerical and physical modeling for industrial multiphase flows", Cargèse, France, 2000

[32] A. QUARTERONI, S. RAGNI, A. VENEZIANI.*Coupling between lumped and distributed models for blood flow problems*, in "Comput. Visual Sci.", 2001, vol. 4, p. 111–124

[33] F. SACHSE.*Computational Cardiology: Modeling of Anatomy, Electrophysiology, and Mechanics*, Springer-Verlag, 2004

[34] J. SUNDNES, G. LINES, X. CAI, B. NIELSEN, K.-A. MARDAL, A. TVEITO.*Computing the electrical activity in the heart*, Springer-Verlag, 2006

[35] K. TERZAGHI.*Theoretical Soil Mechanics*, John Wiley and Sons, New-York, 1943

[36] F. WILLIAMS.*Combustion theory*, 2nd, Benjamin Cummings, 1985

Project-Team RITS

Robotics & Intelligent Transportation Systems

RESEARCH CENTER
Paris

THEME
Robotics and Smart environments

Table of contents

Project-Team RITS

Creation of the Team: 2014 February 17, updated into Project-Team: 2015 July 01

Keywords:

Computer Science and Digital Science:

A1.5. - Complex systems
A1.5.1. - Systems of systems
A1.5.2. - Communicating systems
A2.3. - Embedded and cyber-physical systems
A3.4. - Machine learning and statistics
A3.4.1. - Supervised learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.8. - Deep learning
A5.3. - Image processing and analysis
A5.3.4. - Registration
A5.4. - Computer vision
A5.4.1. - Object recognition
A5.4.4. - 3D and spatio-temporal reconstruction
A5.4.5. - Object tracking and motion analysis
A5.4.6. - Object localization
A5.5.1. - Geometrical modeling
A5.9. - Signal processing
A5.10. - Robotics
A5.10.2. - Perception
A5.10.3. - Planning
A5.10.4. - Robot control
A5.10.5. - Robot interaction (with the environment, humans, other robots)
A5.10.6. - Swarm robotics
A5.10.7. - Learning
A6. - Modeling, simulation and control
A6.1. - Methods in mathematical modeling
A6.2.3. - Probabilistic methods
A6.2.6. - Optimization
A6.4.1. - Deterministic control
A6.4.3. - Observability and Controlability
A6.4.4. - Stability and Stabilization
A8.6. - Information theory
A8.9. - Performance evaluation
A9.2. - Machine learning
A9.5. - Robotics
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B5.2.1. - Road vehicles
B5.6. - Robotic systems
B6.6. - Embedded systems
B7.1.2. - Road traffic
B7.2. - Smart travel
B7.2.1. - Smart vehicles
B7.2.2. - Smart road
B9.5.6. - Data science

1. Team, Visitors, External Collaborators

Research Scientists

Fawzi Nashashibi [Team leader, Inria, Senior Researcher, HDR]
Guy Fayolle [Inria, Emeritus]
Jean-Marc Lasgouttes [Inria, Researcher]
Gérard Le Lann [Inria, Emeritus]
Anne Verroust-Blondet [Inria, Researcher, HDR]

External Collaborator

Itheri Yahiaoui [Université de Reims Champagne-Ardennes]

Technical Staff

Mohammad Abualhoul [Inria, Engineer, until Jun 2019]
Said Alexander Alvarado Marin [Inria, Engineer, from Mar 2019]
Sylia Baraka [Inria, Engineer, from Dec 2019]
Pierre Bourre [Inria, Engineer, until Sep 2019]
Pierre de Beaucorps [Inria, Engineer, from Jul 2019]
Raoul de Charette [Inria, Engineer]
Carlos Flores [Inria, Engineer, until Jul 2019]
Kathia Melbouci [Inria, Engineer, from Mar 2019]
Danut Ovidiu Pop [Inria, Engineer, from Sep 2019]
Armand Yvet [Inria, Engineer]

PhD Students

Pierre de Beaucorps [Inria, until Jun 2019]
Farouk Ghallabi [CIFRE Renault]
Maximilian Jaritz [CIFRE Valeo]
Imane Mahtout [CIFRE Renault]
Kaouther Messaoud [Inria]
Fabio Pizzati [University of Bologna, from Nov 2019]
Renaud Poncelet [ENS Rennes]
Danut Ovidiu Pop [Inria, until Aug 2019]
Luis Roldao [CIFRE AKKA Technologies]

Visiting Scientist

Plamen Petrov [Technical University of Sofia, Professor, from Jul 2019 until Sep 2019]

2. Overall Objectives

2.1. Overall Objectives

The focus of the project-team is to develop the technologies linked to Intelligent Transportation Systems (ITS) with the objective to achieve sustainable mobility by the improvement of the safety, the efficiency of road transport according to the recent "Intelligent Vehicle Initiative" launched by the DG Information Society of the European Commission (for "Smarter, Cleaner, and Safer Transport"). More specifically, we want to develop, demonstrate and test some innovative technologies under the framework of LaRA, "La Route Automatisée [0]" which covers all the advanced driver assistance systems (ADAS) and the traffic management systems going all the way to fully automated vehicles.

These developments are all based on the sciences and technologies of information and communications (STIC) and have the objective to bring significant improvements in the road transport sector through incremental or breakthrough innovations. The project-team covers fundamental R&D work on key technologies, applied research to develop techniques that solve specific problems, and demonstrator activities to evaluate and disseminate the results.

The scientific approach is focused on the analysis and optimization of road transport systems through a double approach:

1. the control of individual road vehicles to improve locally their efficiency and safety,
2. the design and control of large transportation systems.

The first theme on vehicle control is broadly based on signal processing and data fusion in order to have a better machine understanding of the situation a vehicle may encounter, and on robotics techniques to control the vehicle in order to help (or replace) the driver to avoid accidents while improving the performance of the vehicle (speed, comfort, mileage, emissions, noise...). The theme also includes software techniques needed to develop applications in a real-time distributed and complex environment with extremely high safety standards. In addition, data must be exchanged between the vehicles; communication protocols have thus to be adapted to and optimized for vehicular networks characteristics (e.g. mobility, road safety requirements, heterogeneity, density), and communication needs (e.g. network latency, quality of service, network security, network access control).

The second theme on modeling and control of large transportation systems is also largely dependent on STIC. The objective, there, is to improve significantly the performance of the transportation system in terms of throughput but also in terms of safety, emissions, energy while minimizing nuisances. The approach is to act on demand management (e.g. through information, access control or road charging) as well as on the vehicles coordination. Communications technologies are essential to implement these controls and are an essential part of the R&D, in particular in the development of technologies for highly dynamic networks.

In order to address those issues simultaneously, RITS is organized into three research axes, each of which being driven by a separate sub-team. The first axis addresses the traditional problem of vehicle guidance and autonomous navigation. The second axis focuses on the large scale deployment and the traffic analysis and modeling. The third axis deals with the problem of telecommunications from two points of view:

- *Technical*: design certified architectures enabling safe vehicle-to-vehicle and vehicle-to-vehicle communications obeying to standards and norm;
- *Fundamental*, design and develop appropriate architectures capable of handling thorny problems of routing and geonetworking in highly dynamic vehicular networks and high speed vehicles.

Of course, these three research sub-teams interact to build intelligent cooperative mobility systems.

[0] LaRA is a Joint Research Unit (JRU) associating three French research teams: Inria's project-team RITS, Mines ParisTech's CAOR and LIVIC.

3. Research Program

3.1. Vehicle guidance and autonomous navigation

Participants: Mohammad Abualhoul, Pranav Agarwal, Said Alexander Alvarado Marin, Sylia Baraka, Pierre de Beaucorps, Fares Bessam, Pierre Bourre, Raoul de Charette, Carlos Flores, Farouk Ghallabi, Manuel Gonzalez, Maximilian Jaritz, Manohar Kv, Imane Mahtout, Kathia Melbouci, Kaouther Messaoud, Fawzi Nashashibi, Fabio Pizzati, Renaud Poncelet, Danut Ovidiu Pop, Luis Roldao, Anne Verroust-Blondet, Leonardo Ward, Itheri Yahiaoui.

There are three basic ways to improve the safety of road vehicles and these ways are all of interest to the project-team. The first way is to assist the driver by giving him better information and warning. The second way is to take over the control of the vehicle in case of mistakes such as inattention or wrong command. The third way is to completely remove the driver from the control loop.

All three approaches rely on information processing. Only the last two involve the control of the vehicle with actions on the actuators, which are the engine power, the brakes and the steering. The research proposed by the project-team is focused on the following elements:

- perception of the environment,
- planning of the actions,
- real-time control.

3.1.1. Perception of the road environment

Participants: Raoul de Charette, Maximilian Jaritz, Farouk Ghallabi, Manohar Kv, Kaouther Messaoud, Fawzi Nashashibi, Fabio Pizzati, Danut Ovidiu Pop, Luis Roldao, Anne Verroust-Blondet, Itheri Yahiaoui.

Either for driver assistance or for fully automated guided vehicle purposes, the first step of any robotic system is to perceive the environment in order to assess the situation around itself. Proprioceptive sensors (accelerometer, gyrometer,...) provide information about the vehicle by itself such as its velocity or lateral acceleration. On the other hand, exteroceptive sensors, such as video camera, laser or GPS devices, provide information about the environment surrounding the vehicle or its localization. Obviously, fusion of data with various other sensors is also a focus of the research.
The following topics are already validated or under development in our team:

- relative ego-localization with respect to the infrastructure, i.e. lateral positioning on the road can be obtained by mean of vision (lane markings) and the fusion with other devices (e.g. GPS);
- global ego-localization by considering GPS measurement and proprioceptive information, even in case of GPS outage;
- road detection by using lane marking detection and navigable free space;
- detection and localization of the surrounding obstacles (vehicles, pedestrians, animals, objects on roads, etc.) and determination of their behavior can be obtained by the fusion of vision, laser or radar based data processing;
- simultaneous localization and mapping as well as mobile object tracking using laser-based and stereovision-based (SLAMMOT) algorithms.

Scene understanding is a large perception problem. In this research axis we have decided to use only computer vision as cameras have evolved very quickly and can now provide much more precise sensing of the scene, and even depth information. Two types of hardware setups were used, namely: monocular vision or stereo vision to retrieve depth information which allow extracting geometry information.

We have initiated several works:

- estimation of the ego motion using monocular scene flow. Although in the state of the art most of the algorithms use a stereo setup, researches were conducted to estimate the ego-motion using a novel approach with a strong assumption.
- bad weather conditions evaluations. Most often all computer vision algorithms work under a transparent atmosphere assumption which assumption is incorrect in the case of bad weather (rain, snow, hail, fog, etc.). In these situations the light ray are disrupted by the particles in suspension, producing light attenuation, reflection, refraction that alter the image processing.
- deep learning for object recognition. New works are being initiated in our team to develop deep learning recognition in the context of heterogeneous data.
- deep learning for vehicle motion prediction.

3.1.2. Planning and executing vehicle actions

Participants: Pierre de Beaucorps, Carlos Flores, Imane Mahtout, Fawzi Nashashibi, Renaud Poncelet, Anne Verroust-Blondet.

From the understanding of the environment, thanks to augmented perception, we have either to warn the driver to help him in the control of his vehicle, or to take control in case of a driverless vehicle. In simple situations, the planning might also be quite simple, but in the most complex situations we want to explore, the planning must involve complex algorithms dealing with the trajectories of the vehicle and its surroundings (which might involve other vehicles and/or fixed or moving obstacles). In the case of fully automated vehicles, the perception will involve some map building of the environment and obstacles, and the planning will involve partial planning with periodical recomputation to reach the long term goal. In this case, with vehicle to vehicle communications, what we want to explore is the possibility to establish a negotiation protocol in order to coordinate nearby vehicles (what humans usually do by using driving rules, common sense and/or non verbal communication). Until now, we have been focusing on the generation of geometric trajectories as a result of a maneuver selection process using grid-based rating technique or fuzzy technique. For high speed vehicles, Partial Motion Planning techniques we tested, revealed their limitations because of the computational cost. The use of quintic polynomials we designed, allowed us to elaborate trajectories with different dynamics adapted to the driver profile. These trajectories have been implemented and validated in the JointSystem demonstrator of the German Aerospace Center (DLR) used in the European project HAVEit, as well as in RITS's electrical vehicle prototype used in the French project ABV. HAVEit was also the opportunity for RITS to take in charge the implementation of the Co-Pilot system which processes perception data in order to elaborate the high level command for the actuators. These trajectories were also validated on RITS's cybercars. However, for the low speed cybercars that have pre-defined itineraries and basic maneuvers, it was necessary to develop a more adapted planning and control system. Therefore, we have developed a nonlinear adaptive control for automated overtaking maneuver using quadratic polynomials and Lyapunov function candidate and taking into account the vehicles kinematics. For the global mobility systems we are developing, the control of the vehicles includes also advanced platooning, automated parking, automated docking, etc. For each functionality a dedicated control algorithm was designed (see publication of previous years).

3.2. Mobile wireless communications for vehicular networks

Participants: Gérard Le Lann, Mohammad Abualhoul, Fawzi Nashashibi.

Wireless communications are expected to play an essential role in ensuring road safety, road efficiency, and driving comfort o. Road safety applications often require relatively short response time and reliable information exchange between neighboring vehicles and road-side units in any road density condition. Because of the performance of the existing radio communications technology largely degrades with the increase of the traffic density, the challenge of designing wireless communications solution suitable for safety applications is enabling reliable communications in highly dense scenarios.

To investigate this open problem and trade-off situations, RITS has been working on medium access control design for the IEEE 802.11p radio communication and the deployment of supportive solutions such as visible light communications and testing the use-cases for extreme traffic conditions and highly dense scenarios. The works have been carried out considering the vehicle behavior such as autonomous and connected vehicles merging, sharing, and convoy forming as platoon scenarios with considering the hard-safety requirements.

Unlike many of the road safety applications, the applications regarding road efficiency and comfort of road users, often require connectivity to the Internet. Based on our expertise in both Internet-based communications in the mobility context and in ITS, we are investigating the use of IPv6 (Internet Protocol version 6 which is going to replace the current version, IPv4, IoT) for vehicular communications, in a combined architecture supporting both V2V and V2I.

Communication contributions at RITS team have been working on channel modeling for both radio and visible light communications, and design of communications mechanisms, especially for security, service discovery, multicast, and Geo-Cast message delivery, and access point selection.

RITS-team has one of the latest certified standard communication hardware and tools supported by the partnership with the YoGoKo Company. All platforms (connected and autonomous vehicles) are equipped with state-of-art communication units On-Board-Units (OBU), where the Rocquencourt site equipped with two stationary Road-Side-Units (RSU) enabling all kind of tests and projects requirements

Below follows a more detailed description of the related research issues.

3.2.1. Regulation study for interoperability tests for cooperative driving

Participants: Mohammad Abualhoul, Fawzi Nashashibi.

The technological advances of autonomous and connected road vehicles have been shown an accelerating pace in the recent years. On the other hand, the regulations for autonomous, or driverless, road vehicles across Europe still deserve much attention and discussion

Therefore, RITS-Inria team plays a key element in one of the European demonstration-based projects (AUTOC-ITS), which aims to contribute to the regulation study for interoperability in the adoption of autonomous driving in European urban nodes. The regulation study done by RITS team and project partners meant to conduct a deployment of Cooperative Intelligent Transport Systems (C-ITS) in Europe by enhancing interoperability for autonomous vehicles [29]. The project activities and RITS contributions will also boost the role of C-ITS as the primary catalyst for any future implementation of autonomous driving scenarios in Europe. The final demonstration of different European partners will require the implementation and preparations of three pilots sites in three major European cities: Paris, Madrid, and Lisbon. Pilot locations in these major cities are chosen to be located along the European Atlantic Corridor for interoperability evaluation.

RITS-Inria is coordinating the French contribution by evaluating the deployment of C-ITS services in the A13-Paris, which belongs to the French part of the Atlantic Corridor.

Team Core contributions:

- Provide up to date feedback to contribute to the present EU and international regulations on autonomous vehicles.
- Build and evaluate the pilots experimentally by deploying fully autonomous vehicles and a Cooperative Intelligent Transport Systems (C-ITS).
- Define and evaluate a safety autonomous driving services, such as:
  - Roadworks warning.
  - Weather conditions.
  - Other hazardous notifications.
- Define and perform communication interoperability tests between deferent partners for different scenarios, messaging and hardware to ensure the compatibility in using the IEEE 802.11p standard.
- Study the extension of the results on large-scale deployment in other European countries.
- Contribute to the European standards organizations such as C-Roads, C-ITS platforms.

AUTOC-ITS project brings the road authorities from France, Span, and Portugal (DGT, ANSR, SANEF) and C-ITS experts from research institutes and universities (Inria, INDRA, UPM, UC, IPN) to carry out a cooperative work and contributes to the C-ITS Platform by bringing answers to the field of automation driving.

3.2.2. V2X radio communications for road safety applications

Participants: Mohammad Abualhoul, Fawzi Nashashibi.

The development work and generating proper components to facilitate communication requirements and to be deployed in different projects scenarios is one of the main ongoing activities by all RITS team members.

There are continuous activities on both theoretical modeling and experimental evaluation of the radio channel characteristics in vehicular networks, especially the radio quality, channel congestion, load allocations, congestion, and bandwidth availability.

Based on our previous expertise and studies, we develop mechanisms for efficient and reliable V2X communications, access point selection, handover algorithms which are especially dedicated to road safety and autonomous driving applications.

3.2.3. Cyberphysical constructs and mobile communications for fully automated networked vehicles

Participant: Gérard Le Lann.

Intelligent vehicular networks (IVNs) are constituents of ITS. IVNs range from platoons with a lead vehicle piloted by a human driver to fully ad-hoc vehicular networks, a.k.a. VANETs, comprising autonomous/automated vehicles. Safety issues in IVNs appear to be the least studied in the ITS domain. The focus of our work is on safety-critical (SC) scenarios, where accidents and fatalities inevitably occur when such scenarios are not handled correctly. In addition to on-board robotics, inter-vehicular radio communications have been considered for achieving safety properties. Since both technologies have known intrinsic limitations (in addition to possibly experiencing temporary or permanent failures), using them redundantly is mandatory for meeting safety regulations. Redundancy is a fundamental design principle in every SC cyber-physical domain, such as, e.g., air transportation. (Optics-based inter-vehicular communications may also be part of such redundant constructs.) The focus of our on-going work is on safety-critical (SC) communications. We consider IVNs on main roads and highways, which are settings where velocities can be very high, thus exacerbating safety problems acceptable delays in the cyber space, and response times in the physical space, shall be very small. Human lives being at stake, such delays and response times must have strict (non-stochastic) upper bounds under worst-case conditions (vehicular density, concurrency and failures). Consequently, we are led to look for deterministic solutions.

Rationale

In the current ITS literature, the term *safety* is used without being given a precise definition. That must be corrected. In our case, a fundamental open question is: what is the exact meaning of *SC communications*? We have devised a definition, referred to as space-time bounds acceptability (STBA) requirements. For any given problem related to SC communications, those STBA requirements serve as yardsticks for distinguishing acceptable solutions from unacceptable ones with respect to safety. In conformance with the above, STBA requirements rest on the following worst-case upper bounds: λ for channel access delays, and Δ for distributed inter-vehicular coordination (message dissemination, distributed agreement).

Via discussions with foreign colleagues, notably those active in the IEEE 802 Committee, we have comforted our early diagnosis regarding existing standards for V2V/V2I/V2X communications, such as IEEE 802.11p and ETSI ITS-G5: they are totally inappropriate regarding SC communications. A major flaw is the choice of CSMA/CA as the MAC-level protocol. Obviously, there cannot be such bounds as λ and Δ with CSMA/CA. Another flaw is the choice of medium-range omnidirectional communications, radio range in the order of 250 m, and interference range in the order of 400 m. Stochastic delays achievable with existing standards are just unacceptable in moderate/worst-case contention conditions. Consider the following setting, not uncommon in many countries: a highway, 3 lanes each direction, dense traffic, i.e. 1 vehicle per 12.5 m. A simple calculation leads to the following result: any vehicle may experience (destructive) interferences from up to 384 vehicles. Even if one assumes some reasonable communications activity ratio, say 25%, one finds that up to 96 vehicles may be contending for channel access. Under such conditions, MAC-level delays and string-wide dissemination/agreement delays achieved by current standards fail to meet the STBA requirements by huge margins.

Reliance on V2I communications via terrestrial infrastructures and nodes, such as road-side units or WiFi hotspots, rather than direct V2V communications, can only lead to poorer results. First, reachability is not guaranteed: hazardous conditions may develop anywhere anytime, far away from a terrestrial node. Second, mixing SC communications and ordinary communications within terrestrial nodes is a violation of the very fundamental segregation principle: SC communications and processing shall be isolated from ordinary communications and processing. Third, security: it is very easy to jam or to spy on a terrestrial node; moreover, terrestrial nodes may be used for launching all sorts of attacks, man-in-the-middle attacks for example. Fourth, delays can only get worse than with direct V2V communications, since transiting via a node inevitably introduces additional latencies. Fifth, the delivery of every SC message must be acknowledged, which exacerbates the latency problems. Sixth, availability: what happens when a terrestrial node fails?

Trying to tweak existing standards for achieving SC communications is vain. That is also unjustified. Clearly, medium-range omnidirectional communications are unjustified for the handling of SC scenarios. By definition, accidents can only involve vehicles that are very close to each other. Therefore, short-range directional communications suffice. The obvious conclusion is that novel protocols and inter-vehicular coordination algorithms based on short-range direct V2V communications are needed. It is mandatory to check whether these novel solutions meet the STBA requirements. Future standards specifically aimed at SC communications in IVNs may emerge from such solutions.

Naming and privacy

Additionally, we are exploring the (re)naming problem as it arises in IVNs. Source and destination names appear in messages exchanged among vehicles. Most often, names are IP addresses or MAC addresses (plate numbers shall not be used for privacy reasons). A vehicle which intends to communicate with some vehicle, denoted V here, must know which name *name(V)* to use in order to reach/designate V. Existing solutions are based on multicasting/broadcasting existential messages, whereby every vehicle publicizes its existence (name and geolocation), either upon request (replying to a Geocast) or spontaneously (periodic beaconing). These solutions have severe drawbacks. First, they contribute to overloading communication channels (leading to unacceptably high worst-case delays). Second, they amount to breaching privacy voluntarily. Why should vehicles reveal their existence and their time dependent geolocations, making tracing and spying much easier? Novel solutions are needed. They shall be such that:

- At any time, a vehicle can assign itself a name that is unique within a geographical zone centered on that vehicle (no third-party involved),
- No linkage may exist between a name and those identifiers (plate numbers, IP/MAC addresses, etc.) proper to a vehicle,
- Different (unique) names can be computed at different times by a vehicle (names can be short-lived or long-lived),
- *name(V)* at UTC time t is revealed only to those vehicles sufficiently close to V at time t, notably those which may collide with V.

We have solved the (re)naming problem in string/cohort formations [34]. Ranks (unique integers in any given string/cohort) are privacy-preserving names, easily computed by every member of a string, in the presence of string membership changes (new vehicles join in, members leave). That problem is open when considering arbitrary clusters of vehicles/strings encompassing multiple lanes.

3.3. Probabilistic modeling for large transportation systems

Participants: Guy Fayolle, Jean-Marc Lasgouttes.

This activity concerns the modeling of random systems related to ITS, through the identification and development of solutions based on probabilistic methods and more specifically through the exploration of links between large random systems and statistical physics. Traffic modeling is a very fertile area of application for this approach, both for macroscopic (fleet management [32], traffic prediction) and for microscopic (movement of each vehicle, formation of traffic jams) analysis. When the size or volume of structures grows (leading to the so-called "thermodynamic limit"), we study the quantitative and qualitative (performance, speed, stability, phase transitions, complexity, etc.) features of the system.

In the recent years, several directions have been explored.

3.3.1. Traffic reconstruction

Large random systems are a natural part of macroscopic studies of traffic, where several models from statistical physics can be fruitfully employed. One example is fleet management, where one main issue is to find optimal ways of reallocating unused vehicles: it has been shown that Coulombian potentials might be an efficient tool to drive the flow of vehicles. Another case deals with the prediction of traffic conditions, when the data comes from probe vehicles instead of static sensors.

While the widely-used macroscopic traffic flow models are well adapted to highway traffic, where the distance between junction is long (see for example the work done by the NeCS team in Grenoble), our focus is on a more urban situation, where the graphs are much denser. The approach we are advocating here is model-less, and based on statistical inference rather than fundamental diagrams of road segments. Using the Ising model or even a Gaussian Random Markov Field, together with the very popular Belief Propagation (BP) algorithm, we have been able to show how real-time data can be used for traffic prediction and reconstruction (in the space-time domain).

This new use of BP algorithm raises some theoretical questions about the ways the make the belief propagation algorithm more efficient:

- find the best way to inject real-valued data in an Ising model with binary variables [36];
- build macroscopic variables that measure the overall state of the underlying graph, in order to improve the local propagation of information [33];
- make the underlying model as sparse as possible, in order to improve BP convergence and quality [35].

3.3.2. Exclusion processes for road traffic modeling

The focus here is on road traffic modeled as a granular flow, in order to analyze the features that can be explained by its random nature. This approach is complementary to macroscopic models of traffic flow (as done for example in the Opale team at Inria), which rely mainly on ODEs and PDEs to describe the traffic as a fluid.

One particular feature of road traffic that is of interest to us is the spontaneous formation of traffic jams. It is known that systems as simple as the Nagel-Schreckenberg model are able to describe traffic jams as an emergent phenomenon due to interaction between vehicles. However, even this simple model cannot be explicitly analyzed and therefore one has to resort to simulation.

One of the simplest solvable (but non trivial) probabilistic models for road traffic is the exclusion process. It lends itself to a number of extensions allowing to tackle some particular features of traffic flows: variable speed of particles, synchronized move of consecutive particles (platooning), use of geometries more complex than plain 1D (cross roads or even fully connected networks), formation and stability of vehicle clusters (vehicles that are close enough to establish an ad-hoc communication system), two-lane roads with overtaking.

The aspect that we have particularly studied is the possibility to let the speed of vehicle evolve with time. To this end, we consider models equivalent to a series of queues where the pair (service rate, number of customers) forms a random walk in the quarter plane $\mathbb{Z}_+^2$.

Having in mind a global project concerning the analysis of complex systems, we also focus on the interplay between discrete and continuous description: in some cases, this recurrent question can be addressed quite rigorously via probabilistic methods.

We have considered in [30] some classes of models dealing with the dynamics of discrete curves subjected to stochastic deformations. It turns out that the problems of interest can be set in terms of interacting exclusion processes, the ultimate goal being to derive hydrodynamic limits after proper scaling. A seemingly new method is proposed, which relies on the analysis of specific partial differential operators, involving variational calculus and functional integration. Starting from a detailed analysis of the Asymmetric Simple Exclusion Process (ASEP) system on the torus $\mathbb{Z}/n\mathbb{Z}$, the arguments a priori work in higher dimensions (ABC, multi-type exclusion processes, etc), leading to systems of coupled partial differential equations of Burgers' type.

3.3.3. Random walks in the quarter plane $\mathbb{Z}_+^2$

This field remains one of the important *violon d'Ingres* in our research activities in stochastic processes, both from theoretical and applied points of view. In particular, it is a building block for models of many communication and transportation systems.

One essential question concerns the computation of stationary measures (when they exist). As for the answer, it has been given by original methods formerly developed in the team (see books and related bibliography). For instance, in the case of small steps (jumps of size one in the interior of $\mathbb{Z}_+^2$), the invariant measure $\{\pi_{i,j}, i,j \geq 0\}$ does satisfy the fundamental functional equation (see [2]):

$$Q(x,y)\pi(x,y) = q(x,y)\pi(x) + \widetilde{q}(x,y)\widetilde{\pi}(y) + \pi_0(x,y). \tag{1}$$

where the unknown generating functions $\pi(x,y), \pi(x), \widetilde{\pi}(y), \pi_0(x,y)$ are sought to be analytic in the region $\{(x,y) \in \mathbb{C}^2 : |x| < 1, |y| < 1\}$, and continuous on their respective boundaries.

The given function $Q(x,y) = \sum_{i,j} p_{i,j} x^i y^j - 1$, where the sum runs over the possible jumps of the walk inside $\mathbb{Z}_+^2$, is often referred to as the *kernel*. Then it has been shown that equation (1) can be solved by reduction to a boundary-value problem of Riemann-Hilbert type. This method has been the source of numerous and fruitful developments. Some recent and ongoing works have been dealing with the following matters.

- *Group of the random walk*. In several studies, it has been noticed that the so-called *group of the walk* governs the behavior of a number of quantities, in particular through its *order*, which is always even. In the case of small jumps, the algebraic curve R defined by $\{Q(x,y) = 0\}$ is either of *genus* 0 (the sphere) or 1 (the torus). In [Fayolle-2011a], when the drift of the random walk is equal to 0 (and then so is the genus), an effective criterion gives the *order* of the group. More generally, it is also proved that whenever the genus is 0, this order is infinite, except precisely for the zero drift case, where finiteness is quite possible. When the *genus* is 1, the situation is more difficult. Recently [31], a criterion has been found in terms of a determinant of order 3 or 4, depending on the arity of the group.
- *Nature of the counting generating functions*. Enumeration of planar lattice walks is a classical topic in combinatorics. For a given set of allowed jumps (or steps), it is a matter of counting the number of paths starting from some point and ending at some arbitrary point in a given time, and possibly restricted to some regions of the plane. A first basic and natural question arises: how many such paths exist? A second question concerns the nature of the associated counting generating functions (CGF): are they rational, algebraic, holonomic (or D-finite, i.e. solution of a linear differential equation with polynomial coefficients)?

  Let $f(i,j,k)$ denote the number of paths in $\mathbb{Z}_+^2$ starting from $(0,0)$ and ending at (i,j) at time k. Then the corresponding CGF

  $$F(x,y,z) = \sum_{i,j,k\geq 0} f(i,j,k) x^i y^j z^k \tag{2}$$

satisfies the functional equation

$$K(x,y)F(x,y,z) = c(x)F(x,0,z) + \widetilde{c}(y)F(0,y,z) + c_0(x,y), \tag{3}$$

where z is considered as a time-parameter. Clearly, equations (2) and (1) are of the same nature, and answers to the above questions have been given in [Fayolle-2010].

- *Some exact asymptotics in the counting of walks in* $\mathbb{Z}_+^2$. A new and uniform approach has been proposed about the following problem: *What is the asymptotic behavior, as their length goes to infinity, of the number of walks ending at some given point or domain (for instance one axis)?* The method in [Fayolle-2012] works for *both* finite or infinite groups, and for walks not necessarily restricted to excursions.

3.3.4. Simulation for urban mobility

We have worked on various simulation tools to study and evaluate the performance of different transportation modes covering an entire urban area.

- Discrete event simulation for collective taxis, a public transportation system with a service quality comparable with that of conventional taxis.
- Discrete event simulation a system of self-service cars that can reconfigure themselves into shuttles, therefore creating a multimodal public transportation system; this second simulator is intended to become a generic tool for multimodal transportation.
- Joint microscopic simulation of mobility and communication, necessary for investigation of cooperative platoons performance.

These two programs use a technique allowing to run simulations in batch mode and analyze the dynamics of the system afterward.

4. Application Domains

4.1. Introduction

While the preceding section focused on methodology, in connection with automated guided vehicles, it should be stressed that the evolution of the problems which we deal with remains often guided by the technological developments. We enumerate three fields of application whose relative importance varies with time and which have strong mutual dependencies: driving assistance, cars available in self-service mode and fully automated vehicles (cybercars).

4.2. Driving assistance

Several techniques will soon help drivers. One of the first immediate goal is to improve security by alerting the driver when some potentially dangerous or dangerous situations arise, i.e. collision warning systems or lane tracking could help a bus driver and surrounding vehicle drivers to more efficiently operate their vehicles. Human factors issues could be addressed to control the driver workload based on additional information processing requirements. Another issue is to optimize individual journeys. This means developing software for calculating optimal (for the user or for the community) paths. Nowadays, path planning software is based on a static view of the traffic: efforts have to be done to take the dynamic component in account.

4.3. New transportation systems

The problems related to the abusive use of the individual car in large cities led the populations and the political leaders to support the development of public transport. A demand exists for a transport of people and goods which associates quality of service, environmental protection and access to the greatest number. Thus the tram and the light subways of VAL type recently introduced into several cities in France conquered the populations, in spite of high financial costs. However, these means of mass transportation are only possible on lines on which there is a keen demand. As soon as one moves away from these "lines of desire" or when one deviates from the rush hours, these modes become expensive and offer can thus only be limited in space and time. To give a more flexible offer, it is necessary to plan more individual modes which approach the car as we know it. However, if one wants to enjoy the benefits of the individual car without suffering from their disadvantages, it is necessary to try to match several criteria: availability anywhere and anytime to all, lower air and soils pollution as well as sound levels, reduced ground space occupation, security, low cost. Electric or gas vehicles available in self-service, as in the Praxitèle system, bring a first response to these criteria. To be able to still better meet the needs, it is however necessary to re-examine the design of the vehicles on the following points:

- ease empty car moves to better distribute them;
- better use of information systems inboard and on ground;
- better integrate this system in the global transportation system.

These systems are now operating. The challenge is to bring them to an industrial phase by transferring technologies to these still experimental projects.

4.4. Automated vehicles

The long term effort of the project is to put automatically guided vehicles (cybercars) on the road. It seems too early to mix cybercars and traditional vehicles, but data processing and automation now make it possible to consider in the relatively short term the development of such vehicles and the adapted infrastructures. RITS aims at using these technologies on experimental platforms (vehicles and infrastructures) to accelerate the technology transfer and to innovate in this field. Other application can be precision docking systems that will allow buses to be automatically maneuvered into a loading zone or maintenance area, allowing easier access for passengers, or more efficient maintenance operations. Transit operating costs will also be reduced through decreased maintenance costs and less damage to the braking and steering systems. Regarding technical topics, several aspects of Cybercars have been developed at RITS this year. First, we have stabilized a generic Cycab architecture involving Inria SynDEx tool and CAN communications. The critical part of the vehicle is using a real-time SynDEx application controlling the actuators via two Motorola's MPC555. Today, we have decided to migrate to the new dsPIC architecture for more efficiency and ease of use. This application has a second feature, it can receive commands from an external source (Asynchronously to this time) on a second CAN bus. This external source can be a PC or a dedicated CPU, we call it high level. To work on the high level, in the past years we have been developing a R&D framework called (Taxi) which used to take control of the vehicle (Cycab and Yamaha) and process data such as gyro, GPS, cameras, wireless communications and so on. In order to rely on a professional and maintained solution, we use the RTMaps SDK development platform for our developments and demonstrations. These demonstrations include: reliable SLAMMOT algorithm using 2 to 4 laser sensors simultaneously, automatic line/road following techniques, PDA remote control, multi sensors data fusion, collaborative perception via ad-hoc network. The second main topic is inter-vehicle communications using ad-hoc networks. We have worked with the EVA team for setting and tuning OLSR, a dynamic routing protocol for vehicles communications. Our goal is to develop a vehicle dedicated communication software suite, running on a specialized hardware. It can be linked also with the Taxi Framework for getting data such GPS information's to help the routing algorithm.

5. New Software and Platforms

5.1. PML-SLAM

KEYWORD: Localization

SCIENTIFIC DESCRIPTION: Simultaneous Localization and Mapping method based on 2D laser data.

- Participants: Fawzi Nashashibi and Zayed Alsayed
- Contact: Fawzi Nashashibi

5.2. V2Provue

Vehicle-to-Pedestrian

FUNCTIONAL DESCRIPTION: It is a software developed for the Vehicle-to-Pedestrian (V2P) communications, risk calculation, and alarming pedestrians of collision risk. This software is made of an Android application dedicated to pedestrians and RtMaps modules for the vehicles.

On the pedestrian side, the application is relying on GPS data to localize the user and Wi-Fi communications are used to receive messages about close vehicles and send information about the pedestrian positioning. Besides, a service has been developed to evaluate the collision risk with the vehicles near the pedestrian and an HMI based on OpenStreetMap displays all the useful information such as pedestrian and vehicles localization and, collision risk.

On the vehicle side, RtMaps modules allowing V2X communications have been developed. These modules contain features such as TCP/UDP socket transmissions, broadcast, multicast, unicast communications, routing, forwarding algorithms, and application specific modules. In the V2ProVu software, a particular application module has been implemented to create data packets containing information about the vehicle state (position, speed, yaw rate,...) and the V2X communication stack is used to broadcast these packets towards pedestrians. Moreover, the V2proVu application can also receive data from pedestrians and create objects structures that can be shared with the vehicle perception tools.

- Contact: Fawzi Nashashibi

5.3. SimConVA

Connected Autonomous Vehicles Simulator

FUNCTIONAL DESCRIPTION: The software provides an interface between the network simulator ns-3 (https://www.nsnam.org/) and the modular prototyping framework RTMaps (https://intempora.com/).

This code allows to create an RTMaps component which activates and controls the ns-3 simulator. The component handles the sending and reception of data packets between ns-3 and RTMaps for each vehicle. It also handles the mobility of vehicles in ns-3 using their known position in RTMaps.

- Authors: Pierre Merdrignac, Oyunchimeg Shagdar and Jean-Marc Lasgouttes
- Contact: Jean-Marc Lasgouttes

6. New Results

6.1. Multi-Task Cross-Modality Deep Learning for Pedestrian Risk Estimation

Participants: Danut Ovidiu Pop, Fawzi Nashashibi.

We want to solve the problem of multi-task pedestrian protection system (PPS) including not only pedestrian classification, detection and tracking,but also pedestrian action-unit classification and prediction, and finally pedestrian risk estimation. The goal of our research work is to develop an intelligent pedestrian protection component based only on single stereo vision system using an optimal cross-modality deep learning architecture in order to fulfill the prior requirements.

The system has to be able not only to detect all the pedestrians with high precision but also to track all the pedestrian paths, to classify the current pedestrian action and to predict their next actions and, finally, to estimate the pedestrian risk by the time to crossing for each pedestrian.

First, we investigate the classification component where we analyzed how learning representations from one modality would enable recognition for other modalitie(s) within various deep learning, which one terms as cross-modality learning. Second, we study how the cross modality learning improves an end-to-end the pedestrian action detection. Third, we analyze the pedestrian action prediction and the estimation of time to cross the street.

This work has been done in collaboration with Alexandrina Rogozan and Abdelaziz Bensrhair of INSA Rouen. More detail can be fund in [12], [13], [20], [11] and in the PhD manuscript of Danut Ovidiu Pop [6].

6.2. Study on the effect of rain on computer vision

Participants: Raoul de Charette, Fabio Pizzati.

Following the works initiated in past years, we have emphasized the need of developing for outdoor-applications to be robust to adverse weather conditions.

Three works were developed this year: two in the the context of the Samuel de Champlain Québec-France collaboration with Jean-François Lalonde from Univ. Laval (Canada) and another in the context of the new co-tutelle PhD thesis of Fabio Pizzati.

- We have first proposed a physically-based rain rendering pipeline for realistically inserting rain into clear weather images. Our research [16] was published at ICCV'19 and relies on a physical particle simulator, an estimation of the scene lighting and an accurate rain photometric modeling to augment images with arbitrary amount of realistic rain or fog. We validate our rendering with a user study, proving our rain is judged 40% more realistic that state-of-the-art. Using our generated weather augmented Kitti and Cityscapes dataset, we conduct a thorough evaluation of deep object detection and semantic segmentation algorithms and show that their performance decreases in degraded weather, on the order of 15% for object detection and 60% for semantic segmentation. Furthermore, we show refining existing networks with our augmented images improves the robustness of both object detection and semantic segmentation algorithms. We experiment on the popular nuScenes dataset and measure an improvement of 15% for object detection and 35% for semantic segmentation compared to original rainy performance.
  Along with the research we have released the full augmented dataset on our project page [0] and the source code will be soon released.
- An alternative proposal is to use generative networks (GANs) to learn the translation of clear weather images to rainy images. This was achieved in the thesis of Fabio Pizzati and led to an accepted conference paper at WACV'20. To overcome the limitation of publicly available annotated datasets, we propose to learn the clear to rain mapping from datasets of different sources. Standard image-to-image translation architectures have limited effectiveness in such case due to the large source / target domain gap and usually fail to model typical traits of rain as water drops, which ultimately impacts the synthetic images realism. We proposed here a new type of domain bridge, that benefits from web-crawled data to reduce the domain gap.
- To circumvent the limitation of physics-based rendering and GANs rendering, we are currently working on extensions of [16] with Maxime Tremblay, PhD student at Univ. Laval. In this work, we are combining data-driven GAN approaches and physics-based driven learning.

6.3. Unsupervised Domain Adaptation

Participants: Raoul de Charette, Maximilian Jaritz, Fawzi Nashashibi, Fabio Pizzati.

There is an evident dead end to the paradigm of supervised learning, as it requires costly human labeling of millions of data frames to learn the appearance models of objects. As of today, the databases are recorded in very narrow conditions (e.g. only clear weather, only USA, only daytime). Adjusting to unseen conditions such as snow, hail, nighttime or unseen cities, require supervised algorithms to be retrained. Conversely, as

[0] https://team.inria.fr/rits/computer-vision/weather-augment/

humans we're capable of generalizing prior knowledge to new tasks. During this year, we initiated two works on transfer learning, typically Unsupervised Domain Adaptation (UDA) which is crucial to tackle the lack of annotations in a new domain. We have conducted two parallel projects on UDA: the first one in the scope of Maximilian Jaritz' thesis [27] (submitted), and the second one in the scope of Fabio Pizzati's work on rainy scenarios:

- **xMUDA**: In the first work, we explore how to learn from multi-modality and proposed cross-modal UDA (xMUDA) where we assumed the presence of 2D images and 3D point clouds for 3D semantic segmentation. This is challenging as the two input spaces are heterogeneous and can be impacted differently by domain shift. In xMUDA, modalities learn from each other through mutual mimicking, disentangled from the segmentation objective, to prevent the stronger modality from adopting false predictions from the weaker one. We evaluated on new UDA scenarios including day-to-night, country-to-country and dataset-to-dataset, leveraging recent autonomous driving datasets. xMUDA brings large improvements over uni-modal UDA on all tested scenarios, and is complementary to state-of-the-art UDA techniques.
- **Weighted Pseudo Labels**: The second work focus specifically on semantic segmentation in rainy scenarios. We benefited from our other work on GANs clear to rain translation to apply a self-supervised domain adaptation (aka UDA) that learns from the use of pseudo labels. Using pseudo labels enables the self-supervision of the learning reinforcing the network belief in its own predictions. To circumvent the use of hard-coded threshold, which is a common practice for pseudo labels, we proposed new Weighted Pseudo Labels that actively learn the ad-hoc threshold in a sort of region-growing techniques.

6.4. 3D completion and surface modeling

Participants: Raoul de Charette, Maximilian Jaritz, Manohar Kv.

Depth sensors (LiDARs, Time-of-flight cameras, stereo) gather geometrical knowledge about the scene which are rich and may be beneficial for many tasks. However, the depth information is usually sparse in nature and do not recover volumes and surfaces of objects.

This year we have conducted three works on the topic: one work to densify the 3D point clouds generated from LiDAR sensors, another work to fuse 2D images and 3D point clouds, and finalized another work to reconstruct 3D deformable objects.

- The first work is in spirit a 3D point completion and was initiated with intern Manohar Kv. We developed a 3D pipeline to process point cloud and densify existing point clouds. It uses a modified version of the popular PointNet++ and it is thus able to reconstruct highly occluded 3D point clouds. The work is not yet published.
- In [17] we introduce a framework to fuse 2D multi-view images and 3D point clouds in an effective way by computing image features in 2D first, lifting them to 3D, and then fuse complementary geometry and image information in canonical 3D space. This work has been done while Maximilian Jaritz was visiting San Diego University.
- In [24], we propose a new algorithm to reconstruct 3D deformable objects heavily occluded. It uses an automatic registration of multiple depth sensors and Gaussian Mixture Modeling in the radial domain to detect and reconstruct object from their symmetrical properties. This research was applied in the context of pottery wheel for the preservation of the cultural heritage and conducted in collaboration with Mines ParisTech. It resulted that our method enabled reconstruction of challenging deformable objects with an average precision of 7.6mm.

6.5. 3D Surface Reconstruction from Voxel-based Lidar Data

Participants: Luis Roldao, Raoul de Charette, Anne Verroust-Blondet.

To achieve fully autonomous navigation, vehicles need to compute an accurate model of their direct surroundings. In fact, imprecise representations may lead to unexpected situations that could endanger the passengers.

This year, we have proposed an algorithm capable to perform a fine and accurate 3D surface reconstruction of the environment from depth sensors. This representation keeps a high level of detail on the reconstruction, while maintaining a high density in the areas close to the vehicle.

Existing methods used for surface reconstruction from 3D data struggle to accommodate to the heterogeneous density of the input data while keeping the reconstruction accuracy. Conversely, our method is capable of handling this variable density by using an adaptive neighborhood kernel that perform local approximations of the data at different levels. This also permit to gain robustness against noise and output a smoother reconstruction. We also introduce a Gaussian confidence function capable to select the most adequate kernel for the local surface estimation. A Truncated Signed Distance Function (TSDF) is then globally estimated from the local surfaces to obtain the final mesh that represents the input scan.

The proposed method was evaluated in both simulated and real data. Reconstruction results show an improvement on the representation when compared with popular methods such as Implicit Moving Least Squares (IMLS), as the average error of our reconstruction is often 50% lower. Furthermore, almost 80% of vertices from our output mesh present an error below $0.2m$, while only 40% of vertices lie below the same threshold for IMLS. Our method is capable to output a higher level of detail on the reconstruction, while keeping a high density in vehicle surroundings, the mesh can be of special interest for both the robotics and the graphics community to perform different tasks, such as terrain traversability assessment or physical modeling.

More details can be found in [21]. This research is partially funded by AKKA Technology.

6.6. Attention mechanisms for vehicle trajectory prediction

Participants: Kaouther Messaoud, Fawzi Nashashibi, Anne Verroust-Blondet, Itheri Yahiaoui.

Scene understanding and future motion prediction of surrounding vehicles are crucial to achieve safe and reliable decision-making and motion planning for autonomous driving in a highway environment. This is a challenging task considering the correlation between the drivers behaviors. Two methods using attention mechanisms have been introduced in this context:

- In [18], we present a new approach based on an LSTM encoder-decoder that uses a social pooling mechanism to model the interactions between all the neighboring vehicles. This social pooling module combines both local and non-local operations: the non-local multi-head attention mechanism captures the relative importance of each vehicle despite the inter-vehicle distances to the target vehicle, while the local blocks represent nearby interactions between vehicles. Evaluations have been performed using two naturalistic driving datasets: Next Generation Simulation (NGSIM) and the highD Dataset [0]. The proposed method outperforms existing ones in terms of RMS values of prediction error, which shows the effectiveness of combining local and non-local operations in such a context.
- In [19] we propose an RRNNs based encoder-decoder architecture where the encoder analyzes the patterns underlying in the past trajectories and the decoder generates the future trajectory sequence. The originality of this network is that it combines the advantages of the LSTM blocks in representing the temporal evolution of trajectories and the attention mechanism to model the relative interactions between vehicles. The proposed method outperforms LSTM encoder decoder in terms of RMSE values of the predicted trajectories on the large scaled naturalistic driving highD dataset.

6.7. A unified framework for robust 2D/3D PML-SLAM

Participants: Kathia Melbouci, Fawzi Nashashibi.

Enhancing the outdoor mapping with SLAM based approaches is still an active research area. The main reason is that a consistent map of the vehicle's surrounding is one of the prerequisites for an effective vehicle interaction with this environment. In this context, and for the VALET project purpose, we have extended the PML-SLAM framework to handle 2D and 3D Lidars by replacing the localization module and designing a

[0] https://www.highd-dataset.com/

sparse pose graph optimizer. The sparse pose graph jointly optimizes the poses of the submaps generated by the local SLAM, which are already used for the mapping task, and the poses of the scans estimated following the scan matching process. This optimization is formulated as a non linear least square problem, and runs online in a background thread. The optimized poses are used to correct the vehicle's trajectory and to update the environment map. Furthermore, the graph-based PML-SLAM can deal with different sensors (IMU, GPS), that is, a sensor fusion "Kalman-filter" based is available to provide a good pose estimate for the local SLAM.

6.8. LIDAR-Based perception For Vehicle Localization in an HD Map

Participants: Farouk Ghallabi, Fawzi Nashashibi.

Self-vehicle localization is one of the fundamental tasks for autonomous driving. Most of current techniques for global positioning are based on the use of GNSS (Global Navigation Satellite Systems). However, these solutions do not provide a localization accuracy that is better than 2-3 m in open sky environments. Alternatively, the use of maps has been widely investigated for localization since maps can be pre-built very accurately. State of the art approaches often use dense maps or feature maps for localization. This year, we tackled to problems:

- In [14] we proposed a road sign perception system for vehicle localization within a third party map. This is challenging since third party maps are usually provided with sparse geometric features, which makes the localization task more difficult in comparison to dense maps. Experiments have been conducted on a Highway-like test track using GNSS/INS with RTK corrections as ground truth (GT).
- In [15] High Reflective Landmarks (HRL) - such as lane markings, road signs and guard rail reflectors (GRR) - are detected from a 3D point cloud. A particle filtering algorithm estimates the position of the vehicle by matching observed HRLs with HD map attributes. Experiments have been conducted on a highway-like test track using GNSS/INS with RTK corrections as a ground truth (GT). Error evaluations are given as cross-track (CT) and along-track (AT) errors defined in the curvilinear coordinates related to the map. The obtained accuracies of the localization system is 18 cm for the cross-track error and 32 cm for the along-track error.

6.9. Motion planning in presence of highly dynamic obstacles with uncertain motion

Participants: Pierre de Beaucorps, Renaud Poncelet, Anne Verroust-Blondet, Fawzi Nashashibi.

Safe motion planning in a dynamic environment is of great importance in many robotics applications. This year, we have worked in two directions:

- The work on reachable interaction sets introduced in [37] has been extended to the case of dynamic obstacles with uncertain motions. We consider that the obstacles have stochastic motions and we use a probabilistic formulation to compute the RIS at each time step. Our approach improves existing methods in such a context (cf. Pierre de Beaucorps PhD thesis [7]).
- Focusing on autonomous vehicles, we begun to study scenarios with occluded dynamic obstacles.

6.10. A vehicle dynamic model corrector with side slip estimation for adding safety capabilities in autonomous vehicle

Participants: Imane Mahtout, Fawzi Nashashibi.

The ability to identify malfunctions on autonomous vehicles is critical for their deployment. As a matter of fact, systems able to identify when the positioning systems are not providing accurate data, or the perception algorithms are not properly detecting the environment, are extremely important to assure a certain safety level for automated vehicles. This is especially true since these systems are finally connected to the control module that provides the adequate commands to vehicle's actuators. For control algorithms to work properly, proper inputs are necessary to reduce noise, increase controllability and avoid system's malfunctions and instability. It is also critical for these algorithms to identify/consider vehicle physical limits for determining when is the automated system still capable of handling the vehicle. From the above, it is clear that automated vehicles are in need of proper inputs to control the vehicle, but also it is necessary to detect critical situations where the nominal control behavior is no longer assured, in order to take the vehicle to a safe state. Slide slip state is an example of a critical situation where the vehicle is no longer able to correct its trajectory. Thus, this part of my thesis work consists on developing a module for providing smooth signals to the controller and, at the same time, detect side slip situations. The lateral controller implemented in our automated driving (AD) system is based on the yaw error minimization between the desired yaw rate (obtained from the road layout information in function of the curvature) and the current vehicle yaw rate. From this, the first step is to provide a proper current yaw rate measurement. The proposed device compares the measured yaw rate value (coming from the vehicle sensor) with a model-based estimated yaw rate value. The idea is to identify vehicle model mismatches, correcting the model in real time. This permits to extend the nominal vehicle planar model to a road layout-independent model, where roll and pitch variations are considered. This first stage consists of a vehicle model compensator that includes unmodeled vehicle dynamics and parameter incertitude in real time when the vehicle is operating in autonomous mode. The vehicle lateral controller is then fed by the compensator output to allow robust performance in all road conditions. Once lateral control is fed with proper inputs, the second stage is the one detecting that vehicle handling physical limits are surpassed. The proposed system based on Youla-Kucera parametrization identifies the vehicle physical limits by estimating front and rear lateral forces, using as input the previous corrected model and on-board vehicle info. This permits to provide an accurate identification of slide slip vehicle states without adding any additional sensor to production vehicle's on-board sensors (see Fig. 1).

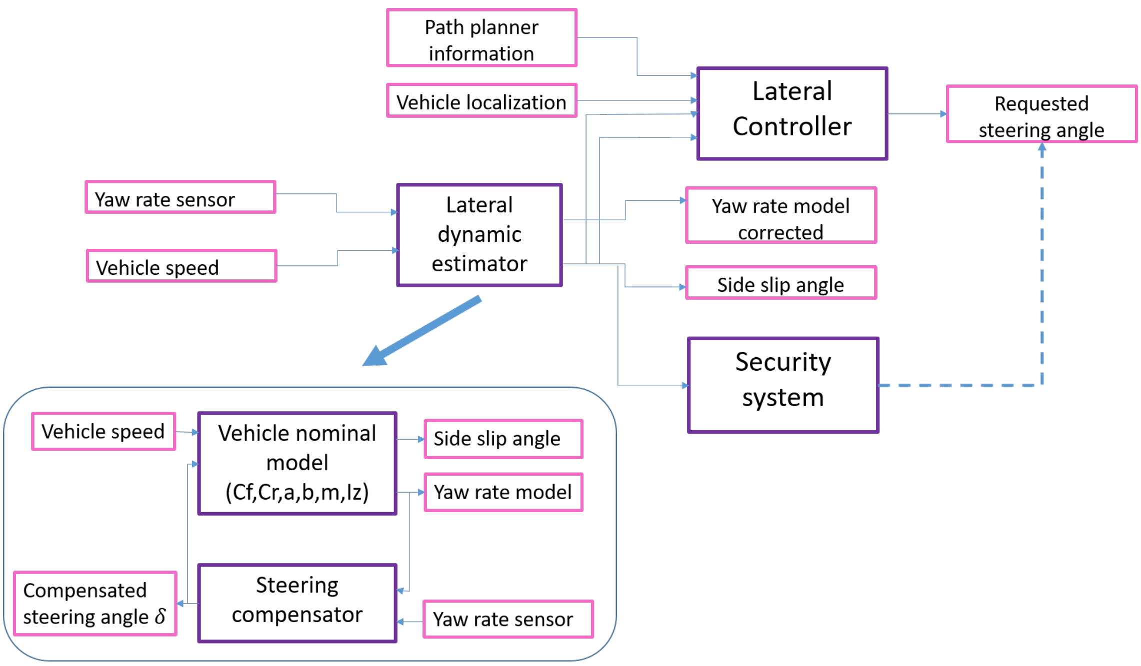

Figure 1. Lateral model compensator

6.11. Perception-adapted controller device for autonomous vehicles

Participants: Imane Mahtout, Fawzi Nashashibi.

Without loss of generality, let us consider a single camera for detecting a preceding vehicle in the road. It is clear that there are two parameters that impact the performance of the perception system:

1) the specific algorithm developed to detect and track the objects providing accurate measurement; and

2) the physical limitation of the sensor itself. For a camera, the number of pixels limits the resolution of the image so the farther away the vehicle is, the lower the accuracy in its detection. This implies a more inaccurate measurement that will degrade the ego-vehicle performance. From the vehicle response point of view, we cannot expect that a single control device can handle for example a camera-based car-following system for all detected vehicle distances.
For the sake of clarity, Figure 2 shows the speed of a preceding vehicle measured from a ground truth (solid blue line); and the measured speed from an on-board perception system. The speed of the preceding vehicle was computer controlled so we can assure a given response for it. In this example, it follows four consecutive reference speed changes from 0 to 5m/s, and finally to 8m/s. The ego-vehicle equipped with the on-board perception system was following that preceding vehicle at a speed dependent distance (i.e. as any on-the-market ACC system), meaning than the higher the speed, the higher the inter-vehicle distance. One can clearly see how the higher the speed, the higher the inaccuracy of the perception system, the more degraded the measurement.

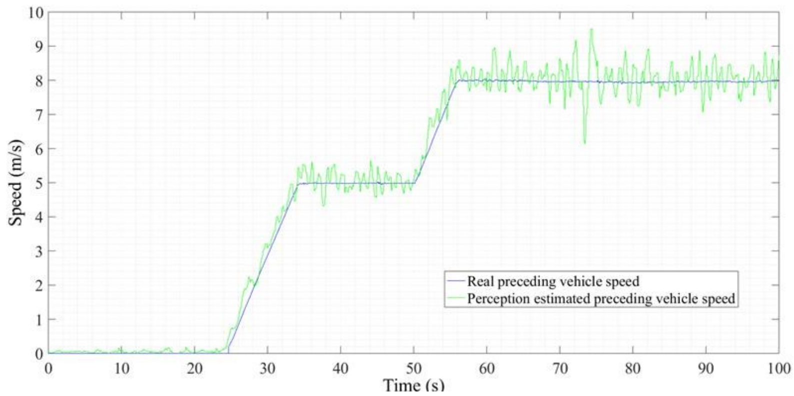

Figure 2. Perception speed profile measurement

We worked on a novel control device able to adapt its response to the perception system capabilities, modifying vehicle response accordingly to the level of accuracy of the perception system. This novel idea redefines and extends the capabilities of any ADAS or autonomous vehicle technology, not only because it improves vehicle's performance but also because we can, a-priori, understand the limitations of the full vehicle performance with a complete closed loop analysis from perception to vehicle control. As additional remark, there are ways of dealing with these problems by filtering the perception signal. This causes two problems: control response stability is no longer guaranteed, and it smooths the response but it's not possible to link the full system performance to that filtering. Figure 3 presents an overview of the method. The block Perception set-of-sensor represents the specific embedded perception system. It can consist either in a single sensor or a combination of them. The characterization of the specific perception setup can be done offline, calibrating the system performance accordingly to the on-board sensors. The module Offline calibration will contain a 3-D look-up table with object distance, preceding-vehicle speed, and as a third parameter the desired measurement inaccuracy (it can be either the speed as shown in Figure 2 or any other relevant parameter as the distance, yaw rate…). This offline calibration allows defining specific design parameters for the vehicle performance. Current control systems linked to perception don't consider this inaccuracy when designing the vehicle performance. We here include two different control design criteria blocks. Assuming that we keep the regular controller design that considers perfect measurement from the perception system; the First control design criteria block includes the current production system controller. On the contrary side, we have also included a Second control design criteria block that can be adapted in function of the specific interest of each application. Following with the example on Figure 2, let us assume that we are interested on developing an application between 0 and 10m/s and the inaccuracy of the perception system is the one presented in the plots. Having this in mind, we can design the second controller with the goal of minimizing the impact of that inaccuracy in the vehicle performance. The system also uses as input the real-time perception value coming from the Perception system measurement block (in the case of Figure 2 would provide the speed of the preceding vehicle in real-time). Then, this measurement feds the Perception-adapted controller block and the Performance degradation block. This last, accordingly to the information from the Offline calibration block, determines the status of the on-board perception system. The output of the Performance degradation block with the output of the first and second design criteria blocks fed the core module of this work: the real-time vehicle performance adaptation module. It is composed by two main blocks: the response corrector block to adapt the vehicle performance and the Perception-adapted controller block that merges both designed controller in a single stable structure.

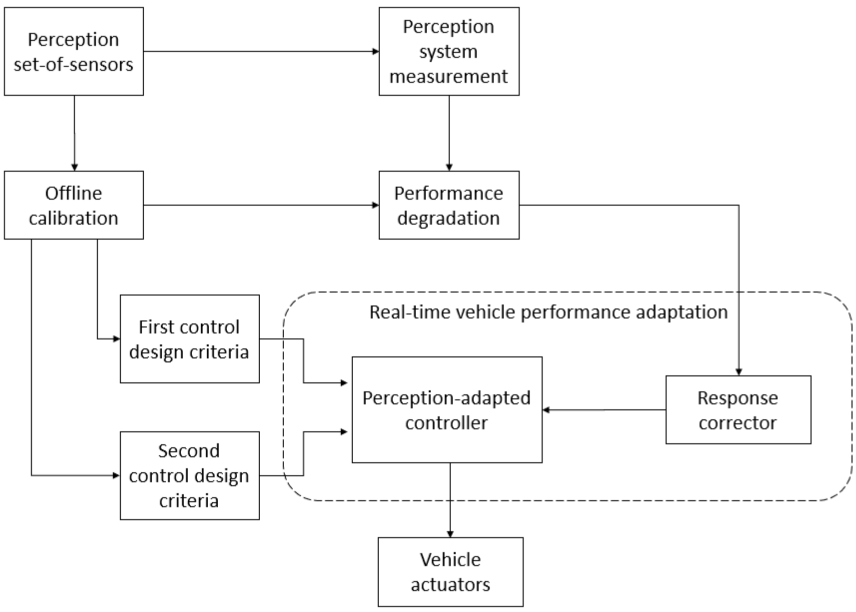

Figure 3. General overview of the block diagram for perception-adapted autonomous vehicle control device.

6.12. Cyberphysical constructs and mobile communications for fully automated networked vehicles

Participant: Gérard Le Lann.

Safety, privacy, efficiency, and cybersecurity (SPEC) properties are key to the advent of self-forming and self-healing networks of fully automated (driverless) terrestrial vehicles. Such vehicles are referred to as Next-Gen Vehicles (NGVs) in order to avoid confusion with Connected Autonomous Vehicles (CAVs). NGVs prefigure SAE level 5 vehicles. CAVs and NGVs rest on robotics capabilities (sensors, motion control laws, actuators, onboard systems, etc.). CAVs are equipped with V2X (vehicle-to-everything) functionalities based on medium range WiFi radio communications. NGVs will be equipped with CMX (coordinated mobility for X) functionalities, X standing for S, P, E, and C, based on very short range communications (cellular radio and optics).

Work in 2019 has been devoted to defining the CMX framework and to comparing V2X and CMX functionalities. The outputs of this work have been published in [23].

Highest SE (safety and efficiency) is one of the most fundamental goals set to designers of onboard systems. It is surmised that onboard robotics must be supplemented with inter-vehicular communication (IVC) capabilities in order to achieve highest SE properties. Thus the question: which IVC capabilities? In the V2X framework, two distinct sets of IVC capabilities are considered, namely DSRC-V2X (WiFi radio) and C-V2X (4G LTE, 5G, cellular radio). IVC capabilities in the CMX framework encompass cellular radio, VLC and passive optics.

Since V2X functionalities rest on medium range radio communications, they are vulnerable to remote and local cyberattacks (message falsification, masquerading, Sybil attacks, injection of bogus messages, DDoS attacks, etc.). It has been amply demonstrated that such cyberattacks can compromise safety (collisions caused by remote and/or local attackers) as well as efficiency (congested roadways and cities). Furthermore, V2X functionalities break down when radio channels are noisy (messages get lost) or/and jammed (intentional remote and local cyberattacks). Finally, owing to decade-old design decisions, there are no privacy properties with V2X functionalities. For example, every CAV must periodically broadcast messages that carry vehicle-centric characteristics and unencrypted current GNSS space coordinates (referred to as beaconing, frequencies ranging between 1 Hz and 10 Hz. Despite certificate-based pseudonymisation, routes followed by vehicles can be tracked and communications can be eavesdropped and recorded. Linkage with passengers) personal data is straightforward.

Therefore, in addition to degrading safety and efficiency properties achieved by onboard robotics, V2X functionalities do not meet elementary requirements regarding privacy and cybersecurity. Some proponents of the V2X approach assert that it is impossible to deliver road safety without breaching passengers' privacy. To be valid, that statement should be backed with an impossibility proof. Such a proof has not appeared yet and will never appear for the simple reason that safety and privacy properties can be achieved jointly, by design, proofs given, as demonstrated with the CMX approach.

From a more theoretical perspective, the V2X and the CMX frameworks can be contrasted as follows. Unquestionably, full asynchrony is the appropriate model for representing the vehicular network universe faithfully. Vehicles are started or stopped at arbitrary times, velocities change unpredictably, ditto for lane changes, on-ramp merging, concurrent traversals of intersections and roundabouts, and so on. Onboard processes that are life/safety critical are run in the presence of fortuitous failures, cyberattacks, and concurrency (due to resource sharing). It follows that even if one postulates the existence of finite bounds for process execution durations, it is impossible to assume any a priori knowledge of values taken by those bounds. That is precisely the definition of full asynchrony.

Numerous impossibility results relative to fully asynchronous systems have been published since the late-1970s. For example, problems akin to distributed consensus (terminating reliable broadcast, consistent multi-copied data structures, exact agreement, leader election, etc.) have no solutions in the presence of a single failure, even when communications are assumed to be perfect (no message losses). Since mobile wireless communications are unreliable, those results hold a fortiori in vehicular networks. Obviously, problems that involve termination in computable/predictable time bounds (a real-time property) have no solutions either.

The above-mentioned problems shall be solved in order to provide vehicles and vehicular networks with the SPEC properties. Knowing that solutions exist when considering synchrony models -such as e.g. partial synchrony, timed asynchrony, full synchrony- the challenge is to show how synchrony models could emerge from full asynchrony. This challenge is ignored in the V2X framework. Conclusion: since V2X designs are conducted considering full asynchrony, none of the SPEC properties may hold true.

The CMX framework results from addressing this challenge. NGVs are endowed with CMX functionalities which are based on specific cyberphysical constructs (cells, cohorts, flocks). These constructs serve to instantiate synchrony models within which it is possible to design protocols and algorithms (e.g., deterministic MAC protocols, time-bounded distributed algorithms for message dissemination, approximate agreement, and consensus) that are needed for establishing and proving the SPEC properties, while matching the real vehicular networks universe.

Concepts at the core of the CMX framework (cyberphysical levels, unfalsifiable vehicle profiles, proactive security modules, privacy-preserving naming, etc.) are detailed in [23]. Regarding SE properties, we show how to achieve theoretical absolute safety (no fatalities, no severe injuries) while keeping smallest safe gaps (highest efficiency) in cohort-structured vehicular networks, under assumptions of high coverage. As for PC properties, we show that passengers' privacy cannot be compromised via cyber eavesdropping and/or physical tracking of vehicles. This is due to the fact that messages do not carry vehicle-centric characteristics or GNSS space coordinates. CMX functionalities are shown to be immune to remote cyberattacks. Thanks to optical communications (in addition to very short range cellular radio), they can withstand radio channel jamming. Owing to controlled cohort admission, external local cyberattacks aimed at cohort members are inoperative. Local cyberattacks launched from the inside of a cohort, i.e. by cohort members themselves, can be thwarted. In the unlikely case of success, dishonest members would be involved in those collisions which they create. Conclusion: the only cyberattacks that may compromise safety in cohort-structured vehicular networks are due to irrational attackers.

6.13. Belief propagation inference for traffic prediction

Participant: Jean-Marc Lasgouttes.

This work [36], [35], in collaboration with Cyril Furtlehner (TAU, Inria), deals with real-time prediction of traffic conditions in a urban setting with incomplete data. The main focus is on finding a good way to encode available information (flow, speed, counts,...) in a Markov Random Field, and to decode it in the form of real-time traffic reconstruction and prediction. Our approach relies in particular on the Gaussian belief propagation algorithm.

Through our collaboration with PTV Sistema, we obtained extensive results on large-scale datasets containing 250 to 2000 detectors. The results show very good ability to predict flow variables and a reasonably good performance on speed or occupancy variables. Some element of understanding of the observed performance are given by a careful analysis of the model, allowing to some extent to disentangle modelling bias from intrinsic noise of the traffic phenomena and its measurement process.

This year we worked on code optimization and submitted our work to *Transportation Research: Part C*.

6.14. Stabilization of traffic through cooperative autonomous vehicles

Participants: Guy Fayolle, Carlos Flores, Jean-Marc Lasgouttes.

We investigate in [26] the transfer function emanating from the linearization of a car-following model, when taking into account a driver reaction time. This leads to stability conditions, which are explicitly given. We also show how this reaction time can introduce a *weak string instability*.

This paper is intended as a foundation of a larger work on traffic stabilization by means of a fleet of cooperative automated vehicles. Contrary to some earlier works, our approach is based on a car-following model with reaction-time delay, rather than on a first order fluid model. The continuation of these studies will concern schockwave analysis and adequate traffic-stabilizing control strategies.

6.15. Random walks in orthants and lattice path combinatorics

Participant: Guy Fayolle.

In the second edition of the book [2], original methods were proposed to determine the invariant measure of random walks in the quarter plane with small jumps (size 1), the general solution being obtained via reduction to boundary value problems. In this framework, number of difficult open problems related to lattice path combinatorics are currently being explored, in collaboration with A. Bostan and F. Chyzak (project-team SPECFUN, Inria-Saclay), both from theoretical and computer algebra points of view: concrete computation of the criteria, utilization of differential Galois theory, genus greater than 1 (i.e. when some jumps are of size ≥ 2), etc. A recent topic deals with the connections between simple product-form stochastic networks (so-called *Jackson networks*) and explicit solutions of functional equations for counting lattice walks, see [25].

6.16. Optimization of test case generation for ADAS via Gibbs sampling algorithms

Participant: Guy Fayolle.

Validating Advanced Driver Assistance Systems (ADAS) is a strategic issue, since such systems are becoming increasingly widespread in the automotive field.

But ADAS validation is a complex issue, particularly for camera based systems, because these functions may be facing a very high number of situations that can be considered as infinite. Building at a low cost level a sufficiently detailed campaign is thus very difficult. Indeed, test case generation faces the crucial question of *inherent combinatorial explosion*. An important constraint is to generate *almost all* situations in the most economical way. This task can be considered from two points of view: deterministic via binary search trees, or stochastic via Markov chain Monte Carlo (MCMC) sampling. We choose the latter probabilistic approach described below, which in our opinion seems to be the most efficient one. Typically, the problem is to produce samples of large random vectors, the components of which are possibly dependent and take a finite number of values with some given probabilities.The following flowchart is proposed.

1. In a first step, starting from the simulation graph generated by the toolboxes of MATLAB, we construct a so-called *Markov Random Field (MRF)*. When the parameters are locally dependent, this can be achieved from the user's specifications and by a systematic application of Bayes' formula.
2. Then, to cope with the combinatorial explosion, test cases are produced by implementing (and comparing) various *Gibbs samplers*, which are fruitfully employed for large systems encountered in physics. In particular, we strive to make a compromise between the convergence rate toward equilibrium, the percentage of generated duplicates and the path coverage, keeping in mind that the speed of convergence is exponential, a classical property deduced from the general theory of Markov chains.
3. The generation of rare events by mixing Gibbs samplers, large deviation techniques (LDT) and cross-entropy method is a work in progress.

The French car manufacturer *Groupe PSA* shows a great interest in these methods and has established a contractual collaboration involving ARMINES-Mines ParisTech (Guy Fayolle as associate researcher) and Can Tho University in Vietnam (Pr. Van Ly Tran).

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

Valeo Group: a very strong partnership is under reinforcement between Valeo and Inria. Several bilateral contracts were signed to conduct joint works on Driving Assistance, some of which Valeo is funding. This joint research includes:

- The PhD thesis of Pierre de Beaucorps under the framework of Valeo project “Daring”,
- A CIFRE like PhD thesis is ongoing between Valeo and Inria (Maximilian JARITZ), dealing with multisensor processing and learning techniques for free navigable road detection.
- Valeo is currently a major financing partner of the “GAT” international Chaire/JointLab in which Inria is a partner. The other partners are: UC Berkeley, Shanghai Jiao-Tong University, EPFL, IFSTTAR, MPSA (Peugeot-Citroën) and SAFRAN.
- Technology transfer is also a major collaboration topic between RITS and Valeo as well as the development of a road automated prototype.
- Finally, Inria and Valeo are partners of the PIA French project CAMPUS (Connected Automated Mobility Platform for Urban Sustainability) including SAFRAN, Invia and Gemalto. The aim of the project is the development of autonomous vehicles and the realization of two canonical uses-cases on highways and urban like environments.

Renault Group: Collaboration between Renault and RITS re-started in 2016. Different research teams in Renault are now working separately with RITS on different topics.

- A CIFRE like PhD thesis is ongoing between Renault and Inria (Farouk GHALLABI) The thesis deals with the accurate localization of an autonomous vehicle on a highway using mainly on-board low-cost perception sensors.
- Another CIFRE PhD thesis is ongoing since November 2017 (Imane MAHTOUT).

AKKA Technologies: Collaboration with AKKA since 2012 (for the Link & Go prototype).

- Inria and AKKA Technologies are partners in the VALET projects (ANR projects).
- A CIFRE PhD thesis (Luis ROLDAO) dealing with 3D-environment modeling for autonomous vehicles begun in October 2017.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

8.1.1.1. VALET

Title: Redistribution automatique d’une flotte de véhicules en partage et valet de parking

Instrument: ANR

Duration: January 2016 - September 2019

Coordinator: Fawzi Nashashibi

Partners: Inria, Ecole Centrale de Nantes (IRCCyN), AKKA Technologies

Inria contact: Fawzi Nashashibi

Abstract: The VALET project proposes a novel approach for solving car-sharing vehicles redistribution problem using vehicle platoons guided by professional drivers. An optimal routing algorithm is in charge of defining platoons drivers’ routes to the parking areas where the followers are parked in a complete automated mode. The main idea of VALET is to retrieve vehicles parked randomly on the urban parking network by users. These parking spaces may be in electric charging stations, parking for car sharing vehicles or in regular parking places. Once the vehicles are collected and guided in a platooning mode, the objective is then to guide them to their allocated parking area or to their respective parking lots. Then each vehicle is assigned a parking place into which it has to park in an automated mode.

8.1.1.2. Hianic

Title: navigation autonome dans les foules inspirée par les humains (Human Inspired Autonomous Navigation In Crowds)

Instrument: ANR

Duration: January 2018 - December 2020

Coordinator: Anne Spalanzani (Inria Rhône-Alpes, Chroma research team)

Partners: Inria Rhône-Alpes, Inria Paris, LIG Laboratoire d'Informatique de Grenoble, LS2N - ECN Laboratoire des Sciences du Numérique de Nantes

Inria contact: Fawzi Nashashibi

Abstract: The HIANIC project will try to address some problems that will arise when these cars are mixed with pedestrians. The HIANIC project will develop new technologies in term of autonomous navigation in dense and human populated traffic. It will explore the complex problem of navigating autonomously in shared-space environments, where pedestrians and cars share the same environment.

Such a system will contribute both to urban safety and intelligent mobility in "shared spaces". Negotiation will help to avoid frozen situations increasing the vehicle's reactivity and optimizing the navigable space. Negotiation, Human-Aware Navigation and Communication will contribute to a better public acceptance of such autonomous systems and facilitate their penetration in the transportation landscape.

8.1.2. FUI

8.1.2.1. PAC V2X

Title: Perception augmentée par coopération véhicule avec l'infrastructure routière

Instrument: FUI

Duration: September 2016 - May 2020

Coordinator: SIGNATURE Group (SVMS)

Partners: DigiMobee, LOGIROAD, MABEN PRODUCTS, SANEF, SVMS, VICI, Inria, VEDECOM

Inria contact: Raoul de Charette

Abstract: The objective of the project is to integrate two technologies currently being deployed in order to significantly increase the time for an automated vehicle to evolve autonomously on European road networks. It is the integration of technologies for the detection of fixed and mobile objects such as radars, lidars, cameras ... etc. And local telecommunication technologies for the development of ad hoc local networks as used in cooperative systems.

8.1.3. Competitivity Clusters

RITS team is a very active partner in the competitivity clusters, especially MOV'EO and System@tic. We are involved in several technical committees like the DAS SUR of MOV'EO for example.
RITS is also the main Inria contributor in the VEDECOM institute (IEED).

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. AUTOCITS

Title: AUTOCITS Regulation Study for Interoperability in the Adoption of Autonomous Driving in European Urban Nodes

Program: CEF- TRANSPORT Atlantic corridor

Duration: November 2016 - March 2019

Coordinator: Indra Sistemas S.A. (Spain)

Partners: Indra Sistemas S.A. (Spain); Universidad Politécnica de Madrid (UPM), Spain; Dirección General de Tráfico (DGT), Spain; Inria (France); Instituto Pedro Nunes (IPN), Portugal; Autoridade Nacional de Segurança Rodoviária (ANSR), Portugal; Universidade de Coimbra (UC), Portugal.

Inria contact: Fawzi Nashashibi, Mohammad Abualhoul

Abstract: The aim of the Study is to contribute to the deployment of C-ITS in Europe by enhancing interoperability for autonomous vehicles as well as to boost the role of C-ITS as catalyst for the implementation of autonomous driving. Pilots will be implemented in 3 major Core Urban nodes (Paris, Madrid, Lisbon) located along the Core network Atlantic Corridor in 3 different Member States. The Action consists of Analysis and design, Pilots deployment and assessment, Dissemination and communication as well as Project Management and Coordination.

8.2.2. Collaborations with Major European Organizations

RITS is member of the **euRobotics AISBL** (Association Internationale Sans But Lucratif) and the Leader of "People transport" Topic. This makes from Inria one of the rare French robotics representatives at the European level. See also: http://www.eu-robotics.net/

RITS is a full partner of **VRA – Vehicle and Road Automation**, a support action funded by the European Union to create a collaboration network of experts and stakeholders working on deployment of automated vehicles and its related infrastructure. VRA project is considered as the cooperation interface between EC funded projects, international relations and national activities on the topic of vehicle and road automation. It is financed by the European Commission DG CONNECT and coordinated by ERTICO – ITS Europe. See also: http://vra-net.eu/

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners

RITS has signed 3 MoU with the following international laboratories:

- Vehicle Dynamics and Control Laboratory, Seoul National University (SNU), S. Korea: international cooperation agreement for Graduate-Level Academic and Research Collaboration
- MICA Lab, Hanoi University of Science and Technology, Vietnam: cooperation agreement for research collaboration and PhD students co-supervision
- Integrated Industrial Design Lab (INDEL) of the Department of Product and Systems Design Engineering, University of the Aegean, Greece: international cooperation agreement for Graduate-Level Academic and Research Collaboration

8.3.2. Participation in Other International Programs

Samuel de Champlain Québec-France collaboration program: "Vision par ordinateur en conditions difficiles", cooperation between Raoul de Charette and Jean-François Lalonde from Laval University.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

Plamen Petrov from Technical University of Sofia, from July to September 2019.

8.4.1.1. Internships

- Pranav Agarwal, from August 2019.
- Fares Bessam, Master student, April-September 2019.
- Manuel Gonzalez and Leonardo Ward, from Simon Bolivar University, Venezuela, from September 2019.
- Manohar KV, May-July 2019.

8.4.2. Visits to International Teams

8.4.2.1. Research Stays Abroad

Maximilian Jaritz was at UC San Diego, visiting SU Lab directed by Hao Su, from October 1st 2018 to February 15th 2019.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Selection

9.1.1.1. Reviewer

Raoul de Charette: *BMVC 2019*, *CVPR 2019*, *ICCV 2019*, *AAAI 2019*, *IEEE ITSC 2019*, *IEEE IV 2019*.

Jean-Marc Lasgouttes: *IEEE ITSC 2019*, *ICVES 2019*.

Fawzi Nashashibi : *IEEE ICRA 2019*, *IEEE IROS 2019*, *IEEE IV 2019*, *IEEE ITSC 2019*, *IEEE VTC 2018*, *IEEE ICVES 2019*.

Anne Verroust-Blondet: *IEEE ICRA 2020*, *IEEE IV 2019*, *IEEE ITSC 2019*, *ICVES 2019*.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

Guy Fayolle: associate editor of the journal *Markov Processes and Related Fields*.

Fawzi Nashashibi: associate editor of the journal *IEEE Transactions on Intelligent Vehicles*, associate editor of the journal *IEEE Transactions on Intelligent Transportation Systems*.

Fawzi Nashashibi: currently Guest Editor of the *IEEE Sensors* journal (ISSN 1424-8220), Special Issue on “The Application of Sensors in Autonomous Vehicles”.

Anne Verroust-Blondet: associate editor of the journal *The Visual Computer*.

9.1.2.2. Reviewer - Reviewing Activities

Raoul de Charette: *IEEE Transactions on Intelligent Transportation Systems*

Guy Fayolle: *AAP, MPRF, PTRF, QUESTA, European Journal of Combinatorics, JSP, Physica A, Springer Science*.

Jean-Marc Lasgouttes: *IEEE Transactions on Intelligent Transportation Systems*, *Probability in the Engineering and Informational Sciences*.

Fawzi Nashashibi: *IEEE Transactions on Intelligent Transportation Systems*, *IEEE Transactions on Intelligent Vehicles*, *Transportation Research Part C*, *IEEE ACCESS*.

Anne Verroust-Blondet: *IEEE Transactions on Intelligent Vehicles*, *The Visual Computer*.

9.1.3. Invited Talks

Raoul de Charette: *Sparse and dense data in computer vision*, Universite Laval, Québec, Canada, July 5th 2019.

Guy Fayolle was invited participant at the conference *Transient Transcendence in Transylvania*, Braov, Romania, May 13–17, 2019.

Gérard Le Lann: *Automated Driving: Ethical/Moral Quandaries*, invitation by Comité d'éthique du CNRS (COMETS), Paris, 26 February 2019.

Gérard Le Lann: *Ethical and Privacy Issues in Autonomic Vehicular Networks*, invited talk, Symposium on Networking Research Topics: Past, Present and Future inspired by Mario Gerla, Politecnico di Milano (Italy), 3 June 2019.

Gérard Le Lann: *Une mise en perspective - Qu'avons-nous appris de l'aventure Arpanet/Internet qui puisse servir aujourd'hui ?*, Colloque SIF Les 50 ans d'Internet, CNAM, Paris, 29 October 2019.

Fawzi Nashashibi: *Connected autonomous vehicles: objectives and challenges*, invited plenary talk, at the JNRR 2019 (national days on robotics research), Vittel, October 18, 2019.

Fawzi Nashashibi: *Connected Autonomous Vehicles: selected challenges*, Keynote speaker, at the IEEE ITSC 2019 conference - workshop, Auckland (New Zealand), October 27, 2019.

Fawzi Nashashibi: *The role of the intelligent infrastructure in the development and deployment of autonomous vehicles*, Keynote speaker, at the SMIV international workshop, organised by Vedecom institute, Paris, November 13, 2019.

Fawzi Nashashibi: *Connected Autonomous Vehicles: next challenges*, Keynote speaker, at the IEEE CIS-RAM conference, Bangkok (Thailand), November 18-20, 2019.

9.1.4. Scientific Expertise

Guy Fayolle is scientific advisor and associate researcher at the *Robotics Laboratory of Mines ParisTech*. He is also collaborating member of the research-team SPECFUN at Inria-Saclay.

Jean-Marc Lasgouttes is member of the Inria *Commission d'Evaluation* of Inria and of the *Conseil Académique* of Université Paris-Saclay.

Gérard Le Lann: Contribution to White Book on Cybersecurity: current challenges and Inria's research directions, Section 6.3.2 on Robotics and connected autonomous vehicles, February 2019.

Fawzi Nashashibi is an associate researcher at the *Robotics Laboratory, Mines ParisTech*. He is an evaluator/reviewer of European H2020 projects.

Anne Verroust-Blondet was a member of the CRCN Recruiting Commission of the CRI Paris in 2019.

9.1.5. Research Administration

Jean-Marc Lasgouttes is a member of the *Comité Technique Inria* and of the *Comité Local Hygiène Sécurité et Conditions de Travail* of Inria Paris.

Guy Fayolle is a member of the working group IFIP WG 7.3.

Fawzi Nashashibi is a member of the international Automated Highway Board Committee of the TRB (AHB30).
He is a member of the Board of Governors of the VEDECOM Institute representing Inria and of the Board of Governors of MOV'EO Competitiveness cluster representing Inria. He is also member of the DAS-SMI of MOV'EO cluster.

Anne Verroust-Blondet is the scientific correspondent of the European affairs and of the International Partnerships for Inria Paris. She was a member of the COST-GTRI committee at Inria (Committee in charge of the evaluation of international projects) and a member of the "emplois scientifiques" committee of Inria Paris in 2019.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Primary school: Raoul de Charette, “Introduction à l’intelligence artificielle”, 10h, -, Puck et Ribambelle, Castelnau-le-Lez, France.

Licence: Fawzi Nashashibi, “Programmation avancée”, 84h, L1, Université Paris-8 Saint-Denis, France.

Engineering, 2nd year: Fawzi Nashashibi, “Image synthesis and 3D Infographics”, 12h, M2, INT Télécom SudParis, IMA4503 “Virtual and augmented reality for autonomy”.

Master: Jean-Marc Lasgouttes, “Analyse de données”, 54h, second year of Magistère de Finance (M1), University Paris 1 Panthéon Sorbonne, France.

Master: Jean-Marc Lasgouttes, “Analyse de données”, 52.5h, Master 1 SIC APP, University Paris 1 Panthéon Sorbonne, France.

Master: Kaouther Messaoud, "Machine learning 1", 22h (TP), M1, ESIEE, Paris.

Master: Fawzi Nashashibi, “Obstacle detection and Multisensor Fusion”, 4h, M2, INSA de Rouen.

Master: Fawzi Nashashibi, “Perception and Image processing for Mobile Autonomous Systems”, 12h, M2, University of Evry.

Mastere : Raoul de Charette, “Scene Understanding with Computer Vision”, 20h, post master, Mines ParisTech, France.

Doctorat: Jean-Marc Lasgouttes, “Introduction au Boosting”, 10.5h, Mastère Spécialisé “Expert en sciences des données”, INSA-Rouen, France.

9.2.2. Supervision

PhD: Pierre de Beaucorps, “Planification de trajectoire dans un environnement peu contraint et fortement dynamique”, UPMC Sorbonne University, July 2019, supervisor: Anne Verroust-Blondet, co-supervisor: Fawzi Nashashibi.

PhD: Danut Ovidiu Pop, "Multi-Task Cross-Modality Deep Learning for Pedestrian Risk Estimation", INSA Rouen, November 2016, supervisor: Abdelaziz Bensrhair, co-supervisor: Fawzi Nashashibi.

PhD in progress: Farouk Ghallabi, “Environment modeling and simultaneous localization of a mobile vehicle on motorways: a multi-sensor approach”, Mines ParisTech, PSL Research University, October 2016, supervisor: Fawzi Nashashibi.

PhD in progress: Maximilian Jaritz, “Perception multi-capteur pour la conduite autonome grâce à l’apprentissage profond”, Mines ParisTech, PSL Research University, January 2017, supervisor: Fawzi Nashashibi, co-supervisor: Raoul de Charette.

PhD in progress: Imane Matout, “Estimation de l’intention des véhicules pour la prise de décision et le contrôle sans faille en navigation autonome”, Mines ParisTech, PSL Research University, December 2017, supervisor: Fawzi Nashashibi, co-supervisor: Vicente Milanés.

PhD in progress: Kaouther Messaoud, “Détermination des manoeuvres et des intentions des véhicules avoisinant un véhicule autonome”, UPMC Sorbonne University, October 2017, supervisor: Anne Verroust-Blondet, co-supervisors: Fawzi Nashashibi, Itheri Yahiaoui.

PhD in progress: Fabio Pizzati, “Style transfer and domain adaptation for semantic segmentation, PSL Research University and University of Bologna, November 2019, co-supervisors: Andrea Prati, Stefano Selleri, Fawzi Nashashibi and Raoul de Charette.

PhD in progress: Renaud Poncelet, “Navigation autonome en présence d’obstacles fortement dynamiques au mouvement incertain”, UPMC Sorbonne University, September 2018, supervisor: Anne Verroust-Blondet, co-supervisor: Fawzi Nashashibi.

PhD in progress: Luis Roldao, “Modélisation 3D de l’environnement et de la manoeuvrabilité d’un véhicule”, UPMC Sorbonne University, October 2017, supervisor: Anne Verroust-Blondet, co-supervisor: Raoul de Charette.

9.2.3. Juries

Fawzi Nashashibi was a reviewer of the HDR thesis of Mr. Romuald Aufrère - *Systèmes de perception bayésienne pour la robotique mobile : application à la localisation de véhicules*, Université Clermont-Ferrand, July 10, 2019.

Fawzi Nashashibi was a reviewer of the PhD thesis of Mrs. Laurène Claussmann - *Motion Planning for Autonomous Highway Driving: A Unified Architecture for Decision-Maker and Trajectory Generator*, Université Versailles, September 27, 2019.

Fawzi Nashashibi was a reviewer of the PhD thesis of Mr. Mohamed-Taha Boudali - *Contributions au guidage d'un véhicule autonome en situations non conventionnelles de conduite*, Université Mulhouse, September 25, 2019.

Fawzi Nashashibi was a jury member of the PhD thesis of Mr. David Pérez Morales - *Multi-sensor-based control in Intelligent Parking applications*, Université Nantes, December 6, 2019.

Fawzi Nashashibi was a reviewer of the PhD thesis of Mr. Mathieu Barbier - *Crossing of Road Intersections: Decision-Making Under Uncertainty for Autonomous Vehicles*, Université Grenoble Alpes, December 11, 2019.

Anne Verroust-Blondet was a jury member of the PhD thesis of Mr. David Sierra Gonzalez - *Towards Human-Like Prediction and Decision-Making for Automated Vehicles in Highway Scenarios*, Université Grenoble Alpes, 1 April 2019.

Anne Verroust-Blondet was a reviewer of the PhD thesis of Mr. Hernan Gonzalez - *Complex Dynamic scene analysis through multi-body motion segmentation - application to intelligent vehicles*, Université Paris-Saclay, 13 December 2019.

9.3. Popularization

9.3.1. Internal or external Inria responsibilities

- Fawzi Nashashibi was audited by a joint committee from the Parliament-Senat and cited in a OPECST report titled *"End of thermal vehicles in 2040: what realistic scenarios"*. Authors: Mrs. Huguette Tiegna (Parliament Deputy, Vice President) and Mr. Stéphane Piednoir (Senator). ISSN: 1249-3872.
- Participation of several team members to PAC V2X "Forum des acteurs" through oral scientific presentations - Paris, September 11, 2019.
- Participation of several team members in oral presentations and demonstrations with real mobile platforms during the *Journées du Patrimoine (Heritage Days)*, Rocquencourt, September 21, 2019.

9.3.2. Articles and contents

Fawzi Nashashibi was interviewed by Eric Gibory in *Le Monde*[0]: "Flottes d'entreprise : "De nouveaux services pour la voiture connectée vont apparaître avec la 5G"", October 24th 2019.

Fawzi Nashashibi was interviewed by Muriel Jaoën in *Les Echos*[0]: "Le véhicule autonome en site ouvert ? Pas avant 2050", February 8th 2019.

[0](https://www.lemonde.fr/economie/article/2019/10/24/flottes-d-entreprise-de-nouveaux-services-pour-la-voiture-connectee-vont-apparaitre-avec-la-5g_6016755_3234.html)

[0](https://www.lesechos.fr/tech-medias/intelligence-artificielle/fawzi-nashashibi-le-vehicule-autonome-en-site-ouvert-pas-avant-2050-963047)

9.3.3. Interventions

Raoul de Charette: *What is the limitation of artificial intelligence ?*, Arbre des Connaissances, Paris, France, February 12th 2019.
Raoul de Charette: *Dangers de l'intelligence artificielle ?*, Crégy-lès-Meaux, France, March 2019.
Raoul de Charette: *Qu'est-ce que l'intelligence artificielle ?*, Oissery, France, May 28th 2019.
Raoul de Charette: *Intelligence artificielle*, Arbre des Connaissances, Paris, France, June 17th 2019.
Fawzi Nashashibi: panelist, round table on the “Vehicle of the Future”, Maison de la Recherche, Paris, April 2019.
Fawzi Nashashibi: panelist, round table on the “Autonomous Vehicle”, Centrales-Supélec, Paris, September 16, 2019.

10. Bibliography

Major publications by the team in recent years

[1] Z. ALSAYED, G. BRESSON, A. VERROUST-BLONDET, F. NASHASHIBI.*2D SLAM Correction Prediction in Large Scale Urban Environments*, in "ICRA 2018 - International Conference on Robotics and Automation 2018", Brisbane, Australia, May 2018, https://hal.inria.fr/hal-01829091

[2] G. FAYOLLE, R. IASNOGORODSKI, V. A. MALYSHEV. , S. ASMUSSEN, P. W. GLYNN, Y. L. JAN (editors)*Random Walks in the Quarter Plane: Algebraic Methods, Boundary Value Problems, Applications to Queueing Systems and Analytic Combinatorics*, Probability Theory and Stochastic Modelling, Springer International Publishing, February 2017, vol. 40, 255, The first edition was published in 1999 [*DOI :* 10.1007/978-3-319-50930-3], https://hal.inria.fr/hal-01651919

[3] C. FLORES-PINO, P. MERDRIGNAC, R. DE CHARETTE, F. NAVAS, V. MILANÉS, F. NASHASHIBI.*A Cooperative Car-Following/Emergency Braking System With Prediction-Based Pedestrian Avoidance Capabilities*, in "IEEE Transactions on Intelligent Transportation Systems", June 2018, p. 1 - 10 [*DOI :* 10.1109/TITS.2018.2841644], https://hal.inria.fr/hal-01835121

[4] D. GONZALEZ BAUTISTA, J. PÉREZ, V. MILANÉS, F. NASHASHIBI.*A Review of Motion Planning Techniques for Automated Vehicles*, in "IEEE Transactions on Intelligent Transportation Systems", April 2016 [*DOI :* 10.1109/TITS.2015.2498841], https://hal.inria.fr/hal-01397924

[5] M. JARITZ, R. DE CHARETTE, M. TOROMANOFF, E. PEROT, F. NASHASHIBI.*End-to-End Race Driving with Deep Reinforcement Learning*, in "ICRA 2018 - IEEE International Conference on Robotics and Automation", Brisbane, Australia, May 2018, https://arxiv.org/abs/1807.02371 , https://hal.inria.fr/hal-01848067

[6] D. O. POP.*Multi-Task Cross-Modality Deep Learning for Pedestrian Risk Estimation*, INSA Rouen, November 2019

[7] P. DE BEAUCORPS.*Planification de trajectoire dans un environnement peu contraint et fortement dynamique*, Sorbonne Université, July 2019

Publications of the year

Doctoral Dissertations and Habilitation Theses

[8] D. O. POP.*Multi-Task Cross-Modality Deep Learning for Pedestrian Risk Estimation*, INSA Rouen, November 2019

[9] P. DE BEAUCORPS.*Planification de trajectoire dans un environnement peu contraint et fortement dynamique*, Sorbonne Université, July 2019

Articles in International Peer-Reviewed Journal

[10] F. NAVAS, V. MILANÉS.*Mixing V2V- and non-V2V-equipped vehicles in car following*, in "Transportation research. Part C, Emerging technologies", November 2019, vol. 108, p. 167-181 [*DOI :* 10.1016/J.TRC.2019.08.021], https://hal.inria.fr/hal-02392487

[11] D. O. POP.*Detection of pedestrian actions based on deep learning approach*, in "Studia Universitatis Babeş-Bolyai. Informatica", 2019, https://hal.inria.fr/hal-02414015

[12] D. O. POP, A. ROGOZAN, C. CHATELAIN, F. NASHASHIBI, A. BENSRHAIR.*Multi-Task Deep Learning for Pedestrian Detection, Action Recognition and Time to Cross Prediction*, in "IEEE Access", 2019, vol. 7, p. 149318-149327 [*DOI :* 10.1109/ACCESS.2019.2944792], https://hal.archives-ouvertes.fr/hal-02352800

[13] D. O. POP, A. ROGOZAN, F. NASHASHIBI, A. BENSRHAIR.*Pedestrian Recognition using Cross-Modality Learning in Convolutional Neural Networks*, in "IEEE Intelligent Transportation Systems Magazine", March 2019, https://hal.inria.fr/hal-02115347

International Conferences with Proceedings

[14] F. GHALLABI, G. EL-HAJ-SHHADE, M.-A. MITTET, F. NASHASHIBI.*LIDAR-Based road signs detection For Vehicle Localization in an HD Map*, in "IV'19 - IEEE Intelligent Vehicles Symposium", Paris, France, IEEE, June 2019, https://hal.archives-ouvertes.fr/hal-02117395

[15] F. GHALLABI, M.-A. MITTET, G. EL-HAJ-SHHADE, F. NASHASHIBI.*LIDAR-Based High Reflective Landmarks (HRL)s For Vehicle Localization in an HD Map*, in "ITSC 2019 - IEEE Intelligent Transportation Systems Conference", Auckland, New Zealand, October 2019, https://hal.archives-ouvertes.fr/hal-02185979

[16] S. S. HALDER, J.-F. LALONDE, R. DE CHARETTE.*Physics-Based Rendering for Improving Robustness to Rain*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, https://arxiv.org/abs/1908.10335 - ICCV 2019. Supplementary pdf / videos available on project page, https://hal.inria.fr/hal-02385436

[17] M. JARITZ, J. GU, H. SU.*Multi-view PointNet for 3D Scene Understanding*, in "Proceedings of the IEEE International Conference on Computer Vision Workshops", Seoul, South Korea, October 2019, https://arxiv.org/abs/1909.13603 - Geometry Meets Deep Learning Workshop, ICCV 2019, https://hal.inria.fr/hal-02387461

[18] K. MESSAOUD, I. YAHIAOUI, A. VERROUST-BLONDET, F. NASHASHIBI.*Non-local Social Pooling for Vehicle Trajectory Prediction*, in "IV 19 - IEEE Intelligent Vehicles Symposium 2019", Paris, France, June 2019, https://hal.inria.fr/hal-02160409

[19] K. MESSAOUD, I. YAHIAOUI, A. VERROUST-BLONDET, F. NASHASHIBI.*Relational Recurrent Neural Networks For Vehicle Trajectory Prediction*, in "ITSC 2019 - IEEE Intelligent transportation systems conference", Auckland, New Zealand, October 2019, https://hal.inria.fr/hal-02195180

[20] D. O. Pop, A. Rogozan, F. Nashashibi, A. Bensrhair.*Improving Pedestrian Recognition using Incremental Cross Modality Deep Learning*, in "European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning", Bruges, Belgium, April 2019, https://hal.inria.fr/hal-02115319

[21] L. Roldão, R. De Charette, A. Verroust-Blondet.*3D Surface Reconstruction from Voxel-based Lidar Data*, in "ITSC 2019 - IEEE Intelligent Transportation Systems Conference", Auckland, New Zealand, October 2019, https://arxiv.org/abs/1906.10515 - IEEE Intelligent Transportation Systems Conference (ITSC) 2019, https://hal.inria.fr/hal-02165672

[22] M. Sangare, D.-V. N. Nguyen, S. Banerjee, P. Muhlethaler, S. Bouzefrane.*Comparing different Machine-Learning techniques to predict Vehicles' Positions using the received Signal Strength of periodic messages*, in "WMNC 2019. 12th IFIP Wireless and Mobile Networking Conference", Paris, France, September 2019, https://hal.archives-ouvertes.fr/hal-02178360

Research Reports

[23] G. Le Lann.*Cyberphysical Constructs and Concepts for Fully Automated Networked Vehicles*, Inria Paris-Rocquencourt, October 2019, n^{O} RR-9297, https://hal.inria.fr/hal-02318242

Other Publications

[24] R. De Charette, S. Manitsaris.*3D Reconstruction of Deformable Revolving Object under Heavy Hand Interaction*, November 2019, https://arxiv.org/abs/1908.01523 - Submitted to journal, https://hal.inria.fr/hal-02385443

[25] G. Fayolle.*A note on the connection between product-form Jackson networks and counting lattice walks in the quarter plane*, January 2020, working paper or preprint, https://hal.inria.fr/hal-02415746

[26] G. Fayolle, J.-M. Lasgouttes, C. Flores-Pino.*Stability and String Stability of Car-following Models with Reaction-time Delay*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02413721

[27] M. Jaritz, T.-H. Vu, R. De Charette, E. Wirbel, P. Pérez.*xMUDA: Cross-Modal Unsupervised Domain Adaptation for 3D Semantic Segmentation*, December 2019, https://arxiv.org/abs/1911.12676 - For a demo video, see http://tiny.cc/xmuda, https://hal.inria.fr/hal-02388974

[28] F. Pizzati, R. De Charette, M. Zaccaria, P. Cerri.*Domain Bridge for Unpaired Image-to-Image Translation and Unsupervised Domain Adaptation*, January 2020, https://arxiv.org/abs/1910.10563 - Accepted at WACV 2020 [*DOI :* 10.10563], https://hal.archives-ouvertes.fr/hal-02436218

References in notes

[29] R. Castiñeira, J. E. Naranjo, M. Gil, F. Jimenez, P. Serra, A. Valejo, A. Asvadi, C. Premebida, M. Abualhoul, F. Nashashibi.*AUTOCITS – Regulation study for interoperability in the adoption of autonomous driving in European urban nodes*, in "TRA 2018 - Transport Research Arena", Vienna, Austria, April 2018, https://hal.inria.fr/hal-01898256

[30] G. Fayolle, C. Furtlehner.*About Hydrodynamic Limit of Some Exclusion Processes via Functional Integration*, in "Int. Math. Conf. “50 Years of IPPI”", Moscow, Institute for Information Transmission Problems (Russian Academy of Sciences), July 2011, Proceedings on CD. ISBN 978-5-901158-15-9, http://hal.inria.fr/hal-00662674

[31] G. FAYOLLE, R. IASNOGORODSKI.*Random Walks in the Quarter-Plane: Advances in Explicit Criterions for the Finiteness of the Associated Group in the Genus 1 Case*, in "Markov Processes and Related Fields", December 2015, vol. 21, n° 4, Accepted for publication in the journal MPRF (Markov Processes and Related Fields), https://hal.inria.fr/hal-01086684

[32] G. FAYOLLE, J.-M. LASGOUTTES.*Asymptotics and Scalings for Large Product-Form Networks via the Central Limit Theorem*, in "Markov Processes and Related Fields", 1996, vol. 2, n° 2, p. 317-348

[33] C. FURTLEHNER, Y. HAN, J.-M. LASGOUTTES, V. MARTIN, F. MARCHAL, F. MOUTARDE.*Spatial and Temporal Analysis of Traffic States on Large Scale Networks*, in "13th International IEEE Conference on Intelligent Transportation Systems ITSC'2010", Madère, Portugal, September 2010, https://hal-mines-paristech.archives-ouvertes.fr/hal-00527481

[34] G. LE LANN.*Safety in Vehicular Networks-On the Inevitability of Short-Range Directional Communications*, in "14th International Conference ADHOC-NOW, 2015", Athens, Greece, S. PAPAVASSILIOU, S. RUEHRUP (editors), Ad Hoc, Mobile, and Wireless Networks, Springer, June 2015, vol. Lecture Notes in Computer Science (LNCS), n° 9143, 14, Mobile Ad Hoc Networks [*DOI :* 10.1007/978-3-319-19662-6_24], https://hal.inria.fr/hal-01172595

[35] V. MARTIN, C. FURTLEHNER, Y. HAN, J.-M. LASGOUTTES.*GMRF Estimation under Topological and Spectral Constraints*, in "7th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases", Nancy, France, T. CALDERS, F. ESPOSITO, E. HÜLLERMEIER, R. MEO (editors), Lecture Notes in Computer Science, Springer Berlin Heidelberg, September 2014, vol. 8725, p. 370-385 [*DOI :* 10.1007/978-3-662-44851-9_24], https://hal.archives-ouvertes.fr/hal-01065607

[36] V. MARTIN, J.-M. LASGOUTTES, C. FURTLEHNER.*Latent binary MRF for online reconstruction of large scale systems*, in "Annals of Mathematics and Artificial Intelligence", 2016, vol. 77, n° 1, p. 123-154

[37] P. DE BEAUCORPS, A. VERROUST-BLONDET, R. PONCELET, F. NASHASHIBI.*RIS : A Framework for Motion Planning among Highly Dynamic Obstacles*, in "ICARCV 2018 - 15th International Conference on Control, Automation, Robotics and Vision", Singapour, Singapore, November 2018, https://hal.inria.fr/hal-01903318

Project-Team SECRET

Security, Cryptology and Transmissions

RESEARCH CENTER
Paris

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team SECRET

Creation of the Project-Team: 2008 July 01, end of the Project-Team: 2019 November 30

Keywords:

Computer Science and Digital Science:

A3.1.5. - Control access, privacy
A4. - Security and privacy
A4.2. - Correcting codes
A4.3. - Cryptography
A4.3.1. - Public key cryptography
A4.3.2. - Secret key cryptography
A4.3.3. - Cryptographic protocols
A4.3.4. - Quantum Cryptography
A7.1. - Algorithms
A7.1.4. - Quantum algorithms
A8.1. - Discrete mathematics, combinatorics
A8.6. - Information theory

Other Research Topics and Application Domains:

B6.4. - Internet of things
B6.5. - Information systems
B9.5.1. - Computer science
B9.5.2. - Mathematics
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

Anne Canteaut [Team leader, Inria, Senior Researcher, HDR]
André Chailloux [Inria, Researcher]
Pascale Charpin [Inria, Emeritus, HDR]
Gaëtan Leurent [Inria, Researcher]
Anthony Leverrier [Inria, Researcher, HDR]
María Naya Plasencia [Inria, Senior Researcher, HDR]
Léo Perrin [Inria, Researcher, from Sep 2019]
Nicolas Sendrier [Inria, Senior Researcher, HDR]
Jean-Pierre Tillich [Inria, Senior Researcher, HDR]

Faculty Members

Magali Bardet [Univ de Rouen, Associate Professor, from Sept 2019]
Christina Boura [Univ de Versailles Saint-Quentin-en-Yvelines, Associate Professor, from Sept 2019]

PhD Students

Xavier Bonnetain [Sorbonne Université, PhD Student]
Rémi Bricout [Sorbonne Université, PhD Student]
Kevin Carrier [Ministère de la Défense, PhD Student9]
Daniel Coggia [DGA, PhD Student]

Thomas Debris [Sorbonne Université PhD Student, until Sep 2019]
Simona Etinski [Univ Denis Diderot, PhD Student, from Oct 2019]
Antonio Florez Gutierrez [Inria, PhD Student, from Sep 2019]
Shouvik Ghorai [Sorbonne Université]
Antoine Grospellier [Sorbonne Université, PhD Student, until Aug 2019]
Lucien Grouès [Sorbonne Université, PhD Student, from Oct 2019]
Matthieu Lequesne [Univ Pierre et Marie Curie, PhD Student]
Vivien Londe [Univ de Bordeaux, PhD Student, until Aug 2019]
Rocco Mora [Sorbonne Université, PhD Student, from Oct 2019]
Andrea Olivo [Inria, PhD Student]
André Schrottenloher [Inria, PhD Student]
Ferdinand Sibleyras [Inria, PhD Student]
Valentin Vasseur [Univ René Descartes, PhD Student]

Post-Doctoral Fellows

Christophe Vuillot [Inria, Post-Doctoral Fellow, from Nov 2019]
Simon Apers [Inria, Post-Doctoral Fellow]
Ivan Bardet [Inria, Post-Doctoral Fellow, from Mar 2019]
Léo Perrin [Inria, Post-Doctoral Fellow, until Aug 2019]

Visiting Scientists

Akinori Hosoyamada [NTT Secure Platform Labaratories, Japan, Mar 2019 and Nov 2019]
Mustafa Mahmoud Mohammed Kairallah [NTU, Singapore, Jul 2019]
Thomas Peyrin [NTU, Singapore, Jan and June 2019]
Yu Sasaki [NTT Secure Platform Labaratories, Japan, Nov 2019]
Shizhu Tian [Chinese Academy of Sciences, until Sep 2019]

Administrative Assistant

Christelle Guiziou [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Presentation and scientific foundations

The research work within the project-team is mostly devoted to the design and analysis of cryptographic algorithms, in the classical or in the quantum setting. This work is essential since the current situation of cryptography is rather fragile. Many cryptographic protocols are now known whose security can be formally proved assuming that the involved cryptographic primitives are ideal (random oracle model, ideal cipher model...). However, the security of the available primitives has been either threatened by recent progress in cryptanalysis or by the possible invention of a large quantum computer. In other words, there is usually no concrete algorithm available to instantiate in practice the ideal "black boxes" used in these protocols!

In this context, our research work focuses on both families of cryptographic primitives, *symmetric* and *asymmetric* primitives.

2.2. Main topics

Our domain in cryptology includes the analysis and the design of

- symmetric primitives (a.k.a. secret-key algorithms),
- public-key primitives based on hard problems coming from coding theory which are likely to be resistant against a quantum computer,
- quantum cryptographic protocols whose security does not rely on computational assumptions but on the laws of quantum physics.

3. Research Program

3.1. Scientific foundations

Our approach relies on a competence whose impact is much wider than cryptology. Our tools come from information theory, discrete mathematics, probabilities, algorithmics, quantum physics... Most of our work mixes fundamental aspects (study of mathematical objects) and practical aspects (cryptanalysis, design of algorithms, implementations). Our research is mainly driven by the belief that discrete mathematics and algorithmics of finite structures form the scientific core of (algorithmic) data protection.

3.2. Symmetric cryptology

Symmetric techniques are widely used because they are the only ones that can achieve some major features such as high-speed or low-cost encryption, fast authentication, and efficient hashing. It is a very active research area which is stimulated by a pressing industrial demand. The process which has led to the new block cipher standard AES in 2001 was the outcome of a decade of research in symmetric cryptography, where new attacks have been proposed, analyzed and then thwarted by some appropriate designs. However, even if its security has not been challenged so far, it clearly appears that the AES cannot serve as a Swiss knife in all environments. In particular an important challenge raised by several new applications is the design of symmetric encryption schemes with some additional properties compared to the AES, either in terms of implementation performance (low-cost hardware implementation, low latency, resistance against side-channel attacks...) or in terms of functionalities (like authenticated encryption). The past decade has then been characterized by a multiplicity of new proposals. This proliferation of symmetric primitives has been amplified by several public competitions (eSTREAM, SHA-3, CAESAR...) which have encouraged innovative constructions and promising but unconventional designs. We are then facing up to a very new situation where implementers need to make informed choices among more than 40 lightweight block ciphers [0] or 57 new authenticated-encryption schemes [0]. Evaluating the security of all these proposals has then become a primordial task which requires the attention of the community.

In this context we believe that the cryptanalysis effort cannot scale up without an in-depth study of the involved algorithms. Indeed most attacks are described as ad-hoc techniques dedicated to a particular cipher. To determine whether they apply to some other primitives, it is then crucial to formalize them in a general setting. Our approach relies on the idea that a unified description of generic attacks (in the sense that they apply to a large class of primitives) is the only methodology for a precise evaluation of the resistance of all these new proposals, and of their security margins. In particular, such a work prevents misleading analyses based on wrong estimations of the complexity or on non-optimized algorithms. It also provides security criteria which enable designers to guarantee that their primitive resists some families of attacks. The main challenge is to provide a generic description which captures most possible optimizations of the attack.

3.3. Code-based cryptography

Public-key cryptography is one of the key tools for providing network security (SSL, e-commerce, e-banking...). The security of nearly all public-key schemes used today relies on the presumed difficulty of two problems, namely factorization of large integers or computing the discrete logarithm over various groups. The hardness of those problems was questioned in 1994 [0] when Shor showed that a quantum computer could solve them efficiently. Though large enough quantum computers that would be able to threaten the existing cryptosystems do not exist yet, the cryptographic research community has to get ready and has to prepare alternatives. This line of work is usually referred to as *post-quantum cryptography*. This has become a prominent research field. Most notably, an international call for post-quantum primitives [0] has been launched by the NIST, with a submission deadline in November 2017.

[0] 35 are described on https://www.cryptolux.org/index.php/Lightweight_Block_Ciphers.
[0] see http://competitions.cr.yp.to/caesar-submissions.html
[0] P. Shor, *Algorithms for quantum computation: Discrete logarithms and factoring*, FOCS 1994.
[0] http://csrc.nist.gov/groups/ST/post-quantum-crypto/

The research of the project-team in this field is focused on the design and cryptanalysis of cryptosystems making use of coding theory. Code-based cryptography is one the main techniques for post-quantum cryptography (together with lattice-based, multivariate, or hash-based cryptography).

3.4. Quantum information

The field of quantum information and computation aims at exploiting the laws of quantum physics to manipulate information in radically novel ways. There are two main applications:

(i) quantum computing, that offers the promise of solving some problems that seem to be intractable for classical computers such as for instance factorization or solving the discrete logarithm problem;

(ii) quantum cryptography, which provides new ways to exchange data in a provably secure fashion. For instance it allows key distribution by using an authenticated channel and quantum communication over an unreliable channel with information-theoretic security, in the sense that its security can be proven rigorously by using only the laws of quantum physics, even with all-powerful adversaries.

Our team deals with quantum coding theoretic issues related to building a large quantum computer and with quantum cryptography. The first part builds upon our expertise in classical coding theory whereas the second axis focuses on obtaining security proofs for quantum protocols or on devising quantum cryptographic protocols (and more generally quantum protocols related to cryptography). A close relationship with partners working in the whole area of quantum information processing in the Parisian region has also been developed through our participation to the Fédération de Recherche "PCQC" (Paris Centre for Quantum Computing).

4. Application Domains

4.1. Cryptographic primitives

Our major application domain is the design of cryptographic primitives, especially for platforms with restricting implementation requirements. For instance, we aim at recommending (or designing) low-cost (or extremely fast) encryption schemes, or primitives which remain secure against quantum computers.

4.2. Code Reconstruction

To evaluate the quality of a cryptographic algorithm, it is usually assumed that its specifications are public, as, in accordance with Kerckhoffs principle, it would be dangerous to rely, even partially, on the fact that the adversary does not know those specifications. However, this fundamental rule does not mean that the specifications are known to the attacker. In practice, before mounting a cryptanalysis, it is necessary to strip off the data. This reverse-engineering process is often subtle, even when the data formatting is not concealed on purpose. A typical case is interception: some raw data, not necessarily encrypted, is observed out of a noisy channel. To access the information, the whole communication system has first to be disassembled and every constituent reconstructed. A transmission system actually corresponds to a succession of elements (symbol mapping, scrambler, channel encoder, interleaver...), and there exist many possibilities for each of them. In addition to the "preliminary to cryptanalysis" aspect, there are other links between those problems and cryptology. They share some scientific tools (algorithmics, discrete mathematics, probability...), but beyond that, there are some very strong similarities in the techniques.

5. Highlights of the Year

5.1. Highlights of the Year

- **Keynote at FSE 2019:** María Naya Plasencia bas been an invited keynote speaker at FSE 2019 in Paris.

- **NIST competition on post-quantum cryptography:** The members of the project-team have submitted 5 candidates to the NIST competition on post-quantum cryptography. After a first selection, three of our candidates have been moved to the second round of the competition, which includes a total of 26 candidates.
- **NIST competition on lightweight cryptography:** The members of the project-team are involved in the design of 3 authenticated encryption schemes submitted to the NIST lightweight competition. These three ciphers are among the 32 candidates which have been move to the second round of the competition.

5.1.1. Awards

- María Naya Plasencia was awarded the Inria - Académie des Sciences prize for young researchers https://www.academie-sciences.fr/fr/Laureats/prix-inria-academie-des-sciences-2019-vincent-hayward-equipe-scikit-learn-et-maria-naya-plasencia.html
- Anne Canteaut has been made doctor honoris causa of the University of Bergen (Norway), October 2019 https://www.uib.no/en/news/129910/ten-new-honorary-doctorates

BEST PAPERS AWARDS :

[31]
L. PERRIN.*Partitions in the S-Box of Streebog and Kuznyechik*, in "IACR Transactions on Symmetric Cryptology", March 2019, vol. 2019, n^{o} 1, p. 302-329 [*DOI :* 10.13154/TOSC.V2019.I1.302-329], https://hal.inria.fr/hal-02396814

[49]
T. DEBRIS-ALAZARD, N. SENDRIER, J.-P. TILLICH.*Wave: A New Family of Trapdoor One-Way Preimage Sampleable Functions Based on Codes*, in "ASIACRYPT 2019 - 25th International Conference on the Theory and Application of Cryptology and Information Security", Kobe, Japan, LNCS, Springer, November 2019, vol. 11921, p. 21-51 [*DOI :* 10.1007/978-3-030-34578-5_2], https://hal.inria.fr/hal-02424057

6. New Software and Platforms

6.1. CFS

FUNCTIONAL DESCRIPTION: Reference implementation of parallel CFS (reinforced version of the digital signature scheme CFS). Two variants are proposed, one with a « bit-packing » finite field arithmetic and an evolution with a « bit-slicing » finite-field arithmetic (collaboration with Peter Schwabe). For 80 bits of security the running time for producing one signature with the « bit-packing » variant is slightly above one second. This is high but was still the fastest so far. The evolution with the « bit-slicing » arithmetic produces the same signature in about 100 milliseconds.

- Participants: Grégory Landais and Nicolas Sendrier
- Contact: Nicolas Sendrier
- URL: https://gforge.inria.fr/projects/cfs-signature/

6.2. Collision Decoding

KEYWORDS: Algorithm - Binary linear code

FUNCTIONAL DESCRIPTION: Collision Decoding implements two variants of information set decoding : Stern-Dumer, and MMT. To our knowledge it is the best full-fledged open-source implementation of generic decoding of binary linear codes. It is the best generic attack against code-based cryptography.

- Participants: Grégory Landais and Nicolas Sendrier
- Contact: Nicolas Sendrier
- URL: https://gforge.inria.fr/projects/collision-dec/

6.3. ISDF

FUNCTIONAL DESCRIPTION: Implementation of the Stern-Dumer decoding algorithm, and of a variant of the algorithm due to May, Meurer and Thomae.

- Participants: Grégory Landais and Nicolas Sendrier
- Contact: Anne Canteaut
- URL: https://gforge.inria.fr/projects/collision-dec/

7. New Results

7.1. Symmetric cryptology

Participants: Xavier Bonnetain, Christina Boura, Anne Canteaut, Daniel Coggia, Pascale Charpin, Daniel Coggia, Gaëtan Leurent, María Naya Plasencia, Léo Perrin, André Schrottenloher, Ferdinand Sibleyras.

7.1.1. Block ciphers

Our recent results mainly concern either the analysis or the design of lightweight block ciphers.

Recent results:

- Design of SATURNIN a new lightweight block cipher for authenticated encryption [74], which is resistant to quantum cryptanalysis. SATURNIN has been submitted to the NIST competition for lightweight cryptography, and has been selected for the 2nd round of the competition [0].
- Mixture-differential distinguishers on AES-like ciphers [18].
- Cryptanalysis of the Sbox of the Russian standards, Streebog and Kuznyechik [31], [56]. This work by L. Perrin received the best paper award at *FSE 2019*. Moreover, L. Perrin has been invited to present his results to AFNOR. He is involved in the international standardization processes in symmetric cryptography [50], [86] and has been invited to ISO meetings on this topic.
- The work on the Streebog Sbox has led to a more general study on tools for quantifying anomalies in Sboxes [44].
- Design of BISON, the first concrete block cipher following the whitened swap-or-not construction [46].

7.1.2. MACs and hash functions

The international research effort related to the selection of the new hash function standard SHA-3 has led to many important results and to a better understanding of the security offered by hash functions. However, hash functions are used in a huge number of applications with different security requirements, and also form the building-blocks of some other primitives, like MACs.

Recent results:

- Chosen-prefix collision attack on SHA-1 [52]: A chosen-prefix collision attack is a stronger variant of a collision attack, where an arbitrary pair of challenge prefixes are turned into a collision. Chosen-prefix collisions are usually significantly harder to produce than (identical-prefix) collisions, but the practical impact of such an attack is much larger. G. Leurent and T. Peyrin proposed new techniques to turn collision attacks into chosen-prefix collision attacks, and present such an attack against SHA-1 with complexity between $2^{66.9}$ and $2^{69.4}$ (depending on assumptions about the cost of finding near-collision blocks).
- Design of lighweight MACs from universal hash functions [51]. Many constructions of MACs used in practice (such as GMAC or Poly1305-AES) follow the Wegman-Carter-Shoup construction, which is only secure up to 2^{64} queries with a 128-bit state. S. Duval and G. Leurent proposed new constructions to reach security beyond the birthday bound, and proposed a concrete instantiation, with very good performances on ARM micro-controllers.

[0] https://csrc.nist.gov/CSRC/media/Projects/lightweight-cryptography/documents/round-2/spec-doc-rnd2/saturnin-spec-round2.pdf

7.1.3. Cryptographic properties and construction of appropriate building blocks

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

Recent results:

- Differential Equivalence of Sboxes: C. Boura, A. Canteaut and their co-authors have studied two notions of differential equivalence of Sboxes corresponding to the case when the functions have the same difference table, or when their difference tables have the same support [19]. They proved that these two notions do not coincide, and that they are invariant under some classical equivalence relations like EA and CCZ equivalence. They also proposed an algorithm for determining the whole equivalence class of a given function.
- Boomerang Uniformity of Sboxes: The boomerang attack is a cryptanalysis technique against block ciphers which combines two differentials for the upper part and the lower part of the cipher. The Boomerang Connectivity Table (BCT) is a tool introduced by Cid *et al.* at Eurocrypt 2018 for analysing the dependency between these two differentials. C. Boura and A. Canteaut have provided an in-depth analysis of BCT, by studying more closely differentially 4-uniform Sboxes. More recently, C. Boura, L. Perrin and S. Tian have obtained new results on the boomerang uniformity of several constructions of Sboxes [57].
- CCZ equivalence of Sboxes: A. Canteaut and L. Perrin have characterized CCZ-equivalence as a property of the zeroes in the Walsh spectrum of an Sbox (or equivalently in their DDT). They used this framework to show how to efficiently upper bound the number of distinct EA-equivalence classes in a given CCZ-equivalence class. More importantly, they proved that CCZ-equivalence can be reduced to the association of EA-equivalence and an operation called twisting. They then revisited several results from the literature on CCZ-equivalence and showed how they can be interpreted in light of this new framework [21], [58].
- Links between linear and differential properties of Sboxes: P. Charpin together with J. Peng has established new links between the differential uniformity and the nonlinearity of some Sboxes in the case of two-valued functions and quadratic functions. More precisely, they have exhibited a lower bound on the nonlinearity of monomial permutations depending on their differential uniformity, as well as an upper bound in the case of differentially two-valued functions [27]
- Study of the properties of the error-correcting codes associated to differentially 4-uniform Sboxes [26]. Most notably, this work analyzes the relationship between the number of low-weight codewords and the nonlinearity of the corresponding Sbox.
- Study of crooked and weakly-crooked functions [35]: Crooked functions form a family of APN functions whose derivaties take they values in an (affine) hyperplane.
- APN functions with the butterfly construction [22], [34]: the butterfly construction, originally introduced by Perrin et al., is a general construction which includes the only known example of APN permutation operating on an even number of variables. A. Canteaut, L. Perrin and S. Tian have proved that the most recent generalization of this construction does not include any other APN function when the number of variables exceeds six.

7.1.4. Modes of operation and generic attacks

In order to use a block cipher in practice, and to achieve a given security notion, a mode of operation must be used on top of the block cipher. Modes of operation are usually studied through provable security, and we know that their use is secure as long as the underlying primitive is secure, and we respect some limits on the amount of data processed. The analysis of generic attack helps us understand what happens when the hypotheses of the security proofs do not hold, or the corresponding limits are not respected. Comparing proofs and attacks also shows gaps where our analysis is incomplete, and when improved proof or attacks are required.
Recent results:

- Low-memory attacks against the 2-round Even-Mansour construction, using the 3-xor problem [41]: G. Leurent and F. Sibleyras proved that attacking the 2-round Even-Mansour construction with blocksize n is related to the 3-XOR problem with elements on size $2n$. Then, they exhibited the first generic attacks on this construction where both the data and the memory complexity are significantly lower than 2^n.
- Generic attacks against the tweakable FX-construction [55]: F. Sibleyras exhibited a generic attack on the general tweakable iterated FX-construction, which provides an upper-bound on its security. Most notably, for two rounds, this upper bound matches the proof of the particular case of XHX2 by Lee and Lee at Asiacrypt 2018, thus proving for the first time its tightness.
- Modes for authenticated encryption: Besides the design of new lightweight authenticated encryption schemes, we also analyzed some modes of operation in case of release of unverified plaintext (RUP). Indeed, in this setting, an adversary gets separated access to the decryption and verification functionality, and has more power in breaking the scheme. Our results include a forgery attack against the GCM-RUP mode of operation [54], and the design of a new lightweight deterministic scheme, named ANYDAE, which is particularly efficient for short messages, and achieves both conventional security and RUP security [24].
- Generic attacks on hash combiners [15]: G. Leurent and his co-authors analyzed the security of hash combiners, i.e. of procedures that combine two or more hash functions in a way that is hopefully more secure than each of the underlying hash functions, or at least remains secure as long as one of them is secure. They found generic attacks on the XOR combiner, on the concatenation of two Merkle-Damgård hash functions and on the Zipper hash and on the Hash-Twice combiners when they both use Merkle-Damgård hash constructions.

7.2. Code-based cryptography

Participants: Magali Bardet, Kevin Carrier, André Chailloux, Thomas Debris, Matthieu Lequesne, Rocco Mora, Nicolas Sendrier, Jean-Pierre Tillich, Valentin Vasseur.

In recent years, there has been a substantial amount of research on quantum computers. Such computers would be a major threat for all the public-key cryptosystems used in practice, since all these systems rely on the hardness of integer factoring or discrete logarithms, and these problems are easy on a quantum computer. This has prompted NIST to launch a standardization process in 2017 for quantum-safe alternatives to those cryptosystems. This concerns all three major asymmetric primitives, namely public-key encryption schemes, key-exchange protocols and digital signatures. There were 69 valid submissions to this call in November 2017, with numerous lattice-based, code-based and multivariate-cryptography submissions and some submission based either on hashing or on supersingular elliptic curve isogenies. NIST expects to perform multiple rounds of evaluation, over a period of three to five years. The goal of this process is to select a number of acceptable candidate cryptosystems for standardization. The second round of evaluation started in February 2019.

The research of the project-team in this field is focused on the design and cryptanalysis of cryptosystems making use of coding theory. The first cryptosystem based on error-correcting codes was a public-key encryption scheme proposed by McEliece in 1978; a dual variant was proposed in 1986 by Niederreiter. We proposed the first (and only) digital signature scheme in 2001. Those systems enjoy very interesting features (fast encryption/decryption, short signature, good security reduction) but also have their drawbacks (large public key, encryption overhead, expensive signature generation). Our recent work on code-based cryptography has to be seen in the context of the recently launched NIST competition for quantum-safe primitives. We have proposed five code-based candidates to the NIST call for the first two primitives, namely public key encryption and key exchange protocols. Our contributions in this area are two-fold and consist in:

- designing and analysis new code-based solutions;
- cryptanalyzing code-based schemes, especially candidates to the NIST competition.

We have also been organizing since 2015 a working group held every month or every two months on code-based cryptography that structures the French efforts on this topic: every meeting is attended by most of the groups working in France on this topic (project-team GRACE, University of Bordeaux, University of Limoges, University of Rennes and University of Rouen).

7.2.1. Design of new code-based solutions

The members of the project-team have submitted several candidates to the NIST competition and have designed new code-based primitives.

Recent results:

- Design of a new code-based signature scheme [49]: T. Debris, N. Sendrier and JP Tillich recently proposed a "hash-and-sign" code-based signature scheme called WAVE, which uses a family of ternary generalized (U, U + V) codes. WAVE achieves existential unforgeability under adaptive-chosen-message attacks in the random oracle model with a tight reduction to two assumptions from coding theory: one is a distinguishing problem that is related to the trapdoor inserted in the scheme, the other one is a multiple-target version of syndrome decoding. This scheme enjoys efficient signature and verification algorithms. For 128-bit security, signature are 8000-bit long and the public-key size is slightly smaller than one megabyte.
- Analysis of the ternary Syndrome Decoding problem [45]: R. Bricout, A. Chailloux, T. Debris and M. Lequesne have performed an algorithmic study of this decoding problem in large weight, which corresponds to the underlying problem in the WAVE signature scheme. Most notably, their study results in an update of the Wave parameters. It also shows that ternary Syndrome Decoding with large weight is a really harder problem than the binary Syndrome Decoding problem, and could have several applications for the design of code-based cryptosystems.

7.2.2. Cryptanalysis of code-based schemes

Recent results:

- Attack against RLCE [48]: M. Lequesne and JP Tillich, together with A. Couvreur, recently presented a key-recovery attack against the Random Linear Code Encryption (RLCE) scheme recently submitted by Y. Wang to the NIST competition. This attack recovers the secret-key for all the short key-parameters proposed by the author. It uses a polynomial-time algorithm based on a square code distinguisher.
- Analysis of an encryption scheme based on the rank syndrome decoding problem [61]: D. Coggia and A. Couvreur presented an attack against a cryptosystem proposed proposed by Loidreau, which used an intermediary version between Gabidulin codes and LRPC codes. This attack has polynomial time for some parameters of the scheme.
- Decoding algorithm for codes with a non-trivial automorphism group [47]: R. Canto-Torres and JP Tillich presented an algorithm which is able to speed up the decoding of a code with a non-trivial automorphism group. For a certain range of parameters, this results in a decoding that is faster by an exponential factor in the code length when compared to the best algorithms for decoding generic linear codes. This algorithm was then used to break several proposals of public-key cryptosystems based on codes with a non-trivial automorphism group.

7.3. Quantum Information

Participants: Simon Apers, Ivan Bardet, Xavier Bonnetain, Rémi Bricout, André Chailloux, Simona Etinski, Antonio Florez Gutierrez, Shouvik Ghorai, Antoine Grospellier, Lucien Grouès, Anthony Leverrier, Vivien Londe, María Naya Plasencia, Andrea Olivo, Jean-Pierre Tillich, André Schrottenloher, Christophe Vuillot.

Our research in quantum information focusses on several axes: quantum codes with the goal of developing better error-correction strategies to build large quantum computers, quantum cryptography which exploits the laws of quantum mechanics to derive security guarantees, relativistic cryptography which exploits in addition the fact that no information can travel faster than the speed of light and finally quantum cryptanalysis which investigates how quantum computers could be harnessed to attack classical cryptosystems.

7.3.1. Quantum codes

Protecting quantum information from external noise is an issue of paramount importance for building a quantum computer. It also worthwhile to notice that all quantum error-correcting code schemes proposed up to now suffer from the very same problem that the first (classical) error-correcting codes had: there are constructions of good quantum codes, but for the best of them it is not known how to decode them in polynomial time.

Two PhD theses have been defended this year within the project-team on this topic. First, Antoine Grospellier, co-advised by A. Leverrier and O. Fawzi (Ens Lyon), studied efficient decoding algorithms for quantum LDPC codes [13]. Beyond their intrinsic interest for channel-coding problems, such algorithms would be particularly relevant in the context of quantum fault-tolerance, since they would allow to considerably reduce the required overhead to obtain fault-tolerance in quantum computation. Vivien Londe, co-advised by A. Leverrier and G. Zémor (IMB), worked on the design of better quantum LDPC codes [14]: the main idea is to generalize the celebrated toric code of Kitaev by considering cellulations of manifolds in higher dimensions. A surprising result was that this approach leads to a much better behaviour than naively expected and a major challenge is to explore the mathematics behind this phenomenon in order to find even better constructions, or to uncover potential obstructions.

Lucien Grouès, who did an internship this summer in the project-team, has recently started a PhD with A. Leverrier and O. Fawzi on decoding quantum LDPC codes, and preliminary numerical results have already appeared in [62].

Ivan Bardet joined the project-team as a postdoc in March 2019, and will start a starting research position in 2020. His research focusses on the study of open-system dynamics as well as mixing times of Markovian dissipative evolutions with the goal of better understanding the lifetime of quantum memories.

Recent results:

- Decoding algorithms for Hypergraph Product Codes [62]: this work deals with numerical simulation of several variants of the SMALL-SET-FLIP decoder for hypergraph product codes. While this decoder had already been studied analytically in previous work in the regime of extremely low noise, we are focussing here on understanding its performance for a realistic noise model.
- Towards Low Overhead Magic State Distillation [30]: the major source of overhead in quantum fault-tolerance usually lies in the primitive called magic state distillation which takes a number of noisy versions of a specific quantum state and prepares a new state with less noise. An important question is to understand how efficient this procedure can be. In this work, we prove that magic state distillation can perform much more efficiently than expected when working with quantum systems of large dimension instead of qubits.

7.3.2. Quantum cryptography

Quantum cryptography exploits the laws of quantum physics to establish the security of certain cryptographic primitives. The most studied one is certainly quantum key distribution, which allows two distant parties to establish a secret using an untrusted quantum channel. Our activity in this field is particularly focussed on protocols with continuous variables, which are well-suited to implementations. The interest of continuous variables for quantum cryptography was recently recognized by being awarded a 10 M€ funding from the Quantum Flagship and SECRET contributes to this project by studying the security of new key distribution protocols.

Recent results:

- Security proof for two-way continuous-variable quantum key distribution [28]: while many quantum key distribution protocols are one-way in the sense that quantum information is sent from one party to the other, it can be beneficial in terms of performance to consider two-way protocols where the quantum states perform a round-trip between the two parties. In this paper, we show how to exploit the symmetries of the protocols in phase-space to establish their security against the most general attacks allowed by quantum theory.
- Asymptotic security of continuous-variable quantum key distribution with a discrete modulation [29]: in this work, we establish a lower bound on the secret key rate of a practical quantum key distribution protocol that will be implemented in the context of the H2020 project CiViQ.

7.3.3. Quantum cryptanalysis of symmetric primitives and quantum algorithms

Symmetric cryptography seems at first sight much less affected in the post-quantum world than asymmetric cryptography: its main known threat seemed for a long time Grover's algorithm, which allows for an exhaustive key search in the square root of the normal complexity. For this reason, it was usually believed that doubling key lengths suffices to maintain an equivalent security in the post-quantum world. However, a lot of work is certainly required in the field of symmetric cryptography in order to "quantize" the classical families of attacks in an optimized way, as well as to find new dedicated quantum attacks. M. Naya Plasencia has been awarded an ERC Starting grant for her project named QUASYModo on this topic.

In parallel to this work, S. Apers is developing generic quantum algorithms solving combinatorial problems, notably in graphs. He also recently proposed a unified framework of quantum walk search, that will likely find applications in the context of quantum cryptanalysis.

Recent results:

- Quantum algorithm for the k-XOR problem and for list merging: The k-XOR (or generalized birthday) problem aims at finding k elements of n-bits, drawn at random, such that the XOR of all of them is 0. The algorithms proposed by Wagner more than 15 years ago remain the best known classical algorithms for solving it, when disregarding logarithmic factors. A. Chailloux, M. Naya-Plasencia and A. Schrottenloher, together with L. Grassi, studied this problem in the quantum setting and provided algorithms with the best known quantum time-complexities [38], [39].
- Quantum security of AES [17]: In order to determine the post-quantum secuirty margin of AES-256, X. Bonnetain and M. Naya-Plasencia have proposed generalized and quantized versions of the best known cryptanalysis on reduced-round versions of AES-256, including a quantum Demirci-Selçuk meet-in-the-middle attack.
- Quantum attacks without superposition queries : In symmetric cryptanalysis, the model of superposition queries has led to surprising results, but the practical implications of these attacks remain blurry. In contrast, the results obtained so far for a quantum adversary making classical queries only were less impressive. For the first time, M. Naya-Plasencia and A. Schrottenloher, together with A. Hosoyamada and Y. Sasaki, managed to leverage the algebraic structure of some cryptosystems in the context of a quantum attacker limited to classical queries and offline quantum computations. Most notably, they are able to break the Even-Mansour construction in quantum time $\widetilde{\mathcal{O}}(2n/3)$ with $\mathcal{O}(2n/3)$ classical queries and $\mathcal{O}(n^2)$ qubits only.
- Quantum cryptanalysis of CSIDH and Ordinary Isogeny-based Schemes [16]: CSIDH is a recent proposal by Castryck et al. for post-quantum non-interactive key-exchange. It is similar in design to a scheme by Couveignes, Rostovtsev and Stolbunov, but it replaces ordinary elliptic curves by supersingular elliptic curves. Although CSIDH uses supersingular curves, it can attacked by a quantum subexponential hidden shift algorithm due to Childs et al. While the designers of CSIDH

claimed that the parameters they suggested ensures security against this algorithm, X. Bonnetain and A. Schrottenloher showed that these security parameters were too optimistic: they improved the hidden shift algorithm and gave a precise complexity analysis in this context, which greatly reduced the complexity. For example, they showed that only 2^{35} quantum equivalents of a key-exchange are sufficient to break the 128-bit classical, 64-bit quantum security parameters proposed, instead of 2^{62}. They also extended their analysis to ordinary isogeny computations, and showed that an instance proposed by De Feo, Kieffer and Smith and expected to offer 56 bits of quantum security can be broken in 2^{38} quantum evaluations of a key exchange.

- New graph-related quantum algorithms. A first paper presents an approach to improve expansion testing using quantum Fast-Forwarding and growing seed sets [64]. A second paper introduces a graph sparsification algorithm [65], which when combined with existing classical algorithms yields the first quantum speedup for approximating the max cut, min cut, min st-cut, sparsest cut and balanced separator of a graph. Moreover, combining it with a classical Laplacian solver yields a similar speedup for Laplacian solving, for approximating effective resistances, cover times and eigenvalues of the Laplacian, and for spectral clustering.
- Quantum walks: in a first work, S. Apers describes a new quantum algorithm for quantum walk sampling using growing seed sets [42] with applications for st-connectivity and problems related to graph isomorphism. A second work [66] introduces a new quantum walk search framework that unifies and strengthens the existing ones.
- Quantum query lower bounds [59], [60]: Many computational problems, such as finding collisions in a function, are symmetric in their inputs. A. Chailloux showed that for this class of problems, any quantum algorithm can have at most a cubic advantage over the best classical algorithm in the query model, while the previously known bound gave up to 7th root advantage. This result enhances our understanding on the limitations of quantum algorithms.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

- **ANR DEREC** ($10/16 \rightarrow 09/21$)
  Relativistic cryptography
  ANR Program: jeunes chercheurs
  244 kEuros
  The goal of project DEREC is to demonstrate the feasibility of guaranteeing the security of some cryptographic protocols using the relativistic paradigm, which states that information propagation is limited by the speed of light. We plan to study some two party primitives such as bit commitment and their security against classical and quantum adversaries in this model. We then plan to the integration of those primitives into larger cryptosystems. Finally, we plan on performing a demonstration of those systems in real life conditions.
- **ANR CBCRYPT** ($10/17 \rightarrow 09/21$)
  Code-based cryptography
  ANR Program: AAP Générique 2017
  Partners: Inria SECRET (coordinator), XLIM, Univ. Rouen, Univ. Bordeaux.
  197 kEuros
  The goal of CBCRYPT is to propose code-based candidates to the NIST call aiming at standardizing public-key primitives which resist to quantum attacks. These proposals are based either on code-based schemes relying on the usual Hamming metric or on the rank metric. The project does not deal solely with the NIST call. We also develop some other code-based solutions: these are either

primitives that are not mature enough to be proposed in the first NIST call or whose functionalities are not covered by the NIST call, such as identity-based encryption, broadcast encryption, attribute based encryption or functional encryption. A third goal of this project is of a more fundamental nature: namely to lay firm foundations for code-based cryptography by developing thorough and rigorous security proofs together with a set of algorithmic tools for assessing the security of code-based cryptography.

- **ANR quBIC** ($10/17 \rightarrow 09/21$)
  Quantum Banknotes and Information-Theoretic Credit Cards
  ANR Program: AAP Générique 2017
  Partners: Univ. Paris-Diderot (coordinator), Inria SECRET, UPMC (LIP6), CNRS (Laboratoire Kastler Brossel)
  87 kEuros
  For a quantum-safe future, classical security systems as well as quantum protocols that guarantee security against all adversaries must be deployed. Here, we will study and implement one of the most promising quantum applications, namely unforgeable quantum money. A money scheme enables a secure transaction between a client, a vendor and a bank via the use of a credit card or via the use of banknotes, with maximal security guarantees. Our objectives are to perform a theoretical analysis of quantum money schemes, in realistic conditions and for encodings in both discrete and continuous variables, and to demonstrate experimentally these protocols using state-of-the-art quantum memories and integrated detection devices.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. QCALL

Title: Quantum Communications for ALL

Programm: H2020-MSCA-ITN-2015

Duration: December 2016 - November 2020

Coordinator: University of Leeds (UK)

Other partners: see http://www.qcall-itn.eu/

Inria contact: Anthony Leverrier

QCALL is a European Innovative Training Network that endeavors to take the next necessary steps to bring the developing quantum technologies closer to the doorsteps of end users. QCALL will empower a nucleus of 15 doctoral researchers in this area to provide secure communications in the European continent and, in the long run, to its connections worldwide.

8.2.1.2. ERC QUASYModo

Title: QUASYModo *Symmetric Cryptography in the Post-Quantum World*

Program: ERC starting grant

Duration: September 2017 - August 2022

PI: María Naya Plasencia

As years go by, the existence of quantum computers becomes more tangible and the scientific community is already anticipating the enormous consequences of the induced breakthrough in computational power. Cryptology is one of the affected disciplines. Indeed, the current state-of-the-art asymmetric cryptography would become insecure, and we are actively searching for alternatives. Symmetric cryptography, essential for enabling secure communications, seems much less affected at first sight: its biggest known threat is Grover's algorithm, which allows exhaustive key searches in the square root of the normal complexity. Thus, so far, it is believed that doubling key lengths suffices to maintain an equivalent security in the post- quantum world. The security of symmetric cryptography is completely based on cryptanalysis: we only gain confidence in the security of a symmetric primitive through extensive and continuous scrutiny. It is therefore not possible to determine whether a symmetric primitive might be secure or not in a post-quantum world without first understanding how a quantum adversary could attack it. Correctly evaluating the security of symmetric primitives in the post-quantum world cannot be done without a corresponding cryptanalysis toolbox, which neither exists nor has ever been studied. This is the big gap I have identified and that I plan to fill with this project. Next, doubling the key length is not a trivial task and needs to be carefully studied. My ultimate aim is to propose efficient solutions secure in the post-quantum world with the help of our previously obtained quantum symmetric cryptanalysis toolbox. This will help prevent the chaos that big quantum computers would generate: being ready in advance will definitely save a great amount of time and money, while protecting our current and future communications. The main challenge of QUASYModo is to redesign symmetric cryptography for the post-quantum world.

8.2.1.3. H2020 FET Flagship on Quantum Technologies - CiViQ

Title: CiViQ *Continuous Variable Quantum Communications*

Program: H2020 FET Flagship on Quantum Technologies

Duration: October 2018 - September 2021

PI: Anthony Leverrier

The goal of the CiViQ project is to open a radically novel avenue towards flexible and cost-effective integration of quantum communication technologies, and in particular Continuous-Variable QKD, into emerging optical telecom- munication networks. CiViQ aims at a broad technological impact based on a systematic analysis of telecom-defined user-requirements. To this end CiViQ unites for the first time a broad interdisciplinary community of 21 partners with unique breadth of experience, involving major telecoms, integrators and developers of QKD. The work targets advancing both the QKD technology itself and the emerging "software network" approach to lay the foundations of future seamless integration of both. CiViQ will culminate in a validation in true telecom network environment. Project-specific network integration and software development work will empower QKD to be used as a physical-layer-anchor securing critical infrastruc- tures, with demonstration in QKD-extended software-defined networks.

8.2.2. Collaborations in European Programs, Except FP7 & H2020

8.2.2.1. QCDA

Program: QuantERA ERA-NET Cofund in Quantum Technologies

Project acronym: QCDA

Project title: Quantum Code Design and Architecture

Duration: February 2018 - January 2021

Coordinator: Earl Campbell, University of Sheffield, UK

Other partners: University of Sheffield (UK), TU Delft (Netherlands), TU Munich (Germany), University College London (UK)

Inria contact: Anthony Leverrier

General purpose quantum computers must follow a fault-tolerant design to prevent ubiquitous decoherence processes from corrupting computations. All approaches to fault-tolerance demand extra physical hardware to perform a quantum computation. Kitaev's surface, or toric, code is a popular idea that has captured the hearts and minds of many hardware developers, and has given many people hope that fault-tolerant quantum computation is a realistic prospect. Major industrial hardware developers include Google, IBM, and Intel. They are all currently working toward a fault-tolerant architecture based on the surface code. Unfortunately, however, detailed resource analysis points towards substantial hardware requirements using this approach, possibly millions of qubits for commercial applications. Therefore, improvements to fault-tolerant designs are a pressing near-future issue. This is particularly crucial since sufficient time is required for hardware developers to react and adjust course accordingly.

This consortium will initiate a European co-ordinated approach to designing a new generation of codes and protocols for fault-tolerant quantum computation. The ultimate goal is the development of high-performance architectures for quantum computers that offer significant reductions in hardware requirements; hence accelerating the transition of quantum computing from academia to industry. Key directions developed to achieve these improvements include: the economies of scale offered by large blocks of logical qubits in high-rate codes; and the exploitation of continuous-variable degrees of freedom.

The project further aims to build a European community addressing these architectural issues, so that a productive feedback cycle between theory and experiment can continue beyond the lifetime of the project itself. Practical protocols and recipes resulting from this project are anticipated to become part of the standard arsenal for building scalable quantum information processors.

8.3. International Initiatives

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

8.3.1.1. CHOCOLAT

Title: Chosen-prefix Collision Attack on SHA-1 with ASICs Cluster

International Partner (Institution - Laboratory - Researcher):

NTU (Singapore) - SYLLAB - Peyrin Thomas

Start year: 2017

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

The hash function SHA-1 is one of the most widely used hash functions in the industry, but it has been shown to not be collision-resistant by a team of Chinese researchers led by Prof. Wang in 2005. However, nobody has publicly produced a real pair of colliding messages so far, because the estimated attack complexity is around 2^{63} SHA-1 computations (this represents about 70000 years of computation on a normal PC).

While a collision of SHA-1 would clearly demonstrate the weakness of the algorithm, a much more powerful attack would be to find a collision such that the prefix of the colliding messages is chosen by some challenger beforehand. In particular, this would allow creating a rogue certificate authority certificate that would be accepted by browsers. Such an attack has already been deployed for certificates using the MD5 hash function, but MD5 is much weaker than SHA-1 and it has already been removed from most security applications. SHA-1 is still widely used and performing such an attack for certificates using SHA-1 would have a very big impact.

The objective of the project is to design a chosen-prefix collision attack against the SHA-1 hash function, and to implement the attack in practice. We estimate this will require 2^{70} computations, and we will use an ASIC cluster to perform such a computation.

8.3.2. Inria International Partners

8.3.2.1. Declared Inria International Partners

Title: Discrete Mathematics, Codes and Cryptography

International Partner (Institution - Laboratory - Researcher):

Indian Statistical Institute (India) - Cryptology Research Group - Bimal Roy

Duration: 2014 - 2019

Start year: 2014

Today's cryptology offers important challenges. Some are well-known: Can we understand existing cryptanalysis techniques well enough to devise criterion for the design of efficient and secure symmetric cryptographic primitives? Can we propose cryptographic protocols which offer provable security features under some reasonable algorithmic assumptions? Some are newer: How could we overcome the possible apparition of a quantum computer with its devastating consequences on public key cryptography as it is used today? Those challenges must be addressed, and some of the answers will involve tools borrowed to discrete mathematics, combinatorics, algebraic coding theory, algorithmic. The guideline of this proposal is to explore further and enrich the already well established connections between those scientific domains and their applications to cryptography and its challenges.

8.3.2.2. Informal International Partners

- Nanyang Technological University (Singapore): cryptanalysis of symmetric primitives.
- Ruhr-Universität Bochum (Germany): design and cryptanalysis of symmetric primitives.
- NTT Secure Platforms Laboratories (Japan): quantum cryptanalysis, symmetric cryptography.
- University of Sherbrooke (Canada): quantum codes.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

- Thomas Peyrin, NTU Singapore, January 2019 and July 2019
- Mustafa Mahmoud Mohammed Kairallah, NTU Singapore, July 2019
- Léo Ducas, CWI Amsterdam, NL, March 2019
- Akinori Hosoyamada, NTT Secure Platform Labaratories, Tokyo, Japan, March 2019 and November 2019
- Yu Sasaki, NTT Secure Platform Labaratories, Tokyo, Japan, November 2019
- Gregor Leander, Ruhr Université Bochum, Germany, November 2019

8.4.1.1. Internships

- Pierre Briaud, MPRI, March-Aug. 2019
- Lucien Grouès, Telecom ParisTech, March-Sept. 2019
- Antonio Florez Gutierrez, Université Paris Saclay, March-Aug. 2019
- Sohaïb Ouzineb, Telecom ParisTech, July-Aug. 2019
- Elodie Rohart-Barbey, INSA Rouen, June-Aug. 2019
- Augustin Bariant, Ecole Polytechnique, April-Aug. 2019

8.4.2. Visits to International Teams

8.4.2.1. Research Stays Abroad

- Bar-Ilan Unversity, Israel, June 16-18, invitation by Nathan Keller (A. Canteaut and G. Leurent)
- Rostock University, Rostock, Germany, June 23-28, invitation to the Institut für Mathematik by Gohar Kyureghyan, (L. Perrin).
- NTT, Tokyo, Japan, August 27-September 27, invitation by Yu Sasaki (F. Sibleyras)

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. General Chair, Scientific Chair

- FSE 2020, March 2020, Athens, Greece: C. Boura, general chair ; G. Leurent, program co-chair
- PQCrypto 2020, April 15 - April 17, 2020, Paris, France: N. Sendrier, general co-chair ; J.P. Tillich, program co-chair
- Dagstuhl seminar on Quantum Cryptanalysis 2019, October 13-18, 2019, Dagstuhl, Germany: M. Naya-Plasencia co-organizer.

9.1.1.2. Member of the Organizing Committees

- FSE 2019: March 25-28, 2019, Paris, France: Gaëtan Leurent.

9.1.2. Scientific Events: Selection

9.1.2.1. Chair of Conference Program Committees

- Co-editor-in-chief of *IACR Transactions on Symmetric Cryptology* starting from 2019: Gaëtan Leurent.
- Co-chair of the Program Committee of *Eurocrypt 2020*, Zagreb, Croatia, May 2020: Anne Canteaut
- Co-chair of the Program Committee of *FSE 2020*, Athens, Greece, March 2020: Gaëtan Leurent.
- Co-chair of the Program Committee of *WCC 2019*, March-April 2019, St Jacut-de-la-Mer, France: A. Canteaut
- Chair of the Program Committee of *QCrypt 2019*, August 26-30 2019, Montreal, Canada: A. Leverrier

9.1.2.2. Member of the Conference Program Committees

- CT-RSA 2019: March 4-8, 2019, San Francisco, USA (L. Perrin);
- FSE 2019: March 25-28, 2019, Paris, France (C. Boura, A. Canteaut, G. Leurent, M. Naya-Plasencia)
- WCC 2019: March 31 - April 5, 2019, St Jacut-de-la-Mer, France, (A. Canteaut chair, P. Charpin, N. Sendrier, J.P. Tillich);
- PQCrypto 2019: May 8-10, 2019, Chongqing, China, (M. Naya-Plasencia, N. Sendrier, J.P. Tillich);
- CBC 2019: May 18-19, 2019, Darmstadt, Germany, (J.-P. Tillich);
- TQC 2019: June 3-9, 2019, University of Maryland, USA, (A. Chailloux);
- ISIT 2019: July 7-12, 2019, Paris, France, (A. Leverrier, J.-P. Tillich);
- CHES 2019: August 25-28, 2019, Atlanta, USA, (G. Leurent);
- PQC: Perspectives on Quantum Computing, NISQ and beyond: November 20, 2019, Paris, France (A. Leverrier);
- SPACE 2019: December 3-7, 2019, Gandhinagar, India (L. Perrin);
- FSE 2020: March 22-26, 2020, Athens, Greece, (C. Boura, A. Canteaut, G. Leurent co-chair, L. Perrin)
- PQCrypto 2020: April 15-17, 2020, Paris, France (A. Chailloux, M. Naya-Plasencia, N. Sendrier, J.P. Tillich);
- CBCrypto 2020: May 9-10, 2020, Zagreb, Croatia, (J.-P. Tillich);
- Eurocrypt 2020: May 10-14, 2020, Zagreb, Croatia, (A. Canteaut co-chair, M. Naya-Plasencia);
- ISIT 2020: June 21-26, 2020, Los Angeles, USA, (J.-P. Tillich).

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- *Advances in Mathematics of Communications*, associate editors: N. Sendrier and J.P. Tillich
- *Applicable Algebra in Engineering, Communication and Computing*, associate editor: A. Canteaut.
- *Designs, Codes and Cryptography*, associate editor: P. Charpin.
- *Finite Fields and Applications*, associate editors: A. Canteaut, P. Charpin.
- *IACR Transactions on Symmetric Cryptology*, associate editors: C. Boura, A. Canteaut, G. Leurent (co-editor in chief), M. Naya-Plasencia, L. Perrin.
- *IACR Transactions on Cryptographic Hardware and Embedded Systems*, associate editor: G. Leurent.
- *Quantum (the open journal for quantum science)*, associate editor: A. Leverrier.
- *IEEE Transactions on Information Theory*, associate editor: A. Canteaut.

9.1.4. Invited Talks

- P. Charpin, *Crooked and weakly crooked functions*, Finite Fields and Applications - Fq14, Vancouver, Canada, June 3-7, 2019.
- M. Naya-Plasencia, *Preparing Symmetric Cryptography for the Quantum World*, FSE 2019, Paris, France, 24-28 March, 2019.

9.1.5. Leadership within the Scientific Community

- A. Canteaut serves as a chair of the steering committee of *Fast Software Encryption (FSE)*, M. Naya-Plasencia and G. Leurent also serve on the committee.
- N. Sendrier serves on the steering committee of *Post-quantum cryptography (PQCrypto)*.
- P. Charpin, N. Sendrier and JP Tillich serve on the steering committee of the WCC conference series.

9.1.6. Research Administration

- A. Canteaut served as Head of Science of the Inria Paris research center until August 2019.
- A. Canteaut serves as Head of *Inria Evaluation Committee* since September 2019.
- P. Charpin serves on the *Comité Parité* at Inria.
- G. Leurent and M. Naya-Plasencia are members of *Inria Paris CSD Committee* (Comité de suivi doctoral).
- M. Naya-Plasencia serves on the *Inria Evaluation Committee* since September 2019.
- A. Leverrier serves on the steering committee of the Domaine D'Intérêt Majeur SIRTEQ (Quantum Technologies in IdF).

9.1.7. Committees for the selection of professors, assistant professors and researchers

- 2019 Jury Inria Research Positions (ARP/SRP) (M. Naya-Plasencia);
- 2019 Jury d'admission Inria DR2, (M. Naya-Plasencia);
- 2019 Jury d'admissibilité Inria DR2, (A. Canteaut);
- 2019 Jury d'admissibilité Inria CRCN national (A. Canteaut);
- Committee for the nomination to a permanent position, Radboud University Nijmegen, Netherlands (M. Naya-Plasencia).
- Committee for a professorship in cryptography, TU Graz, Austria (A. Canteaut)
- Reviewer for a promotion to professorhip, ISI Kolkata, India (A. Canteaut)
- Committee for 3 assistant professor positions in computer science at Université de Paris: A. Leverrier

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: A. Canteaut, *Error-correcting codes and applications to cryptology*, 12 hours, M2, University Paris-Diderot (MPRI), France;

Master: A. Chailloux, *Quantum information*, 12 hours, M2, University Paris-Diderot (MPRI), France;

Master: A. Chailloux, *Quantum algorithms*, 4 hours, M2, Ecole Normale Supérieure de Lyon, France;

Master: A. Leverrier, *Quantum information and quantum cryptography*, 12 hours, M2, University Paris-Diderot (MPRI), France;

Master: L. Perrin, *Application Web et Sécurité*, 24 hours, M1, UVSQ, France;

Bachelor: L. Perrin, *Cryptographie*, 29 hours, L3, UVSQ, France;

Master: J.-P. Tillich, *Introduction to Information Theory*, 36 hours, M2, Ecole Polytechnique, France;

Master: J.-P. Tillich, *Quantum Information and Applications*, 36 hours, M2, Ecole Polytechnique, France.

The members of the project-team were also invited to give courses at training schools for PhD students and young researchers:

- N. Sendrier, *Code-Based Cryptography*, PQCrypto 2019 Summer School, Chongqing, China, May 6, 2019.

9.2.2. Supervision

PhD: Xavier Bonnetain, *Cryptanalysis of symmetric primitives in the post-quantum world*, Sorbonne Université, November 15, 2019, supervisor: M. Naya Plasencia.

PhD: Thomas Debris, *Quantum algorithms for decoding linear codes*, Sorbonne Université, December 17, 2019, supervisor: JP Tillich.

PhD: Antoine Grospellier, *LDPC codes: constructions and decoding*, Sorbonne Université, November 8, 2019, supervisors: A. Leverrier and O. Fawzi (ENS Lyon)

PhD: Vivien Londe, *Study of quantum LDPC codes*, University of Bordeaux, December 6, 2019, supervisors: G. Zémor (Univ. Bordeaux) and A. Leverrier

PhD in progress: Kevin Carrier, *Presque-collisions et applications au décodage générique et à la reconnaissance de codes correcteurs d'erreurs*, since October 2016, supervisor: N. Sendrier and JP Tillich.

PhD in progress: Matthieu Lequesne, *Attaques par canaux cachés sur les cryptosystèmes à base de codes MDPC quasi-cycliques*, since September 2017, supervisor: N. Sendrier

PhD in progress: André Schrottenloher, *Long-term security of symmetric primitives*, since February 2018 supervisor: M. Naya-Plasencia

PhD in progress: Ferdinand Sibleyras, *Security of modes of operation*, since October 2017, supervisor: G. Leurent and A. Canteaut

PhD in progress: Valentin Vasseur, *Etude du décodage des codes QC-MDPC*, since October 2017, supervisor: N. Sendrier

PhD in progress: Rémi Bricout, *Etude de scénarios non-locaux quantiques à l'aide d'outils de la théorie de l'information quantique*, since September 2017, supervisor: A. Chailloux and A. Leverrier

PhD in progress: Shouvik Ghorai, *Beyond-QKD continuous-variable quantum cryptographic protocols*, since October 2017, supervisors: E. Diamanti (UPMC), A. Leverrier

PhD in progress: Andrea Olivo, *Partir de contraintes relativistes pour faire de la cryptographie quantique*, since November 2017, supervisors: A. Chailloux and F. Grosshans (laboratoire Aimé Cotton).

PhD in progress: Daniel Coggia, *Cryptanalysis techniques for lightweight ciphers*, since September 2018, supervisors: A. Canteaut and C. Boura.

PhD in progress: Simona Etinski, *Quantum algorithms and protocols*, since October 2019, supervisors: A. Chailloux, A. Leverrier and F. Magniez (Université de Paris)

PhD in progress: A. Florez Gutierrez, *Secure Symmetric Primitives and the Post-Quantum World*, since September 2019, supervisor: M. Naya Plasencia

PhD in progress: Lucien Grouès, *Decoding algorithms for quantum LDPC codes*, since October 2019, supervisors: A. Leverrier, O. Fawzi

PhD in progress: Rocco Mora, *Algebraic structures in code-based cryptography*, since October 2019, supervisor: JP Tillich

9.2.3. Juries

- Lorenzo Grassi, *Cryptanalysis of AES-like ciphers and reviving old design ideas for new constructions*, TU Graz, Austria, April 26, 2019, committe: A. Canteaut (reviewer).
- Louiza Khati, *Full Disk Encryption and Beyond*, Université Paris PSL, July 15, 2019, committee: M. Naya-Plasencia (chair).
- T. van Himbeeck, *Quantum Cryptography with Partially Trusted Devices*, Université Libre de Bruxelles, October 18, 2019, committee: A. Leverrier
- A. Aloshious, *Higher dimensional topological codes : structural properties and decoders*, Indian Institute of Technology Madras, November, 2019, committee: J.-P. Tillich (reviewer).
- A. Grospellier, *Décodage des codes expanseurs quantiques et application au calcul quantique tolérant aux fautes*, Sorbonne University, November 8, 2019, committee: A. Leverrier (supervisor), J.-P. Tillich.
- Xavier Bonnetain, *Hidden structures and quantum cryptanalysis*, Sorbonne Université, November 15, 2019, committee: A. Chailloux, M. Naya-Plasencia (Supervisor).
- Loïc Ferreira, *Secure Tunnels for Constrained Environments*, Université Bretagme Loire, November 18, 2019, committee: M. Naya-Plasencia.
- O. Fawzi, HdR, *Contributions to quantum information theory*, École Normale Supérieure de Lyon, December 2, committee: J.-P. Tillich (reviewer).
- V. Londe, *Topological Quantum Error-Correcting Codes beyond dimension 2*, Bordeaux University, December 6, 2019, committee: A. Leverrier (supervisor), JP Tillich
- M. Bozzio, *Security and implementation of advanced quantum cryptography: Quantum money - Quantum weak coin flipping*, Université Paris-Saclay, December 10, 2019, committee: A. Leverrier (reviewer)
- Alain Couvreur, HdR, *Codes algébriques et géométriques, applications à la cryptographie et à l'information quantique*; Université Paris-Diderot, December 16, 2020, committee: N. Sendrier (reviewer), J.P. Tillich
- Patrick Lacharme, HdR, *Études en Securité Informatique*, Université de Normandie, February 6, 2020, committee: M. Naya-Plasencia (reviewer).
- Thomas Debris, *Cryptographie fondée sur les codes : nouvelles approches pour constructions et preuves ; contribution en cryptanalyse*; Sorbonne Université, December 17, 2019, committee: N. Sendrier, J.P. Tillich (supervisor)

9.3. Popularization

9.3.1. Internal or external Inria responsibilities

- **Association Animath**: M. Lequesne serves on the board of Animath.
- M. Lequesne is also member of the scientific committee of the French Tournament of Young Mathematicians: redaction of the problems for the competition, jury member (chair of a jury) ; member of the scientific committee of the International Tournament of Young Mathematicians: redaction of the problems for the competition, jury member (chair of a jury) ; Member of the scientific committee of the Correspondances des Jeunes Mathématicien.ne.s: redaction of the problems for the competition.

9.3.2. Articles and contents

- *La fragilité inattendue du chiffrement symétrique dans le monde post-quantique*, Gaëtan Leurent and Maria Naya-Plasencia, María, Interstices [63]

9.3.3. Education

- **Alkindi cipher challenge:** Several members of the project-team are involved in the cipher challenge for high-school students "concours Alkindi" http://www.concours-alkindi.fr/. Mathieu Lequesne serves as a co-organizer of the challenge, preparing the three rounds and the final. Together with C. Boura, he was also involved in the redaction of the exercises.
- Organization of the event “Rendez-vous des Jeunes Mathématiciennes et Informaticiennes” at Inria Paris (October 28-29) by M. Lequesne, a 2-day camp for 20 high-school girls interested in mathematics and computer science. C. Boura, A. Canteaut and L. Perrin gave talks at this event.
- **Project Algo'scape with a high school (Lycée Anna Judic, Semur en Auxois):** design of an escape game on cryptography (L. Perrin), visit of 15 students at Inria, May 2019 (A. Canteaut and L. Perrin).

10. Bibliography

Major publications by the team in recent years

[1] C. BEIERLE, A. CANTEAUT, G. LEANDER, Y. ROTELLA.*Proving Resistance Against Invariant Attacks: How to Choose the Round Constants*, in "Crypto 2017 - Advances in Cryptology", Santa Barbara, United States, J. KATZ, H. SHACHAM (editors), LNCS - Lecture Notes in Computer Science, Springer, August 2017, vol. 10402, p. 647–678 [*DOI :* 10.1007/978-3-319-63715-0_22], https://hal.inria.fr/hal-01631130

[2] K. BHARGAVAN, G. LEURENT.*On the Practical (In-)Security of 64-bit Block Ciphers*, in "ACM CCS 2016 - 23rd ACM Conference on Computer and Communications Security", Vienna, Austria, ACM, October 2016 [*DOI :* 10.1145/2976749.2978423], https://hal.inria.fr/hal-01404208

[3] A. CANTEAUT, J. ROUÉ.*On the behaviors of affine equivalent Sboxes regarding differential and linear attacks*, in "Advances in Cryptology - Eurocrypt 2015", Sofia, Bulgaria, Lecture Notes in Computer Science, Springer, April 2015, https://hal.inria.fr/hal-01104051

[4] A. CHAILLOUX, M. NAYA-PLASENCIA, A. SCHROTTENLOHER.*An Efficient Quantum Collision Search Algorithm and Implications on Symmetric Cryptography*, in "Asiacrypt 2017 - Advances in Cryptology", Hong Kong, China, T. TAKAGI, T. PEYRIN (editors), LNCS - Lecture Notes in Computer Science, Springer, December 2017, vol. 10625, p. 211–240 [*DOI :* 10.1007/978-3-319-70697-9_8], https://hal.inria.fr/hal-01651007

[5] K. CHAKRABORTY, A. CHAILLOUX, A. LEVERRIER.*Arbitrarily long relativistic bit commitment* , in "Physical Review Letters", 2015 [*DOI :* 10.1103/PHYSREVLETT.115.250501], https://hal.inria.fr/hal-01237241

[6] P. CHARPIN, G. M. KYUREGHYAN, V. SUDER.*Sparse Permutations with Low Differential Uniformity*, in "Finite Fields and Their Applications", March 2014, vol. 28, p. 214-243 [*DOI :* 10.1016/J.FFA.2014.02.003], https://hal.archives-ouvertes.fr/hal-01068860

[7] N. COURTOIS, M. FINIASZ, N. SENDRIER.*How to achieve a McEliece-based Digital Signature Scheme*, in "Advances in Cryptology - Asiacrypt 2001", LNCS, Springer-Verlag, 2001, n° 2248, p. 157–174

[8] A. COUVREUR, A. OTMANI, J.-P. TILLICH.*Polynomial Time Attack on Wild McEliece Over Quadratic Extensions*, in "IEEE Transactions on Information Theory", January 2017, vol. 63, n° 1, p. 404–427 [*DOI :* 10.1109/TIT.2016.2574841], https://hal.inria.fr/hal-01661935

[9] M. KAPLAN, G. LEURENT, A. LEVERRIER, M. NAYA-PLASENCIA.*Breaking Symmetric Cryptosystems Using Quantum Period Finding*, in "Crypto 2016 - 36th Annual International Cryptology Conference", Santa Barbara, United States, M. ROBSHAW, J. KATZ (editors), LNCS - Lecture Notes in Computer Science, Springer, August 2016, vol. 9815, p. 207 - 237 [*DOI :* 10.1007/978-3-662-53008-5_8], https://hal.inria.fr/hal-01404196

[10] R. MISOCZKI, J.-P. TILLICH, N. SENDRIER, P. S. L. M. BARRETO.*MDPC-McEliece: New McEliece Variants from Moderate Density Parity-Check Codes*, in "IEEE International Symposium on Information Theory - ISIT 2013", Istanbul, Turkey, July 2013, p. 2069-2073, https://hal.inria.fr/hal-00870929

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] X. BONNETAIN.*Hidden Structures and Quantum Cryptanalysis*, Sorbonne Université, November 2019, https://tel.archives-ouvertes.fr/tel-02400328

[12] T. DEBRIS-ALAZARD.*Code-based Cryptography: New Approaches for Design and Proof ; Contribution to Cryptanalysis*, Sorbonne Universites, UPMC University of Paris 6, December 2019, https://hal.inria.fr/tel-02424234

[13] A. GROSPELLIER.*Constant Time Decoding of Quantum Expander Codes and Application to Fault-Tolerant Quantum Computation*, Sorbonne universités, November 2019, https://hal.inria.fr/tel-02422585

[14] V. LONDE.*Topological Quantum Error-Correcting Codes beyond dimension 2*, Inria Paris ; Université de bordeaux, December 2019, https://hal.inria.fr/tel-02429868

Articles in International Peer-Reviewed Journal

[15] Z. BAO, I. DINUR, J. GUO, G. LEURENT, L. WANG.*Generic Attacks on Hash Combiners*, in "Journal of Cryptology", 2020, 82, forthcoming [*DOI :* 10.1007/S00145-019-09328-W], https://hal.inria.fr/hal-02424905

[16] J.-F. BIASSE, X. BONNETAIN, B. PRING, A. SCHROTTENLOHER, W. YOUMANS.*A trade-off between classical and quantum circuit size for an attack against CSIDH*, in "Journal of Mathematical Cryptology", August 2019, p. 1-16, https://hal.inria.fr/hal-02423394

[17] X. BONNETAIN, M. NAYA-PLASENCIA, A. SCHROTTENLOHER.*Quantum Security Analysis of AES*, in "IACR Transactions on Symmetric Cryptology", June 2019, vol. 2019, n^o 2, p. 55-93 [*DOI :* 10.13154/TOSC.V2019.I2.55-93], https://hal.inria.fr/hal-02397049

[18] C. BOURA, A. CANTEAUT, D. COGGIA.*A General Proof Framework for Recent AES Distinguishers*, in "IACR Transactions on Symmetric Cryptology", March 2019, vol. 2019, n^o 1, p. 170-191 [*DOI :* 10.13154/TOSC.V2019.I1.170-191], https://hal.inria.fr/hal-02431695

[19] C. BOURA, A. CANTEAUT, J. JEAN, V. SUDER.*Two Notions of Differential Equivalence on Sboxes*, in "Designs, Codes and Cryptography", 2019, vol. 87, n^o 2-3, p. 185-202, forthcoming [*DOI :* 10.1007/S10623-018-0496-Z], https://hal.inria.fr/hal-01944565

[20] C. BOURA, E. CHAVLI, M. CHLOUVERAKI, K. KARVOUNIS.*The BMM symmetrising trace conjecture for groups G, G, G, G, G*, in "Journal of Symbolic Computation", 2020, vol. 96, p. 62-84, https://arxiv.org/abs/1802.07482 [*DOI :* 10.1016/J.JSC.2019.02.012], https://hal.archives-ouvertes.fr/hal-02147376

[21] A. CANTEAUT, L. PERRIN.*On CCZ-Equivalence, Extended-Affine Equivalence, and Function Twisting*, in "Finite Fields and Their Applications", March 2019, vol. 56, p. 209-246 [*DOI :* 10.1016/J.FFA.2018.11.008], https://hal.inria.fr/hal-01953353

[22] A. CANTEAUT, L. PERRIN, S. TIAN.*If a generalised butterfly is APN then it operates on 6 bits*, in "Cryptography and Communications - Discrete Structures, Boolean Functions and Sequences ", November 2019, vol. 11, n^o 6, p. 1147-1164 [*DOI :* 10.1007/S12095-019-00361-X], https://hal.inria.fr/hal-02420992

[23] K. CARRIER, J.-P. TILLICH.*Identifying an unknown code by partial Gaussian elimination*, in "Designs, Codes and Cryptography", March 2019, vol. 87, n^o 2-3, p. 685-713 [*DOI :* 10.1007/S10623-018-00593-7], https://hal.inria.fr/hal-02424098

[24] D. CHANG, N. DATTA, A. DUTTA, B. MENNINK, M. NANDI, S. SANADHYA, F. SIBLEYRAS.*Release of Unverified Plaintext: Tight Unified Model and Application to ANYDAE*, in "IACR Transactions on Symmetric Cryptology", 2019, forthcoming, https://hal.inria.fr/hal-02424957

[25] P. CHARPIN, P. LANGEVIN.*Obituary of Jacques Wolfmann (1932–2018)*, in "Designs, Codes and Cryptography", May 2019, vol. 87, n^o 5, p. 955-956 [*DOI :* 10.1007/S10623-019-00631-Y], https://hal-univ-tln.archives-ouvertes.fr/hal-02168222

[26] P. CHARPIN, J. PENG.*Differential uniformity and the associated codes of cryptographic functions*, in "Advances in Mathematics of Communications", November 2019, vol. 13, n^o 4, p. 579-600 [*DOI :* 10.3934/AMC.2019036], https://hal.inria.fr/hal-01908336

[27] P. CHARPIN, J. PENG.*New links between nonlinearity and differential uniformity*, in "Finite Fields and Their Applications", March 2019, vol. 56, p. 188-208 [*DOI :* 10.1016/J.FFA.2018.12.001], https://hal.inria.fr/hal-01907499

[28] S. GHORAI, E. DIAMANTI, A. LEVERRIER.*Composable security of two-way continuous-variable quantum key distribution without active symmetrization*, in "Physical Review A", January 2019, vol. 99, n^o 1, 11, https://arxiv.org/abs/1806.11356 [*DOI :* 10.1103/PHYSREVA.99.012311], https://hal.archives-ouvertes.fr/hal-02096575

[29] S. GHORAI, P. GRANGIER, E. DIAMANTI, A. LEVERRIER.*Asymptotic security of continuous-variable quantum key distribution with a discrete modulation*, in "Physical Review X", June 2019, vol. 9, n^{o} 2, 11, https://arxiv.org/abs/1902.01317 [*DOI :* 10.1103/PHYSREVX.9.021059], https://hal.archives-ouvertes.fr/hal-02163714

[30] A. KRISHNA, J.-P. TILLICH.*Towards Low Overhead Magic State Distillation*, in "Physical Review Letters", August 2019, vol. 123, n^{o} 7, 4 [*DOI :* 10.1103/PHYSREVLETT.123.070507], https://hal.inria.fr/hal-02424053

[31] *Best Paper*
L. PERRIN.*Partitions in the S-Box of Streebog and Kuznyechik*, in "IACR Transactions on Symmetric Cryptology", March 2019, vol. 2019, n^{o} 1, p. 302-329 [*DOI :* 10.13154/TOSC.V2019.I1.302-329], https://hal.inria.fr/hal-02396814.

Invited Conferences

[32] A. CANTEAUT.*Algebraic attacks on symmetric primitives for advanced protocols*, in "Frisiacrypt 2019 - Workshop on Symmetric Cryptography", Borkum, Germany, September 2019, https://hal.inria.fr/hal-02431723

[33] A. CANTEAUT.*Cryptanalysis – a Never-Ending Story*, in "Guest lecture for Honorary Doctorate", Bergen, Norway, October 2019, https://hal.inria.fr/hal-02431731

[34] A. CANTEAUT.*Searching for APN permutations with the butterfly construction*, in "CANADAM 2019 - minisymposium on "Finite Fields in Discrete Mathematics"", Vancouver, Canada, May 2019, https://hal.inria.fr/hal-02431757

[35] P. CHARPIN.*Crooked and weakly crooked functions*, in "Fq14 - 14th international conference on Finite Fields and Applications -", Vancouver, Canada, June 2019, https://hal.inria.fr/hal-02431744

[36] A. LEVERRIER, V. LONDE, G. ZÉMOR.*Quantum local testability*, in "Symmetry, Phases of Matter, and Resources in Quantum Computing", Waterloo, Canada, November 2019, https://hal.inria.fr/hal-02432364

[37] M. NAYA-PLASENCIA.*Preparing Symmetric Crypto for the Quantum World*, in "FSE 2019 - 26th Annual Fast Software Encryption Conference", Paris, France, March 2019, https://hal.inria.fr/hal-02424409

[38] M. NAYA-PLASENCIA, A. SCHROTTENLOHER, A. CHAILLOUX, L. GRASSI.*New algorithms for quantum (symmetric) cryptanalysis*, in "QuAC: Quantum Algorithms for Cryptanalysis", Darmstadt, Germany, May 2019, https://hal.inria.fr/hal-02423376

[39] M. NAYA-PLASENCIA, A. SCHROTTENLOHER, A. CHAILLOUX, L. GRASSI.*Quantum Merging Algorithms*, in "Dagstuhl Seminar 19421 Quantum Cryptanalysis", Dagstuhl, Germany, October 2019, https://hal.inria.fr/hal-02423380

[40] F. SIBLEYRAS.*The Missing Difference Problem, and its Applications to Counter Mode Encryption*, in "Invited talk at NTT Secure Platform Laboratories", Musashino, Japan, September 2019, https://hal.inria.fr/hal-02424996

[41] F. SIBLEYRAS.*Low-Memory Attacks Against Two-Round Even-Mansour Using the 3-XOR Problem*, in "Invited talk at NTT Secure Platform Laboratories", Musashino, Japan, September 2020, https://hal.inria.fr/hal-02425000

International Conferences with Proceedings

[42] S. APERS.*Quantum Walk Sampling by Growing Seed Sets*, in "ESA 2019 - 27th Annual European Symposium on Algorithms", Munich/Garching, Germany, September 2019, https://arxiv.org/abs/1904.11446 [*DOI :* 10.4230/LIPIcs.ESA.2019.9], https://hal.inria.fr/hal-02436629

[43] X. BONNETAIN, A. HOSOYAMADA, M. NAYA-PLASENCIA, Y. SASAKI, A. SCHROTTENLOHER.*Quantum Attacks without Superposition Queries: the Offline Simon's Algorithm*, in "ASIACRYPT 2019 - 25th Annual International Conference on the Theory and Application of Cryptology and Information Security", Kobe, Japan, LNCS, Springer, December 2019, vol. 11921, p. 552-583 [*DOI :* 10.1007/978-3-030-34578-5_20], https://hal.inria.fr/hal-02397056

[44] X. BONNETAIN, L. PERRIN, S. TIAN.*Anomalies and Vector Space Search: Tools for S-Box Analysis*, in "ASIACRYPT 2019 - 25th International Conference on the Theory and Application of Cryptology and Information Security", Kobe, Japan, LNCS, Springer, November 2019, vol. 11921, p. 196-223 [*DOI :* 10.1007/978-3-030-34578-5_8], https://hal.inria.fr/hal-02396738

[45] R. BRICOUT, A. CHAILLOUX, T. DEBRIS-ALAZARD, M. LEQUESNE.*Ternary Syndrome Decoding with Large Weight*, in "SAC 2019 - 26th International Conference Selected Areas in Cryptography", Waterloo, Canada, August 2019, https://hal.inria.fr/hal-02420997

[46] A. CANTEAUT, V. LALLEMAND, G. LEANDER, P. NEUMANN, F. WIEMER.*Bison: Instantiating the Whitened Swap-Or-Not Construction*, in "Eurocrypt 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques", Darmstadt, Germany, LNCS, Springer, May 2019, vol. 11478 [*DOI :* 10.1007/978-3-030-17659-4_20], https://hal.inria.fr/hal-02431714

[47] R. CANTO-TORRES, J.-P. TILLICH.*Speeding up decoding a code with a non-trivial automorphism group up to an exponential factor*, in "ISIT 2019 - IEEE International Symposium on Information Theory", Paris, France, IEEE, July 2019, p. 1927-1931 [*DOI :* 10.1109/ISIT.2019.8849628], https://hal.inria.fr/hal-02424101

[48] A. COUVREUR, M. LEQUESNE, J.-P. TILLICH.*Recovering short secret keys of RLCE encryption scheme in polynomial time*, in "PQCrypto 2019 - International Conference on Post-Quantum Cryptography", Chongqing, China, May 2019, https://arxiv.org/abs/1805.11489 [*DOI :* 10.1007/978-3-030-25510-7_8], https://hal.inria.fr/hal-01959617

[49] *Best Paper*
T. DEBRIS-ALAZARD, N. SENDRIER, J.-P. TILLICH.*Wave: A New Family of Trapdoor One-Way Preimage Sampleable Functions Based on Codes*, in "ASIACRYPT 2019 - 25th International Conference on the Theory and Application of Cryptology and Information Security", Kobe, Japan, LNCS, Springer, November 2019, vol. 11921, p. 21-51 [*DOI :* 10.1007/978-3-030-34578-5_2], https://hal.inria.fr/hal-02424057.

[50] O. DUNKELMAN, L. PERRIN.*Adapting Rigidity to Symmetric Cryptography: Towards "Unswerving" Designs*, in "SSR 2019 - Proceedings of the 5th ACM Workshop on Security Standardisation Research Workshop",

London, Royaume-Uni, ACM Press, November 2019, p. 69-80 [*DOI :* 10.1145/3338500.3360335], https://hal.inria.fr/hal-02396695

[51] S. DUVAL, G. LEURENT.*Lightweight MACs from Universal Hash Functions*, in "CARDIS 2019 - 18th Smart Card Research and Advanced Application Conference", Prague, Czech Republic, November 2019, https://hal.inria.fr/hal-02424904

[52] G. LEURENT, T. PEYRIN.*From Collisions to Chosen-Prefix Collisions : Application to Full SHA-1*, in "Eurocrypt 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques", Darmstadt, Germany, LNCS, Springer, April 2019, vol. 11478, p. 527-555 [*DOI :* 10.1007/978-3-030-17659-4_18], https://hal.inria.fr/hal-02424900

[53] G. LEURENT, F. SIBLEYRAS.*Low-Memory Attacks against Two-Round Even-Mansour using the 3-XOR Problem*, in "CRYPTO 2019 - 39th Annual International Cryptology Conference", Santa Barbara, United States, A. BOLDYREVA, D. MICCIANCIO (editors), LNCS, Springer, August 2019, vol. 11693, p. 210-235 [*DOI :* 10.1007/978-3-030-26951-7_8], https://hal.inria.fr/hal-02424902

[54] Y. LI, G. LEURENT, M. WANG, W. WANG, G. ZHANG, Y. LIU.*Universal Forgery Attack against GCM-RUP*, in "CT-RSA 2020 - The Cryptographers' Track at the RSA Conference 2020", San Francisco, United States, February 2020, https://hal.inria.fr/hal-02424899

[55] F. SIBLEYRAS.*Generic Attack on Iterated Tweakable FX Constructions*, in "CT-RSA 2020 - The Cryptographers' Track at the RSA Conference 2020", San Francisco, United States, February 2020, https://hal.inria.fr/hal-02424953

National Conferences with Proceeding

[56] L. PERRIN, X. BONNETAIN.*Russian Style (Lack of) Randomness*, in "Symposium sur la sécurité des technologies de l'information et des communications", Rennes, France, June 2019, https://hal.inria.fr/hal-02396792

Conferences without Proceedings

[57] C. BOURA, L. PERRIN, S. TIAN.*Boomerang Uniformity of Popular S-box Constructions*, in "WCC 2019 - The Eleventh International Workshop on Coding and Cryptography", Saint-Jacut-de-la-Mer, France, March 2019, https://hal.inria.fr/hal-02420970

[58] A. CANTEAUT, L. PERRIN.*On CCZ-Equivalence, Extended-Affine Equivalence, and Function Twisting*, in "Fq14 - 14th international conference on Finite Fields and Applications", Vancouver, Canada, June 2019, https://hal.inria.fr/hal-02431739

[59] A. CHAILLOUX.*A note on the Quantum Query Complexity of Permutation Symmetric Functions*, in "QIP 2019 - 22nd Annual Conference on Quantum Information Processing", Boulder, United States, January 2019, https://hal.inria.fr/hal-02427235

[60] A. CHAILLOUX.*A note on the quantum query complexity of permutation symmetric functions*, in "ITCS 2019 - 10th Annual Innovations in Theoretical Computer Science", San Diego, United States, 2019, https://arxiv.org/abs/1810.01790 [*DOI :* 10.4230/LIPIcs.ITCS.2019.19], https://hal.inria.fr/hal-01950650

[61] D. COGGIA, A. COUVREUR.*On the security of a Loidreau's rank metric code based encryption scheme*, in "WCC 2019 - The Eleventh International Workshop on Coding and Cryptography", Saint Jacut de la mer, France, March 2019, https://hal.archives-ouvertes.fr/hal-02064465

[62] A. GROSPELLIER, L. GROUÈS, A. KRISHNA, A. LEVERRIER.*Combining Hard and Soft Decoders for Hypergraph Product Codes*, in "QEC19 - 5th International Conference on Quantum Error Correction", London, United Kingdom, July 2019, https://hal.inria.fr/hal-02429542

Scientific Popularization

[63] G. LEURENT, M. NAYA-PLASENCIA.*La fragilité inattendue du chiffrement symétrique dans le monde post-quantique*, in "Interstices", December 2019, https://hal.inria.fr/hal-02425716

Other Publications

[64] S. APERS.*Expansion Testing using Quantum Fast-Forwarding and Seed Sets*, January 2020, https://arxiv.org/abs/1907.02369 - 17 pages, 2 figures; v2: fixed error in Lemma 3, with corresponding modifications, https://hal.inria.fr/hal-02436647

[65] S. APERS, R. DE WOLF.*Quantum Speedup for Graph Sparsification, Cut Approximation and Laplacian Solving*, January 2020, https://arxiv.org/abs/1911.07306 - working paper or preprint, https://hal.inria.fr/hal-02436651

[66] S. APERS, A. GILYÉN, S. JEFFERY.*A Unified Framework of Quantum Walk Search*, January 2020, https://arxiv.org/abs/1912.04233 - working paper or preprint, https://hal.inria.fr/hal-02436653

[67] M. BARDET, P. BRIAUD, M. BROS, P. GABORIT, V. NEIGER, O. RUATTA, J.-P. TILLICH.*An Algebraic Attack on Rank Metric Code-Based Cryptosystems*, October 2019, working paper or preprint, https://hal-unilim.archives-ouvertes.fr/hal-02303015

[68] I. BARDET, A. CAPEL, A. LUCIA, D. PÉREZ-GARCÍA, C. ROUZÉ.*On the modified logarithmic Sobolev inequality for the heat-bath dynamics for 1D systems*, January 2020, https://arxiv.org/abs/1908.09004 - 26 pages, 4 figures, https://hal.archives-ouvertes.fr/hal-02436766

[69] I. BARDET, M. JUNGE, N. LARACUENTE, C. ROUZÉ, D. S. FRANÇA.*Group transference techniques for the estimation of the decoherence times and capacities of quantum Markov semigroups*, January 2020, https://arxiv.org/abs/1904.11043 - 39 pages, 2 figures, https://hal.archives-ouvertes.fr/hal-02436767

[70] A. BARIANT.*Cryptanalysis of Tweakable Block Ciphers and Forkciphers*, École Polytechnique, July 2019, INTERNSHIP REPORT, https://hal.inria.fr/hal-02426441

[71] X. BONNETAIN.*Collisions on Feistel-MiMC and univariate GMiMC*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02400343

[72] X. BONNETAIN.*Improved Low-qubit Hidden Shift Algorithms*, December 2019, https://arxiv.org/abs/1901.11428 - working paper or preprint, https://hal.inria.fr/hal-02400414

[73] R. BRICOUT, A. CHAILLOUX, T. DEBRIS-ALAZARD, M. LEQUESNE.*Ternary Syndrome Decoding with Large Weight*, July 2019, Munich Workshop on Coding and Cryptography (MWCC), Poster, https://hal.inria.fr/hal-02421017

[74] A. CANTEAUT, S. DUVAL, G. LEURENT, M. NAYA-PLASENCIA, L. PERRIN, T. PORNIN, A. SCHROTTENLOHER.*Saturnin: a suite of lightweight symmetric algorithms for post-quantum security*, March 2019, Soumission à la compétition "Lightweight Cryptography" du NIST, https://hal.inria.fr/hal-02436763

[75] A. CHAILLOUX.*DEREC - Développement de la cryptographie relativiste*, October 2019, WISG 2019 - 13ème Workshop Interdisciplinaire sur la Sécurité Globale, Poster, https://hal.inria.fr/hal-02427236

[76] A. CHAILLOUX.*Quantum security of the Fiat-Shamir transform of commit and open protocols*, 2019, https://arxiv.org/abs/1906.05415 - working paper or preprint, https://hal.inria.fr/hal-02427223

[77] P. CHARPIN.*Crooked functions*, October 2019, working paper or preprint, https://hal.inria.fr/hal-02337711

[78] N. DAVID.*Quantum impossible differential attack. Applications to CLEFIA, AES and SKINNY*, MPRI, September 2019, https://hal.inria.fr/hal-02424410

[79] T. DEBRIS-ALAZARD, N. SENDRIER, J.-P. TILLICH.*About Wave Implementation and its Leakage Immunity*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02424231

[80] A. FLOREZ GUTIERREZ.*Improving the key recovery in Linear Cryptanalysis: An application to PRESENT*, UVSQ, September 2019, https://hal.inria.fr/hal-02424413

[81] A. KRISHNA, J.-P. TILLICH.*Magic state distillation with punctured polar codes*, May 2019, https://arxiv.org/abs/1811.03112 - working paper or preprint, https://hal.inria.fr/hal-02120563

[82] A. KRISHNA, J.-P. TILLICH.*Towards low overhead magic state distillation*, May 2019, https://arxiv.org/abs/1811.08461 - working paper or preprint, https://hal.inria.fr/hal-02120564

[83] A. LEVERRIER, V. LONDE, G. ZÉMOR.*Towards local testability for quantum coding*, January 2020, https://arxiv.org/abs/1911.03069 - 38 pages, https://hal.inria.fr/hal-02432360

[84] S. OUZINEB.*Towards the Reverse-Engineering of the CaveTable*, Télécom ParisTech, August 2019, https://hal.inria.fr/hal-02275389

[85] L. PERRIN.*Russian Style (Lack of) Randomness*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02396756

[86] L. PERRIN.*Streebog and Kuznyechik: Inconsistencies in the Claims of their Designers*, July 2019, IETF 105, https://hal.inria.fr/hal-02396671

Project-Team SERENA

Simulation for the Environment: Reliable and Efficient Numerical Algorithms

IN COLLABORATION WITH: Centre d'Enseignement et de Recherche en Mathématiques et Calcul Scientifique (CERMICS)

IN PARTNERSHIP WITH:
Ecole des Ponts ParisTech

RESEARCH CENTER
Paris

THEME
Earth, Environmental and Energy Sciences

Table of contents

Project-Team SERENA

Creation of the Team: 2015 June 01, updated into Project-Team: 2017 April 01

Keywords:

Computer Science and Digital Science:

A2.1.3. - Object-oriented programming
A2.1.4. - Functional programming
A2.4.3. - Proofs
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.4. - Multiscale modeling
A6.1.5. - Multiphysics modeling
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.5. - Numerical Linear Algebra
A6.2.8. - Computational geometry and meshes
A6.3.1. - Inverse problems
A6.3.4. - Model reduction
A6.3.5. - Uncertainty Quantification

Other Research Topics and Application Domains:

B3.1. - Sustainable development
B3.3.1. - Earth and subsoil
B3.4.2. - Industrial risks and waste
B3.4.3. - Pollution
B4.1. - Fossile energy production (oil, gas)
B4.2.1. - Fission
B5.5. - Materials

1. Team, Visitors, External Collaborators

Research Scientists

Martin Vohralík [Team leader, Inria, Senior Researcher, HDR]
François Clément [Inria, Researcher]
Alexandre Ern [Ecole Nationale des Ponts et Chaussées, Researcher, HDR]
Jean-Luc Guermond [Inria International Chair, HDR]
Michel Kern [Inria, Researcher]
Jan Papež [Inria, Starting Research Position, from Jul 2019]
Géraldine Pichot [Inria, Researcher]
Pierre Weis [Inria, Senior Researcher, until May 2019]

External Collaborators

Hend Ben Ameur [IPEST and ENIT-Lamsin (Tunisia), Professor, HDR]
Erik Burman [University College London, Professor, HDR]
Guy Chavent [Univ Paris-Dauphine, Professor (retired), HDR]
Jad Dabaghi [CEA, post-doc, from Sep 2019]
Jérôme Jaffré [Inria, Senior Researcher (retired), HDR]
Caroline Japhet [Univ Paris-Nord, Associate Professor]

Vincent Martin [Univ de technologie de Compiègne, Associate Professor]
Jean-Elizabeth Roberts [Inria, Senior Researcher (retired), HDR]
Pierre Weis [Inria, Senior Researcher (retired), from Jun 2019]

Technical Staff
Sébastien Furic [Inria, Engineer, until Jul 2019]
Florent Hédin [Inria, Engineer]

PhD Students
Jad Dabaghi [Inria, until Jan 2019]
Joëlle Ferzly [IFPEN, from Oct 2019]
Nicolas Pignet [EDF, until Oct 2019]
Frédéric Marazzato [CEA]
Riccardo Milani [EDF]
Idrissa Niakh [EDF, from Nov 2019]
Ani Miraci [Inria]
Stefano Piccardo [Ecole Nationale des Ponts et Chaussées, from Oct 2019]

Post-Doctoral Fellows
Guillaume Delay [Ecole Nationale des Ponts et Chaussées, until Sep 2019]
Omar Duran [Ecole Nationale des Ponts et Chaussées, from Nov 2019]
Kenan Kergrene Profit [Inria]
Seyed Mohammad Zakerzadeh [Inria, until Jul 2019]

Visiting Scientists
Roland Becker [University of Pau, Professor, Apr 2019]
Théophile Chaumont-Frelet [Inria Sophia Antipolis, Researcher, Feb, Apr, Jun, Sep & Dec 2019]
Quaanling Deng [Curtin University, PostDoc, from May 2019 until Jun 2019]
Gregor Gantner [Vienna University of Technology, PostDoc, Feb & Sep 2019]
Thirupathi Gudi [Indian Institute of Science, Bangalore, Professor, Jun 2019]
Dirk Praetorius [Vienna University of Technology, Professor, Mar 2019]
Sergio Rojas [Curtin University, PostDoc, Sep 2019]
Ivan Yotov [University of Pittsburgh, Inria invited Professor, Sep – Dec 2019]

Administrative Assistant
Derya Gök [Inria]

2. Overall Objectives

2.1. Overall Objectives

The project-team SERENA is concerned with **numerical methods** for **environmental problems**. The main topics are the conception and analysis of *models* based on *partial differential equations*, the study of their *precise and efficient numerical approximation*, and implementation issues with special concern for *reliability and correctness of programs*. We are in particular interested in *guaranteeing* the *quality* of the *overall simulation process*. SERENA has taken over the project-team POMDAPI2 which ended on May 31, 2015. It has been given an authorization to become a joint project-team between Inria and ENPC at the Committee of Projects, September 1st, 2016, and was created as project-team on April 10, 2017.

3. Research Program

3.1. Multiphysics coupling

Within our project, we start from the conception and analysis of *models* based on *partial differential equations* (PDEs). Already at the PDE level, we address the question of *coupling* of different models; examples are that of simultaneous fluid flow in a discrete network of two-dimensional *fractures* and in the surrounding three-dimensional porous medium, or that of interaction of a compressible flow with the surrounding elastic *deformable structure*. The key physical characteristics need to be captured, whereas existence, uniqueness, and continuous dependence on the data are minimal analytic requirements that we seek to satisfy. At the modeling stage, we also develop model-order reduction techniques, such as the use of reduced basis techniques or proper generalized decompositions, to tackle evolutive problems, in particular in the nonlinear case, and we are also interested in developing reduced-order methods for variational inequalities such as those encountered in solid mechanics with contact and possibly also friction.

3.2. Discretization by hybrid high-order and discrete element methods

We consequently design *numerical methods* for the devised model. Traditionally, we have worked in the context of finite element, finite volume, mixed finite element, and discontinuous Galerkin methods. Novel classes of schemes enable the use of general *polygonal* and *polyhedral meshes* with *nonmatching interfaces*, and we develop them in response to a high demand from our industrial partners (namely EDF, CEA, and IFP Energies Nouvelles). In the lowest-order case, our focus is to design *discrete element* methods for solid mechanics. The novelty is to devise these methods to treat dynamic elastoplasticity as well as quasi-static and dynamic crack propagation. We also develop *structure-preserving* methods for the Navier–Stokes equations, i.e., methods that mimic algebraically at the discrete level fundamental properties of the underlying PDEs, such as conservation principles and preservation of invariants. In the higher-order case, we actively contribute to the development of *hybrid high-order* methods. We contribute to the numerical analysis in nonlinear cases (obstacle problem, Signorini conditions), we apply these methods to challenging problems from solid mechanics involving large deformations and plasticity, and we develop a comprehensive software implementing them. We believe that these methods belong to the future generation of numerical methods for industrial simulations; as a concrete example, the implementation of these methods in an industrial software of EDF has been completed in 2019 in the framework of the PhD thesis of Nicolas Pignet.

3.3. Domain decomposition and Newton–Krylov (multigrid) solvers

We next concentrate an intensive effort on the development and analysis of efficient solvers for the systems of nonlinear algebraic equations that result from the above discretizations. We have in the past developed *Newton–Krylov solvers* like the *adaptive inexact Newton method*, and we place a particular emphasis on *parallelization* achieved via the *domain decomposition* method. Here we traditionally specialize in *Robin transmission conditions*, where an optimized choice of the parameter has already shown speed-ups in orders of magnitude in terms of the number of domain decomposition iterations in model cases. We concentrate in the SERENA project on adaptation of these algorithms to the above novel discretization schemes, on the optimization of the free Robin parameter for challenging situations, and also on the use of the Ventcell transmission conditions. Another feature is the use of such algorithms in time-dependent problems in *space-time* domain decomposition that we have recently pioneered. This allows the use of different time steps in different parts of the computational domain and turns out to be particularly useful in porous media applications, where the amount of diffusion (permeability) varies abruptly, so that the evolution speed varies significantly from one part of the computational domain to another. Our new theme here are *Newton–multigrid solvers*, where the geometric multigrid solver is *tailored* to the specific problem under consideration and to the specific numerical method, with problem- and discretization-dependent restriction, prolongation, and smoothing. Using patchwise smoothing, we have in particular recently developed a first multigrid method whose behavior is both in theory and in practice insensitive of (robust with respect to) the approximation polynomial degree.

With patchwise techniques, we also achieve mass balance at each iteration step, a highly demanded feature in most of the target applications. The solver itself is then *adaptively steered* at each execution step by an a posteriori error estimate (adaptive stepsize, adaptive smoothing).

3.4. Reliability by a posteriori error control

The fourth part of our theoretical efforts goes towards guaranteeing the results obtained at the end of the numerical simulation. Here a key ingredient is the development of rigorous *a posteriori estimates* that make it possible to estimate in a fully computable way the error between the unknown exact solution and its numerical approximation. Our estimates also allow to distinguish the different *components* of the overall *error*, namely the errors coming from modeling, from the discretization scheme, from the nonlinear (Newton) solver, and from the linear algebraic (Krylov, domain decomposition, multigrid) solver. A new concept here is that of *local stopping criteria*, where all the error components are balanced locally within each computational mesh element. This naturally connects all parts of the numerical simulation process and gives rise to novel *fully adaptive algorithms*. We also theoretically address the question of convergence of the new fully adaptive algorithms. We identify theoretical conditions so that the error diminishes at each adaptive loop iteration by a contraction factor and we in particular derive a guaranteed error reduction factor in model cases. We have also proved a numerical optimality of the derived algorithms in model cases in the sense that, up to a generic constant, the smallest possible computational effort to achieve the given accuracy is needed.

3.5. Safe and correct programming

Finally, we concentrate on the issue of computer implementation of scientific computing programs. Increasing complexity of algorithms for modern scientific computing makes it a major challenge to implement them in the traditional imperative languages popular in the community. As an alternative, the computer science community provides theoretically sound tools for *safe* and *correct programming*. We explore here the use of these tools to design generic solutions for the implementation of the class of scientific computing software that we deal with. Our focus ranges from high-level programming via *functional programming* with OCAML through safe and easy parallelism via *skeleton parallel programming* with SKLML to proofs of correctness of numerical algorithms and programs via *mechanical proofs* with COQ.

4. Application Domains

4.1. Multiphase flows and transport of contaminants in the subsurface

- subsurface depollution after chemical leakage
- nuclear waste disposal in deep underground repositories
- flow in large scale discrete fracture networks
- production of oil and gas

4.2. Industrial risks in energy production

- Stokes and Navier–Stokes flows related to nuclear reactor operation
- seismic wave propagation for detection and protection
- electromagnetism for interfaces between dielectrics and negative metamaterials

4.3. Nonlinear mechanics

- quasi-static and dynamic elastoplastic evolutions with small and large deformations
- quasi-static and dynamic crack propagation
- nonlinear contact and friction conditions

- application to engineering components mainly related to nuclear reactor operation and safety analysis

4.4. Computational quantum chemistry

- guaranteed bounds for ground-state energy (eigenvalues) and ground-state density matrix (eigenvectors) in first-principle molecular simulation
- application to Laplace, Gross–Pitaevskii, Kohn–Sham, and Schrödinger models

5. Highlights of the Year

5.1. Highlights of the Year

Many new results of the ERC GATIPOR project in the ERC GATIPOR Gallery.

6. New Software and Platforms

6.1. CELIA3D

KEYWORDS: Fluid mechanics - Multi-physics simulation

FUNCTIONAL DESCRIPTION: The CELIA3D code simulates the coupling between a compressible fluid flow and a deformable structure. The fluid is handled by a Finite Volume method on a structured Cartesian grid. The solid is handled by a Discrete Element method (Mka3d scheme). The solid overlaps the fluid grid and the coupling is carried out with immersed boundaries (cut cells) in a conservative way.

- Partners: Ecole des Ponts ParisTech - CEA
- Contact: Laurent Monasse
- URL: http://cermics.enpc.fr/~monassel/CELIA3D/

6.2. DiSk++

Discontinuous Skeletal C++ Library

KEYWORDS: High order methods - Polyhedral meshes - C++

SCIENTIFIC DESCRIPTION: Discontinuous Skeletal methods approximate the solution of boundary-value problems by attaching discrete unknowns to mesh faces (hence the term skeletal) while allowing these discrete unknowns to be chosen independently on each mesh face (hence the term discontinuous). Cell-based unknowns, which can be eliminated locally by a Schur complement technique (also known as static condensation), are also used in the formulation. Salient examples of high-order Discontinuous Skeletal methods are Hybridizable Discontinuous Galerkin methods and the recently-devised Hybrid High-Order methods. Some major benefits of Discontinuous Skeletal methods are that their construction is dimension-independent and that they offer the possibility to use general meshes with polytopal cells and non-matching interfaces. The mathematical flexibility of Discontinuous Skeletal methods can be efficiently replicated in a numerical software: by using generic programming, the DiSk++ library offers an environment to allow a programmer to code mathematical problems in a way completely decoupled from the mesh dimension and the cell shape.

FUNCTIONAL DESCRIPTION: The software provides a numerical core to discretize partial differential equations arising from the engineering sciences (mechanical, thermal, diffusion). The discretization is based on the "Hybrid high-order" or "Discontinuous Skeletal" methods, which use as principal unknowns polynomials of arbitrary degree on each face of the mesh. An important feature of these methods is that they make it possible to treat general meshes composed of polyhedral cells. The DiSk ++ library, using generic programming techniques, makes it possible to write a code for a mathematical problem independently of the mesh. When a user writes the code for his problem using the basic operations offered by DiSk ++, that code can be executed without modifications on all types of mesh already supported by the library and those that will be added in the future.

- Author: Matteo Cicuttin
- Partner: CERMICS
- Contact: Matteo Cicuttin
- Publication: Implementation of Discontinuous Skeletal methods on arbitrary-dimensional, polytopal meshes using generic programming
- URL: https://github.com/wareHHOuse/diskpp

6.3. NEF-Draw

Numerical Experiments involving Fractures -Visualisation

KEYWORD: Fracture network

SCIENTIFIC DESCRIPTION: This version includes Matlab vectorization of the operations which makes it possible to load flow solution on meshes with more than one million fractures. It includes a text menu allowing the user to choose between different visualisation options (geometry, mesh together with the aspect ratio or together with the flow solution). A selective visualisation of fractures is also possible, by loading only the fractures that carry most of the flow.

FUNCTIONAL DESCRIPTION: This software is a visualization tool of discrete fractured networks. It allows the visualization of the network geometry, the mesh of the network together with several quantities of interest (mesh quality, flow solution including wells) computed with the software NEF-Flow.

NEWS OF THE YEAR: Add vizualization of new quantities (a posteriori estimators, velocity fields,...). New features for user-friendly usage of NEF-Draw.

- Participant: Géraldine Pichot
- Contact: Géraldine Pichot
- URL: https://gitlab.inria.fr/gpichot/NEF

6.4. NEF-Flow

KEYWORDS: Hydrogeology - Numerical simulations - 3D

SCIENTIFIC DESCRIPTION: NEF-Flow is a Matlab software for the simulation of steady state single phase flow in Discrete Fracture Networks (DFNs) using the Mixed Hybrid Finite Element (MHFEM) method for conforming and non conforming discretizations. It includes: wells, sink/source terms boundary conditions, implementation of RT0 and P1 non conforming finite elements, data structures to save the information local to each fracture, a set of non regression tests to check the solution on the Inria continuous integration platform, wells and sink/sources boundary conditions in the function that check the solution.

FUNCTIONAL DESCRIPTION: The software NEF-Flow solves the problem of an incompressible fluid flowing through a network of fractures. The software is interfaced with different mesh generators, among which BLSURF from the GAMMA3 team. A mixed hybrid finite element method is implemented.

NEWS OF THE YEAR: The last version includes new features (new APP v1.1): a full documentation of the code, new 2D test cases with analytical solutions to study the convergence of the non-conforming (Mortar) method, new organization of the non regression tests and automatic run on the Inria continuous integration platform, a refactoring of the solver parts, new developments of an iterative solver to solve flow in fractures networks (research work with Pr. Ludmil Zikatanov, PennState University), the computation of the velocity fields with P1 non conforming FE.

- Participants: Géraldine Pichot, Jean-Raynald De Dreuzy and Jocelyne Erhel
- Contact: Géraldine Pichot
- Publication: A mixed hybrid Mortar method for solving flow in discrete fracture networks
- URL: https://gitlab.inria.fr/gpichot/NEF

6.5. NEF++

KEYWORDS: Fracture network - Finite element modelling - High order methods

SCIENTIFIC DESCRIPTION: NEF++ is able to solve in a very efficient way flow in large scale fractures networks using the HHO method.

FUNCTIONAL DESCRIPTION: The software NEF++ allows to solve flow problems in fractured rocks. It is based on a hybrid high order (HHO) method. It is based on the C++17 standard. It relies on the Eigen library, which is a C++ template library for linear algebra and on the DiSk++ library for HHO https://bil.inria.fr/fr/software/view/3143/tab. The linear systems can be solved with direct solvers or with iterative solvers like the preconditioned conjugate gradient or multigrid solvers.

RELEASE FUNCTIONAL DESCRIPTION: Work in progress, no production release yet.

NEWS OF THE YEAR: NEF++ has been successfully tested on networks that contain up to 500,000 fractures. The following two direct solvers can be called: Pardiso from Intel MKL library or SuiteSparse. Both solvers support tasks parallelism, either using OpenMP or Intel TBB and support SIMD (Single Instruction, Multiple Data) vectorization. The input parameters are JSON files. The mesh data are in .vector files (see Appendix A, publication hal-00735675). The outputs are written in HDF5 files.

- Participants: Florent Hedin, Géraldine Pichot, Alexandre Ern and Nicolas Pignet
- Contact: Géraldine Pichot
- Publication: A hybrid high-order method for flow simulations in discrete fracture networks.
- URL: https://gitlab.inria.fr/nef/NEFpp

6.6. NEF-Flow-a-posteriori

KEYWORD: A posteriori error estimates

SCIENTIFIC DESCRIPTION: The a posteriori error estimates that are implemented are the ones described in the publication https://hal.archives-ouvertes.fr/hal-00019800/document.

FUNCTIONAL DESCRIPTION: Compute a posteriori estimators for diffusion problems.

NEWS OF THE YEAR: New software: implementation of the estimators, management of input files with Json.

- Participants: Géraldine Pichot and Martin Vohralík
- Contact: Géraldine Pichot
- Publication: A posteriori error estimates for lowest-order mixed finite element discretizations of convection-diffusion-reaction equations
- URL: https://gitlab.inria.fr/gpichot/nef-flow-a-posteriori

6.7. SBM

Skew Brownian Motion

KEYWORDS: Monte-Carlo methods - Skew Brownian Motion

FUNCTIONAL DESCRIPTION: SBM is a code allowing exact or approximated simulations of the Skew Brownian Motion. This code is used for the simulation, with a Monte-Carlo approach, of a 1D diffusion process with a discontinuous diffusion coefficient. Several benchmark tests are also implemented.

NEWS OF THE YEAR: - Refactoring and Cmake compilation - Automatic non regression tests on ci-inria.fr - Full documentation - Open source project on gitlab-inria

- Authors: Antoine Lejay and Géraldine Pichot
- Contact: Antoine Lejay
- Publication: Simulating diffusion processes in discontinuous media: Benchmark tests
- URL: https://gitlab.inria.fr/lejay/sbm

7. New Results

7.1. Hybrid high-order methods for nonlinear mechanics

Participants: Alexandre Ern, Nicolas Pignet.

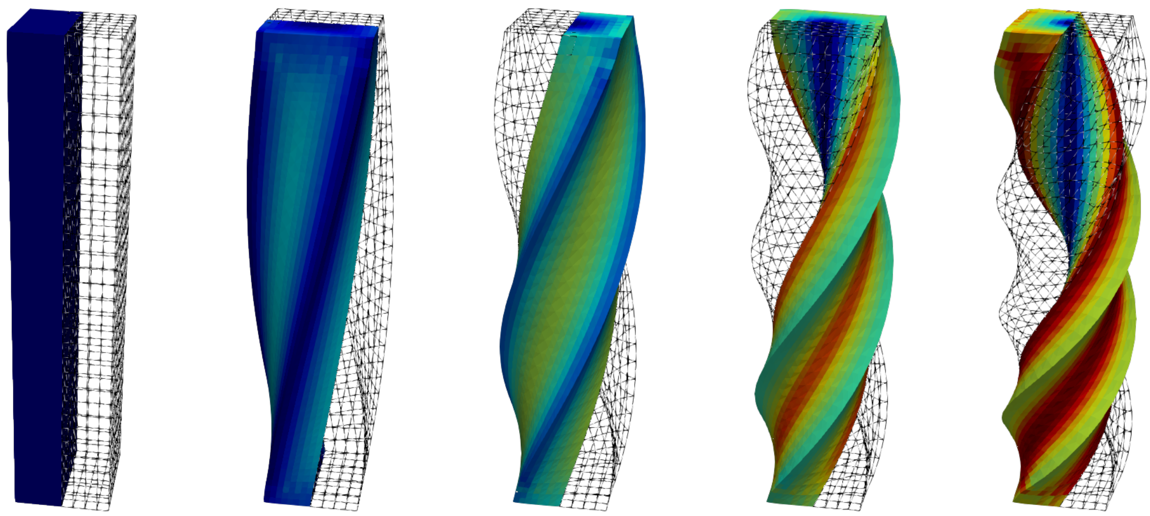

Figure 1. Torsion of a square-section bar: Equivalent plastic strain p (values between 0 (blue) to 0.49 (red)) for HHO with polynomial degree $k = 1$ at the quadrature points for different rotation angles Θ. From left to right: $\Theta = 0°$, $\Theta = 90°$, $\Theta = 180°$, $\Theta = 270°$, and $\Theta = 360°$

Our team contributes actively to the development of hybrid high-order (HHO) methods for nonlinear solid mechanics. Within the PhD of Nicolas Pignet in collaboration with EDF we have addressed several nonlinearities, including plasticity, large deformations, contact, and (Tresca) friction [15], [14], [49]. The advantage with respect to conforming finite elements is the robustness with respect to volumetric locking. The advantage with respect to mixed approaches is computational efficiency avoiding saddle-point formulations and additional unknowns. The advantage with respect to discontinuous Galerkin methods is avoiding the integration of the nonlinear behavior law at face quadrature nodes and the use of symmetric tangent matrices within Newton's method. The torsion of a square-section elastoplastic bar is presented in Figure 1. The color filling reports the equivalent plastic strain. The solution is obtained with the HHO method using the polynomial degree $k = 1$ for the face and the cell unknowns.

7.2. A hybrid high-order method for flow simulations in discrete fracture networks

Participants: Florent Hédin, Géraldine Pichot, Alexandre Ern.

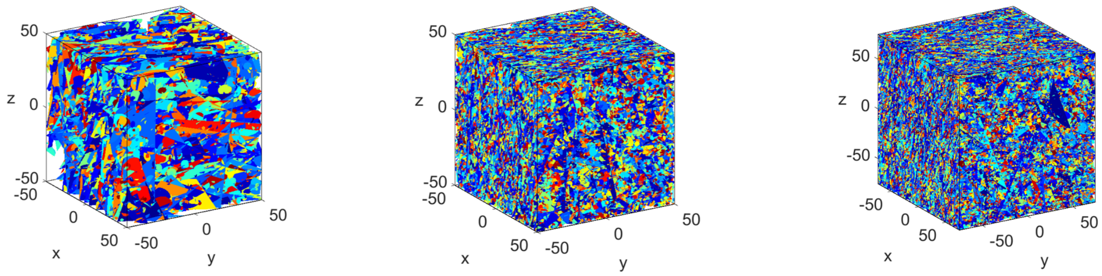

Figure 2. Examples of DFN: (left) B1: 19,007 fractures; (center) B2: 152,399 fractures ; (right) B3: 508,338 fractures.

In [36], we are interested in solving flow in large trimensional Discrete Fracture Networks (DFN) (cf Figure 2) with the hybrid high-order (HHO) method. The objectives of this paper are: (1) to demonstrate the benefit of using a high-order method for computing macroscopic quantities, like the equivalent permeability of fracture rocks; (2) to present the computational efficiency of our C++ software, `NEF++`, which implements the solving of flow in fractures based on the HHO method.

7.3. Analytic expressions of the solutions of advection-diffusion problems in 1D with discontinuous coefficients

Participants: Antoine Lejay, Lionel Lenôtre, Géraldine Pichot.

In [30], we provide a method to compute analytic expressions of the resolvent kernel of differential operators of the diffusion type with discontinuous coefficients in one dimension. Then we apply it when the coefficients are piecewise constant. We also perform the Laplace inversion of the resolvent kernel to obtain expressions of the transition density functions or fundamental solutions. We show how these explicit formula are useful to simulate advection-diffusion problems using particle tracking techniques.

7.4. An exponential timestepping algorithm for diffusion with discontinuous coefficients

Participants: Antoine Lejay, Lionel Lenôtre, Géraldine Pichot.

In [29], we present a new Monte Carlo algorithm to simulate diffusion processes in presence of discontinuous convective and diffusive terms. The algorithm is based on the knowledge of close form analytic expressions of the resolvents of the diffusion processes which are usually easier to obtain than close form analytic expressions of the density. In the particular case of diffusion processes with piecewise constant coefficients, known as Skew Diffusions, such close form expressions for the resolvent are available. Then we apply our algorithm to this particular case and we show that the approximate densities of the particles given by the algorithm replicate well the particularities of the true densities (discontinuities, bimodality, ...) Besides, numerical experiments show a quick convergence.

7.5. Polynomial-degree-robust multilevel algebraic error estimator & solver

Participants: Ani Miraci, Jan Papež, Martin Vohralík.

In [58], we devise a novel multilevel a posteriori estimator of the algebraic error. It delivers a fully computable, guaranteed lower bound on the error between an unknown exact solution of a system of linear algebraic equations and its approximation by an algebraic solver. The bound is also proved to be efficient, i.e., it also gives an upper bound on the algebraic error. Remarkably, the quality of these bounds is independent of the approximation polynomial degree. The derived estimates give immediately rise to a multilevel iterative algebraic solver whose contraction factor is independent of the polynomial degree of the approximation. We actually prove an equivalence between efficiency of the estimator and contraction of the solver. The estimator/solver are based on a global coarsest-level solve of lowest-order ($p = 1$), followed by local patchwise p-degree problems solved on the other levels. It corresponds to a V-cycle geometric multigrid solver with zero pre- and one post-smoothing step via block-Jacobi. A salient feature is the choice of the optimal step size for the descent direction.

7.6. Local- and global-best equivalence, simple projector, and optimal hp approximation in $\mathbf{H}(\mathrm{div})$

Participants: Alexandre Ern, Thirupathi Gudi, Iain Smears, Martin Vohralík.

$$\min_{\substack{\boldsymbol{v}_\mathcal{T}\in \boldsymbol{RTN}_p(\mathcal{T})\cap \boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)\\ \nabla\cdot\boldsymbol{v}_\mathcal{T}=\Pi^p_\mathcal{T}(\nabla\cdot\boldsymbol{v})}} \|\boldsymbol{v}-\boldsymbol{v}_\mathcal{T}\|^2_\Omega$$
$$\approx \sum_{K\in\mathcal{T}} \min_{\boldsymbol{v}_K\in \boldsymbol{RTN}_p(K)} \|\boldsymbol{v}-\boldsymbol{v}_K\|^2_K$$

Figure 3. Equivalence between global-best and local-best approximation for any $\mathbf{H}(\mathrm{div})$ function $\mathbf{v}$ with zero normal flux over part of the boundary and piecewise polynomial divergence

In [53], we prove that a global-best approximation in $\mathbf{H}(\mathrm{div})$, with constraints on normal component continuity and divergence, is equivalent to the sum of independent local-best approximations, without any constraints, as illustrated in Figure 3. This may seem surprising on a first sight since the right term in Figure 3 is seemingly much smaller (since the minimization set is unconstrained and thus much bigger). This result leads to optimal a priori hp-error estimates for mixed and least-squares finite element methods, which were missing in the literature until 2019. Additionally, the construction we devise gives rise to a simple stable local commuting projector in $\mathbf{H}(\mathrm{div})$, which delivers approximation error equivalent to the local-best approximation and applies under the minimal necessary Sobolev $\mathbf{H}(\mathrm{div})$ regularity, which is another result that has been sought for a very long time.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

One new two-part contract with EDF accompanying the PhD thesis of Idrissa Niakh.

One new two-part contract with CEA accompanying the postdoc of Guillaume Delay.

One new two-part contract with IFP Energies Nouvelles accompanying the PhD thesis of Joëlle Ferzly.

Three-part contract Inria–EDF–Sciworks Technologies (2017–2020) on “Form-L for the formalization of constraints of complex systems”. SERENA representants are François Clément, Sébastien Furic and Pierre Weis.

AMIES contract with ITASCA, January 10, 2019 - March 10, 2020. SERENA representants are François Clément, Sébastien Furic, Florent Hédin, Michel Kern, Géraldine Pichot.

9. Partnerships and Cooperations

9.1. Regional Initiatives

MILC (DMI RFSI, 2018–2019): “Mesure et Intégrale de Lebesgue en Coq”, with LIPN (Université de Paris 13), and TOCCATA (Inria Saclay - Île-de-France). SERENA representants are François Clément and Vincent Martin (UTC).

GiS: scientific collaboration network between ten public institutions from the Paris (Ile-de-France) region, focused on natural resources and environment. The project-team SERENA is a member.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

- ERC GATIPOR: “Guaranteed fully adaptive algorithms with tailored inexact solvers for complex porous media flows”. The subject of this consolidator grant are new approaches to porous media multiphase flows: inexact Newton-multigrid solvers, local stopping criteria, adaptivity, and a posteriori error control. The goal is to guarantee the overall simulation error and to speed-up importantly the present-day simulations. SERENA representant is M. Vohralík (grant leader, 75% commitment), period 2015–2020.
- ERC EMC2: “ Extreme-scale Mathematically-based Computational Chemistry”. The goal of this project is to develop physical and chemical models in chemistry, condensed matter physics, molecular biology, materials science, and nanosciences, altogether with mathematically-certified and numerically-efficient algorithms, and to implement them in a scalable way on various computer architectures. There are 4 principal investigators and a little more than 10 co-investigators. SERENA representant is M. Vohralík (co-investigator, 10% commitment), period 2019–2025.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

Erik Burman, Professor, University College London, unfitted methods.

Ulrich Rüde, Professor, University of Erlangen-Nürnberg, multigrid methods.

Iain Smears, Lecturer, University College London, local-global approximations.

Benjamin Stamm, Professor, RTWH Aachen University, eigenvalue problems (first-principle molecular simulation).

Barbara Wohlmuth, professor, Technical University Munich, multigrid methods.

9.3.2. Participation in Other International Programs

9.3.2.1. Inria International Chairs

IIC GUERMOND Jean-Luc

Title: Curved H(div), H(curl) elements, and magnetohydrodynamics & Approximation of hyperbolic systems

International Partner (Institution - Laboratory - Researcher):

Texas A&M University (United States) - Department of Mathematics - Jean-Luc Guermond

Duration: 2019 - 2023

Start year: 2019

See also: https://www.math.tamu.edu/~guermond/

The program is articulated around two themes: (1) Theoretical aspects in finite elements and applications to multi-physics magneto-hydrodynamics; (2) Finite element approximation of hyperbolic systems and applications. The results from this research will have applications in problems related to porous media flows, magnetohydrodynamics, water management, and compressible and incompressible fluid flows.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

Hend Ben Ameur, Professor at IPEST and member of ENIT-Lamsin, Tunis, Tunisia, November 4–15.

Gregor Gantner, Vienna University of Technology. Collaboration on IGA methods. September 23–27.

Thirupathi Gudi, Indian Institute of Science, Bangalore. Collaboration on local-global approximations. June 10–17.

Dirk Praetorius, Vienna University of Technology. Collaboration on cost-optimality of fully adaptive algorithms. March 21–23.

Ivan Yotov, Professor, University of Pittsburgh. Inria Paris invited professor, September 1–December 15, 2019. Collaboration on multilevel and space-time domain decomposition methods.

9.4.1.1. Internships

Théo Kaprélian, internship at Ecole Centrale de Lyon, from September 2019 to February 2020, supervised by Martin Vohralík.

9.4.2. Visits to International Teams

9.4.2.1. Research Stays Abroad

+ Géraldine Pichot was invited for a one week stay at Pennstate University, USA for a collaboration with Pr. Ludmil Zikatanov.
+ Géraldine Pichot was invited for a one week stay at University of Bergen, Norway for a collaboration with Pr. Florin Radu.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. Member of the Organizing Committees

Michel Kern was co-organizer of the SITRAM19 workshop (Advances in the SImulation of reactive flow and TRAnsport in porous Media), December 2–3, Pau, France.

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

Guillaume Enchéry and Martin Vohralík have organized a 1-day workshop Journée contrat cadre IFP Energies Nouvelles/Inria.

10.1.2.2. Member of the Conference Program Committees

Alexandre Ern is a member of the Scientific Committee of the European Finite Element Fair (EFEF) and of the European Numerical Mathematics and Advanced Applications Conference (ENUMATH).

Géraldine Pichot is a member of the Scientific Commitee of FRAME2020 - FRActured MEdia: Numerical methods for fluid flow and mechanics, September 21-24, 2020.

Michel Kern and Martin Vohralík were members of the Scientific Committee of the SimRace 2019 conference.

Michel Kern was a member of the Program Comittee of JCAD2019, Journées Calcul et Données .

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Alexandre Ern is a member of the editorial boards of SIAM Journal on Scientific Computing, ESAIM Mathematical Modelling and Numerical Analysis, IMA Journal of Numerical Analysis, and Computational Methods in Applied Mathematics.

Martin Vohralík is a member of the editorial boards of SIAM Journal on Numerical Analysis, Acta Polytechnica, and Applications of Mathematics.

10.1.3.2. Reviewer - Reviewing Activities

Alexandre Ern and Martin Vohralík served as reviewers for dozens of papers in different journals.

Michel Kern serrved as reviewer for BIT, Numerical Mathematics, Computational Geosciences, SIAM Journal on Scientific Computing, Applicable Analysis, and Journal of Mathematical Study.

Géraldine Pichot served as reviewer for Computational Geosciences.

10.1.4. Invited Talks

Alexandre Ern was invited as one of the two keynote lecturers at the Workshop on Computational Modeling and Numerical Analysis, Petropolis, Brésil, Feb 2019.

Michel Kern was invited as one of the keynote speakers at the HPC TerrSys fall school, Bonn, Germany, Sep. 2019.

Martin Vohralík was an invited speaker at GAMM section on Numerical methods of differential equations, Vienna, Austria, February 2019, Modern Finite Element Technologies 2019, Bad Honnef, Germany, July 2019, and Reliable Methods of Mathematical Modeling, Vienna, Austria, Sep 2019.

10.1.5. Leadership within the Scientific Community

Alexandre Ern is the coordinator of the Master Mathématiques et applications, Ecole nationale des ponts et chaussées.

Michel Kern is a member of the SIAM "Committee on Committees and Appointments".

Michel Kern is a member of the Scientific Board of ORAP (Organisation Associative du Parallélisme).

Michel Kern is a member of the Scientific Advisory Board of GDR CNRS Groupe Calcul.

Michel Kern is a member of the steering committee of Géosciences franciliennes.

Martin Vohralík is a member of the steering committee of Summer schools CEA–EDF–Inria.

Martin Vohralík is in charge of the topic "Numerical schemes, mesh generation algorithms, and error control" in the ANDRA, BRGM, CEA, EDF, IFP Energies Nouvelles, and Total working group on *High-Performance Numerical Simulation in the Geosciences* (identification of common challenges and collaboration opportunities).

10.1.6. Research Administration

François Clément is a member of the *Comité local d'hygiène, de sécurité et des conditions de travail* of the Inria Research Center of Paris.

Michel Kern is a member of the *Comité de Centre* of the Inria Research Center of Paris.

Martin Vohralík is a member of the Inria Paris *Committee on scientific positions* (evaluation of applications for Ph.D. theses (CORDI-S), post-docs, and délégations).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence : Alexandre Ern, Numerical analysis, 77h, L3/M1, Ecole Polytechnique, France.

Licence : Alexandre Ern, Optimal control, 10h, L3, Ecole nationale des ponts et chaussées, France.

Master : Alexandre Ern, Discontinuous Galerkin methods, 20h, M2, Sorbonne University, France.

Master: Michel Kern, Models and numerical methods for subsurface flow, 30h, M2, Université Paris Saclay, France

Master: Michel Kern, Inverse Problems, 24h, M1, Mines-ParisTech, France

Master: Michel Kern, Introduction to Partial Differential Equations, 30h, & Advanced numerical analysis, 30h, M1, Institut Galilée, Université Paris-Nord, France.

10.2.2. Supervision

PhD defended on June 3, 2019: Jad Dabaghi, A posteriori error estimates and adaptive stopping criteria for formulations with complementarity constraints, Martin Vohralík and Vincent Martin.

PhD defended on March 22, 2019: Patrik Daniel, Adaptive hp-finite elements with guaranteed error contraction and inexact multilevel solvers, Martin Vohralík and Alexandre Ern.

PhD in progress: Joëlle Ferzly, A posteriori error estimates and adaptivity for complementarity problems, started 15 October 2019, Martin Vohralík.

PhD in progress: Frédéric Marazzato, Discrete element methods for fracture and fragmentation, started 01 October 2016, Alexandre Ern.

PhD in progress: Riccardo Milani, Compatible Discrete Operator schemes for Navier–Stokes equations, started 01 October 2017, Alexandre Ern.

PhD in progress: Ani Miraci, Robust a posteriori error control and adaptivity with inexact solvers, started 01 October 2017, Martin Vohralík and Alexandre Ern.

PhD in progress: Idrissa Niakh, Reduced models for variational inequalities in computational mechanics, started 01 November 2019, Alexandre Ern and Virginie Ehrlacher (ENPC).

PhD in progress: Stefano Piccardo, High-fidelity simulation of droplets in complex shear flows, started 01 October 2019, Alexandre Ern and Antonio Huerta (UPC).

PhD defended on October 22, 2019: Nicolas Pignet, Hybrid high-order methods for nonlinear solid mechanics, Alexandre Ern.

10.2.3. Juries

Alexandre Ern, Examiner and Chairman, PhD V. Mamet, INP Grenoble, Jan 2019.

Alexandre Ern, Referee, PhD A. Rupp, Erlangen FA University, Mar 2019.

Alexandre Ern, Examiner, HDR C. Millet, University Paris-Sud, Mar 2019.

Michel Kern, Examiner, PhD Daniel Salas, Université de Strasbourg, Sep. 2019.

Michel Kern, Examiner, PhD Kevish Napal, Université Paris Saclay, Dec. 2019.

Géraldine Pichot, Referee, PhD Fabio Vicini, Politecnico di Torino, April 2019.

Martin Vohralík, Examiner and Chairman, PhD Hoang Phuong Thai, ENS Paris-Saclay, Jun 2019.

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

+ Géraldine Pichot is a member of the editorial committee of Interstices since November 2019.

11. Bibliography

Major publications by the team in recent years

[1] 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://dx.doi.org/10.1007/s10817-012-9255-4

[2] 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", August 2014, vol. 68, n^{o} 3, p. 325–352, http://dx.doi.org/10.1016/j.camwa.2014.06.004

[3] E. Cancès, G. Dusson, Y. Maday, B. Stamm, M. Vohralík.*Guaranteed and robust a posteriori bounds for Laplace eigenvalues and eigenvectors: conforming approximations*, in "SIAM J. Numer. Anal.", 2017, vol. 55, n^{o} 5, p. 2228–2254, http://dx.doi.org/10.1137/15M1038633

[4] D. A. Di Pietro, A. Ern.*A hybrid high-order locking-free method for linear elasticity on general meshes*, in "Comput. Methods Appl. Mech. Engrg.", 2015, vol. 283, p. 1–21, http://dx.doi.org/10.1016/j.cma.2014.09.009

[5] A. Ern, J.-L. Guermond.*Finite element quasi-interpolation and best approximation*, in "ESAIM Math. Model. Numer. Anal.", 2017, vol. 51, n^{o} 4, p. 1367–1385, https://doi.org/10.1051/m2an/2016066

[6] A. Ern, M. Vohralík.*Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations*, in "SIAM J. Numer. Anal.", 2015, vol. 53, n^{o} 2, p. 1058–1081, http://dx.doi.org/10.1137/130950100

[7] T.-T.-P. HOANG, J. JAFFRÉ, C. JAPHET, M. KERN, J. E. ROBERTS.*Space-time domain decomposition methods for diffusion problems in mixed formulations*, in "SIAM J. Numer. Anal.", 2013, vol. 51, n^{o} 6, p. 3532–3559, http://dx.doi.org/10.1137/130914401

[8] T.-T.-P. HOANG, C. JAPHET, M. KERN, J. E. ROBERTS.*Space-time domain decomposition for reduced fracture models in mixed formulation*, in "SIAM J. Numer. Anal.", 2016, vol. 54, n^{o} 1, p. 288–316, http://dx.doi.org/10.1137/15M1009651

[9] A. LEJAY, G. PICHOT.*Simulating diffusion processes in discontinuous media: a numerical scheme with constant time steps*, in "J. Comput. Phys.", 2012, vol. 231, n^{o} 21, p. 7299–7314, http://dx.doi.org/10.1016/j.jcp.2012.07.011

[10] G. PICHOT, J. ERHEL, J.-R. DE DREUZY.*A generalized mixed hybrid mortar method for solving flow in stochastic discrete fracture networks*, in "SIAM J. Sci. Comput.", 2012, vol. 34, n^{o} 1, p. B86–B105, http://dx.doi.org/10.1137/100804383

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] J. DABAGHI.*A posteriori error estimates for variational inequalities: application to a two-phase flow in porous media*, Sorbonne Université, Université Pierre et Marie Curie, Paris 6, June 2019, https://hal.archives-ouvertes.fr/tel-02151951

[12] P. DANIEL.*Adaptive hp-finite elements with guaranteed error contraction and inexact multilevel solvers*, Sorbonne Université ; Inria Paris, March 2019, https://hal.inria.fr/tel-02104982

[13] N. PIGNET.*Hybrid High-Order methods for nonlinear solid mechanics*, Université Paris-Est Marne la Vallée, October 2019, https://tel.archives-ouvertes.fr/tel-02318157

Articles in International Peer-Reviewed Journal

[14] M. ABBAS, A. ERN, N. PIGNET.*A Hybrid High-Order method for finite elastoplastic deformations within a logarithmic strain framework*, in "International Journal for Numerical Methods in Engineering", June 2019 [*DOI :* 10.1002/NME.6137], https://hal.archives-ouvertes.fr/hal-01978385

[15] M. ABBAS, A. ERN, N. PIGNET.*A Hybrid High-Order method for incremental associative plasticity with small deformations*, in "Computer Methods in Applied Mechanics and Engineering", April 2019, vol. 346, p. 891-912, https://arxiv.org/abs/1804.06129 [*DOI :* 10.1016/J.CMA.2018.08.037], https://hal.archives-ouvertes.fr/hal-01768411

[16] E. AHMED, S. ALI HASSAN, C. JAPHET, M. KERN, M. VOHRALÍK.*A posteriori error estimates and stopping criteria for space-time domain decomposition for two-phase flow between different rock types*, in "SMAI Journal of Computational Mathematics", December 2019, vol. 5, p. 195-227, https://hal.inria.fr/hal-01540956

[17] L. AMIR, M. KERN.*Preconditioning a coupled model for reactive transport in porous media*, in "International Journal of Numerical Analysis and Modeling", 2019, vol. 16, n^{o} 1, p. 18-48, https://arxiv.org/abs/1710.01483 , https://hal.inria.fr/hal-01327307

[18] H. BARUCQ, G. CHAVENT, F. FAUCHER.*A priori estimates of attraction basins for nonlinear least squares, with application to Helmholtz seismic inverse problem*, in "Inverse Problems", 2019, forthcoming [*DOI :* 10.1088/1361-6420/AB3507], https://hal.archives-ouvertes.fr/hal-02194212

[19] I. BEN GHARBIA, J. DABAGHI, V. MARTIN, M. VOHRALÍK.*A posteriori error estimates and adaptive stopping criteria for a compositional two-phase flow with nonlinear complementarity constraints*, in "Computational Geosciences", December 2019 [*DOI :* 10.1007/S10596-019-09909-5], https://hal.archives-ouvertes.fr/hal-01919067

[20] A. BENACEUR, A. ERN, V. EHRLACHER.*A reduced basis method for parametrized variational inequalities applied to contact mechanics*, in "International Journal for Numerical Methods in Engineering", October 2019 [*DOI :* 10.1002/NME.6261], https://hal.archives-ouvertes.fr/hal-02081485

[21] J. BLECHTA, J. MÁLEK, M. VOHRALÍK.*Localization of the $W^{-1,q}$ norm for local a posteriori efficiency*, in "IMA Journal of Numerical Analysis", March 2019 [*DOI :* 10.1093/IMANUM/DRZ002], https://hal.inria.fr/hal-01332481

[22] S. BLIUDZE, S. FURIC, J. SIFAKIS, A. VIEL.*Rigorous Design of Cyber-Physical Systems : Linking Physicality and Computation*, in "Software and Systems Modeling", 2019, vol. 18, n^{o} 3, p. 1613–1636, forthcoming [*DOI :* 10.1007/S10270-017-0642-5], https://hal.inria.fr/hal-01636392

[23] T. BOIVEAU, V. EHRLACHER, A. ERN, A. NOUY.*Low-rank approximation of linear parabolic equations by space-time tensor Galerkin methods*, in "ESAIM: Mathematical Modelling and Numerical Analysis", May 2019, vol. 53, n^{o} 2, p. 635-658, https://arxiv.org/abs/1712.07256 [*DOI :* 10.1051/M2AN/2018073], https://hal.archives-ouvertes.fr/hal-01668316

[24] V. M. CALO, M. CICUTTIN, Q. DENG, A. ERN.*Spectral approximation of elliptic operators by the Hybrid High-Order method*, in "Mathematics of Computation", 2019, vol. 88, n^{o} 318, p. 1559-1586, https://hal.archives-ouvertes.fr/hal-01628698

[25] M. CICUTTIN, A. ERN, S. LEMAIRE.*A Hybrid High-Order method for highly oscillatory elliptic problems*, in "Computational Methods in Applied Mathematics", 2019, vol. 19, n^{o} 4, p. 723-748 [*DOI :* 10.1515/CMAM-2018-0013], https://hal.archives-ouvertes.fr/hal-01467434

[26] P. DANIEL, A. ERN, M. VOHRALÍK.*An adaptive hp-refinement strategy with inexact solvers and computable guaranteed bound on the error reduction factor*, in "Computer Methods in Applied Mechanics and Engineering", November 2019, vol. 359, 112607 [*DOI :* 10.1016/J.CMA.2019.112607], https://hal.inria.fr/hal-01931448

[27] A. ERN, I. SMEARS, M. VOHRALÍK.*Equilibrated flux a posteriori error estimates in $L^2(H^1)$-norms for high-order discretizations of parabolic problems*, in "IMA Journal of Numerical Analysis", 2019, vol. 39, n^{o} 3, p. 1158-1179 [*DOI :* 10.1093/IMANUM/DRY035], https://hal.inria.fr/hal-01489721

[28] A. ERN, M. VOHRALÍK.*Stable broken H1 and H(div) polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions*, in "Mathematics of Computation", October 2019 [*DOI :* 10.1090/MCOM/3482], https://hal.inria.fr/hal-01422204

[29] A. LEJAY, L. LENÔTRE, G. PICHOT.*An exponential timestepping algorithm for diffusion with discontinuous coefficients*, in "Journal of Computational Physics", November 2019, vol. 396, p. 888-904 [*DOI :* 10.1016/J.JCP.2019.07.013], https://hal.inria.fr/hal-01806465

[30] A. LEJAY, L. LENÔTRE, G. PICHOT.*Analytic expressions of the solutions of advection-diffusion problems in 1D with discontinuous coefficients*, in "SIAM Journal on Applied Mathematics", September 2019, vol. 79, n^{o} 5, p. 1823-1849 [*DOI :* 10.1137/18M1164500], https://hal.inria.fr/hal-01644270

[31] G. MALLIK, M. VOHRALÍK, S. YOUSEF.*Goal-oriented a posteriori error estimation for conforming and nonconforming approximations with inexact solvers*, in "Journal of Computational and Applied Mathematics", August 2019, vol. 366, 112367 [*DOI :* 10.1016/J.CAM.2019.112367], https://hal.inria.fr/hal-01964733

[32] F. MARAZZATO, A. ERN, C. MARIOTTI, L. MONASSE.*An explicit pseudo-energy conserving time-integration scheme for Hamiltonian dynamics*, in "Computer Methods in Applied Mechanics and Engineering", April 2019, vol. 347, p. 906-927 [*DOI :* 10.1016/J.CMA.2019.01.013], https://hal-enpc.archives-ouvertes.fr/hal-01661608

International Conferences with Proceedings

[33] E. BURMAN, A. ERN.*A cut-cell Hybrid High-Order method for elliptic problems with curved boundaries*, in "ENUMATH 2017", Bergen, Norway, EnuMath 2019 - European Numerical Mathematics and Advanced Applications Conference, 2019 [*DOI :* 10.1007/978-3-319-96415-7_14], https://hal.archives-ouvertes.fr/hal-01653685

[34] T. CHAUMONT-FRELET, A. ERN, M. VOHRALÍK.*Asymptotically constant-free, p-robust and guaranteed a posteriori error estimates for the Helmholtz equation*, in "EnuMath 2019 - European Numerical Mathematics and Advanced Applications Conference", Egmond aan Zee, Netherlands, September 2019, https://hal.inria.fr/hal-02321140

[35] M. CICUTTIN, A. ERN, S. LEMAIRE.*On the implementation of a multiscale Hybrid High-Order method*, in "ENUMATH 2017", Bergen, Norway, I. BERRE, K. KUMAR, J. M. NORDBOTTEN, I. S. POP, F. A. RADU (editors), Numerical Mathematics and Advanced Applications - ENUMATH 2017, Springer, Cham, 2019, vol. 126, p. 509-517 [*DOI :* 10.1007/978-3-319-96415-7_46], https://hal.archives-ouvertes.fr/hal-01661925

[36] F. HÉDIN, G. PICHOT, A. ERN.*A hybrid high-order method for flow simulations in discrete fracture networks*, in "ENUMATH - European Numerical Mathematics and Advanced Applications Conference 2019", Egmond aan Zee, Netherlands, September 2019, https://hal.inria.fr/hal-02315491

Research Reports

[37] H. BARUCQ, H. CALANDRA, G. CHAVENT, F. FAUCHER.*A priori estimates of attraction basins for velocity model reconstruction by time-harmonic Full Waveform Inversion and Data Space Reflectivity formulation*, Magique 3D ; Inria Bordeaux Sud-Ouest ; Université de Pau et des Pays de l'Adour, February 2019, n^{o} RR-9253, Material under review by journal Geophysics, https://hal.archives-ouvertes.fr/hal-02016373

Other Publications

[38] E. AHMED, C. JAPHET, M. KERN.*Space-time domain decomposition for two-phase flow between different rock types*, January 2020, working paper or preprint, https://hal.inria.fr/hal-02275690

[39] A. BENACEUR.*A time-dependent Parametrized Background Data-Weak approach*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02330408

[40] A. BENACEUR.*An inexpensive Parametrized Background Data-Weak approach for time-dependent problems*, August 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02265533

[41] E. BURMAN, M. CICUTTIN, G. DELAY, A. ERN.*An unfitted Hybrid High-Order method with cell agglomeration for elliptic interface problems*, September 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02280426

[42] E. BURMAN, G. DELAY, A. ERN.*The unfitted HHO method for the Stokes problem on curved domains*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02438622

[43] V. M. CALO, A. ERN, I. MUGA, S. ROJAS.*An adaptive stabilized conforming finite element method via residual minimization on dual discontinuous Galerkin norms*, 2019, https://arxiv.org/abs/1907.12605 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02196242

[44] E. CANCÈS, G. DUSSON, Y. MADAY, B. STAMM, M. VOHRALÍK.*Guaranteed a posteriori bounds for eigenvalues and eigenvectors: multiplicities and clusters*, May 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02127954

[45] E. CANCÈS, G. DUSSON, Y. MADAY, B. STAMM, M. VOHRALÍK.*Post-processing of the planewave approximation of Schrödinger equations. Part I: linear operators*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01908039

[46] C. CANCÈS, F. NABET, M. VOHRALÍK.*Convergence and a posteriori error analysis for energy-stable finite element approximations of degenerate parabolic equations*, January 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01894884

[47] J.-P. CHANCELIER, S. FURIC, P. WEIS.*Translating Simulink Models to Modelica using the Nsp Platform*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01948681

[48] T. CHAUMONT-FRELET, A. ERN, M. VOHRALÍK.*On the derivation of guaranteed and p-robust a posteriori error estimates for the Helmholtz equation*, July 2019, working paper or preprint, https://hal.inria.fr/hal-02202233

[49] F. CHOULY, A. ERN, N. PIGNET.*A Hybrid High-Order discretization combined with Nitsche's method for contact and Tresca friction in small strain elasticity*, September 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02283418

[50] M. CICUTTIN, A. ERN, T. GUDI.*Hybrid high-order methods for the elliptic obstacle problem*, May 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01718883

[51] J. DABAGHI, V. MARTIN, M. VOHRALÍK.*A posteriori estimates distinguishing the error components and adaptive stopping criteria for numerical approximations of parabolic variational inequalities*, August 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02274493

[52] J. ERHEL, M. OUMOUNI, G. PICHOT, F. SCHOEFS.*Analysis of continuous spectral method for sampling stationary Gaussian random fields*, April 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02109037

[53] A. ERN, T. GUDI, I. SMEARS, M. VOHRALÍK.*Equivalence of local-and global-best approximations, a simple stable local commuting projector, and optimal hp approximation estimates in H(div)*, August 2019, https://arxiv.org/abs/1908.08158 - working paper or preprint, https://hal.inria.fr/hal-02268960

[54] A. ERN, M. VOHRALÍK, M. ZAKERZADEH.*Guaranteed and robust L^2-norm a posteriori error estimates for 1D linear advection problems*, April 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02105418

[55] A. ERN, P. ZANOTTI.*A quasi-optimal variant of the hybrid high-order method for elliptic PDEs with H^{-1} loads*, April 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02114715

[56] F. MARAZZATO, A. ERN, L. MONASSE.*A consistent discrete element method for quasi-static and dynamic elasto-plasticity*, November 2019, https://arxiv.org/abs/1911.00738 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02343280

[57] R. MILANI, J. BONELLE, A. ERN.*Compatible Discrete Operator schemes for the steady incompressible Stokes and Navier-Stokes equations*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02417514

[58] A. MIRAÇI, J. PAPEŽ, M. VOHRALÍK.*A multilevel algebraic error estimator and the corresponding iterative solver with p-robust behavior*, July 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02070981

[59] J. PAPEŽ, U. RÜDE, M. VOHRALÍK, B. WOHLMUTH.*Sharp algebraic and total a posteriori error bounds for h and p finite elements via a multilevel approach. Recovering mass balance in any situation*, December 2019, working paper or preprint, https://hal.inria.fr/hal-01662944

[60] J. PAPEŽ, M. VOHRALÍK.*Inexpensive guaranteed and efficient upper bounds on the algebraic error in finite element discretizations*, January 2020, working paper or preprint, https://hal.inria.fr/hal-02422851

Project-Team SIERRA

Statistical Machine Learning and Parsimony

IN COLLABORATION WITH: Département d'Informatique de l'Ecole Normale Supérieure

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Paris

RESEARCH CENTER
Paris

THEME
Optimization, machine learning and statistical methods

Table of contents

Project-Team SIERRA

Creation of the Team: 2011 January 01, updated into Project-Team: 2012 January 01

Keywords:

Computer Science and Digital Science:

A3.4. - Machine learning and statistics
A5.4. - Computer vision
A6.2. - Scientific computing, Numerical Analysis & Optimization
A7.1. - Algorithms
A8.2. - Optimization
A9.2. - Machine learning

Other Research Topics and Application Domains:

B9.5.6. - Data science

1. Team, Visitors, External Collaborators

Research Scientists

Francis Bach [Team leader, Inria, Senior Researcher, HDR]
Alexandre d'Aspremont [CNRS, Senior Researcher, HDR]
Pierre Gaillard [Inria, Researcher]
Alessandro Rudi [Inria, Starting Research Position]
Adrien Taylor [Inria, Starting Research Position, from Oct 2019]

Faculty Member

Marco Cuturi Cameto [École nationale supérieure de l'électronique et de ses applications, Associate Professor, until Sep 2019]

Technical Staff

Loïc Estève [Inria, Engineer]
Gautier Izacard [CNRS, Engineer, from Nov 2019]

PhD Students

Dmitry Babichev [Inria, PhD Student, until Feb 2019]
Mathieu Barre [École Normale Supérieure de Paris, PhD Student]
Eloise Berthier [DGA, PhD Student, from Sep 2019]
Raphael Berthier [Inria, PhD Student]
Margaux Bregere [EDF, PhD Student]
Vivien Cabannes [Inria, PhD Student, from Sep 2019]
Alexandre Defossez [Facebook, PhD Student]
Radu Alexandru Dragomir [École polytechnique, PhD Student]
Remi Jezequel [Inria, PhD Student, from Sep 2019]
Thomas Kerdreux [École polytechnique, PhD Student]
Ulysse Marteau Ferey [Inria, PhD Student, from Sep 2019]
Gregoire Mialon [Inria, PhD Student]
Alex Nowak Vila [La Caixa fellowship, PhD Student]
Loucas Pillaud Vivien [Ministère de l'Ecologie, de l'Energie, du Développement durable et de la Mer, PhD Student]
Tatiana Shpakova [Inria, PhD Student, until Feb 2019]
Remi Leblond [Inria, until Feb 2019]

Post-Doctoral Fellows

Pierre Yves Masse [CTU Prague, Post-Doctoral Fellow]
Dmitrii Ostrovskii [Inria, Post-Doctoral Fellow, until Jul 2019]
Yifan Sun [École Normale Supérieure de Paris, Post-Doctoral Fellow, from Sep 2019]
Adrien Taylor [Inria, Post-Doctoral Fellow, until Sep 2019]

Visiting Scientists

Quentin Berthet [University of Cambridge, from Feb 2019 until Apr 2019]
Eduard Gorbunov [Moscow Institute of Physics and Technology, Oct 2019]
Samy Jelassi [Princeton University, from Jun 2019 until Sep 2019]
Ulysse Marteau Ferey [École Normale Supérieure de Paris, until Mar 2019]
Song Mei [Stanford University, from Sep 2019 until Oct 2019]
Geoffrey Negiar [U.C. Berkeley, from May 2019 until Jul 2019]
Anant Raj [M.P.I. Tubingen, from Oct 2019]
Aadirupa Saha [Indian Institute of Technology, Bangalore, from Nov 2019]

Administrative Assistants

Helene Bessin Rousseau [Inria, Administrative Assistant]
Helene Milome [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Statement

Machine learning is a recent scientific domain, positioned between applied mathematics, statistics and computer science. Its goals are the optimization, control, and modelisation of complex systems from examples. It applies to data from numerous engineering and scientific fields (e.g., vision, bioinformatics, neuroscience, audio processing, text processing, economy, finance, etc.), the ultimate goal being to derive general theories and algorithms allowing advances in each of these domains. Machine learning is characterized by the high quality and quantity of the exchanges between theory, algorithms and applications: interesting theoretical problems almost always emerge from applications, while theoretical analysis allows the understanding of why and when popular or successful algorithms do or do not work, and leads to proposing significant improvements.

Our academic positioning is exactly at the intersection between these three aspects—algorithms, theory and applications—and our main research goal is to make the link between theory and algorithms, and between algorithms and high-impact applications in various engineering and scientific fields, in particular computer vision, bioinformatics, audio processing, text processing and neuro-imaging.

Machine learning is now a vast field of research and the team focuses on the following aspects: supervised learning (kernel methods, calibration), unsupervised learning (matrix factorization, statistical tests), parsimony (structured sparsity, theory and algorithms), and optimization (convex optimization, bandit learning). These four research axes are strongly interdependent, and the interplay between them is key to successful practical applications.

3. Research Program

3.1. Supervised Learning

This part of our research focuses on methods where, given a set of examples of input/output pairs, the goal is to predict the output for a new input, with research on kernel methods, calibration methods, and multi-task learning.

3.2. Unsupervised Learning

We focus here on methods where no output is given and the goal is to find structure of certain known types (e.g., discrete or low-dimensional) in the data, with a focus on matrix factorization, statistical tests, dimension reduction, and semi-supervised learning.

3.3. Parsimony

The concept of parsimony is central to many areas of science. In the context of statistical machine learning, this takes the form of variable or feature selection. The team focuses primarily on structured sparsity, with theoretical and algorithmic contributions.

3.4. Optimization

Optimization in all its forms is central to machine learning, as many of its theoretical frameworks are based at least in part on empirical risk minimization. The team focuses primarily on convex and bandit optimization, with a particular focus on large-scale optimization.

4. Application Domains

4.1. Applications for Machine Learning

Machine learning research can be conducted from two main perspectives: the first one, which has been dominant in the last 30 years, is to design learning algorithms and theories which are as generic as possible, the goal being to make as few assumptions as possible regarding the problems to be solved and to let data speak for themselves. This has led to many interesting methodological developments and successful applications. However, we believe that this strategy has reached its limit for many application domains, such as computer vision, bioinformatics, neuro-imaging, text and audio processing, which leads to the second perspective our team is built on: Research in machine learning theory and algorithms should be driven by interdisciplinary collaborations, so that specific prior knowledge may be properly introduced into the learning process, in particular with the following fields:

- Computer vision: object recognition, object detection, image segmentation, image/video processing, computational photography. In collaboration with the Willow project-team.
- Bioinformatics: cancer diagnosis, protein function prediction, virtual screening. In collaboration with Institut Curie.
- Text processing: document collection modeling, language models.
- Audio processing: source separation, speech/music processing.
- Neuro-imaging: brain-computer interface (fMRI, EEG, MEG).

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Damien Scieur, Prix de thèse PSL-ADELI
- Francis Bach, Prix Jean-Jacques Moreau

6. New Software and Platforms

6.1. ProxASAGA

KEYWORD: Optimization

FUNCTIONAL DESCRIPTION: A C++/Python code implementing the methods in the paper "Breaking the Nonsmooth Barrier: A Scalable Parallel Method for Composite Optimization", F. Pedregosa, R. Leblond and S. Lacoste-Julien, Advances in Neural Information Processing Systems (NIPS) 2017. Due to their simplicity and excellent performance, parallel asynchronous variants of stochastic gradient descent have become popular methods to solve a wide range of large-scale optimization problems on multi-core architectures. Yet, despite their practical success, support for nonsmooth objectives is still lacking, making them unsuitable for many problems of interest in machine learning, such as the Lasso, group Lasso or empirical risk minimization with convex constraints. In this work, we propose and analyze ProxASAGA, a fully asynchronous sparse method inspired by SAGA, a variance reduced incremental gradient algorithm. The proposed method is easy to implement and significantly outperforms the state of the art on several nonsmooth, large-scale problems. We prove that our method achieves a theoretical linear speedup with respect to the sequential version under assumptions on the sparsity of gradients and block-separability of the proximal term. Empirical benchmarks on a multi-core architecture illustrate practical speedups of up to 12x on a 20-core machine.

- Contact: Fabian Pedregosa
- URL: https://github.com/fabianp/ProxASAGA

6.2. object-states-action

KEYWORD: Computer vision

FUNCTIONAL DESCRIPTION: Code for the paper Joint Discovery of Object States and Manipulation Actions, ICCV 2017: Many human activities involve object manipulations aiming to modify the object state. Examples of common state changes include full/empty bottle, open/closed door, and attached/detached car wheel. In this work, we seek to automatically discover the states of objects and the associated manipulation actions. Given a set of videos for a particular task, we propose a joint model that learns to identify object states and to localize state-modifying actions. Our model is formulated as a discriminative clustering cost with constraints. We assume a consistent temporal order for the changes in object states and manipulation actions, and introduce new optimization techniques to learn model parameters without additional supervision. We demonstrate successful discovery of seven manipulation actions and corresponding object states on a new dataset of videos depicting real-life object manipulations. We show that our joint formulation results in an improvement of object state discovery by action recognition and vice versa.

- Participants: Jean-Baptiste Alayrac, Josef Sivic, Ivan Laptev and Simon Lacoste-Julien
- Contact: Jean-Baptiste Alayrac
- Publication: Joint Discovery of Object States and Manipulation Actions
- URL: https://github.com/jalayrac/object-states-action

7. New Results

7.1. Fast Gradient Methods for Symmetric Nonnegative Matrix Factorization.

We describe fast gradient methods for solving the symmetric nonnegative matrix factorization problem (SymNMF). We use recent results on non-Euclidean gradient methods and show that the SymNMF problem is smooth relative to a well-chosen Bregman divergence. This approach provides a simple hyperparameter-free method which comes with theoretical convergence guarantees. We also discuss accelerated variants. Numerical experiments on clustering problems show that our algorithm scales well and reaches both state of the art convergence speed and clustering accuracy for SymNMF methods.

7.2. Naive Feature Selection: Sparsity in Naive Bayes.

Due to its linear complexity, naive Bayes classification remains an attractive supervised learning method, especially in very large-scale settings. We propose a sparse version of naive Bayes, which can be used for feature selection. This leads to a combinatorial maximum-likelihood problem, for which we provide an exact solution in the case of binary data, or a bound in the multinomial case. We prove that our bound becomes tight as the marginal contribution of additional features decreases. Both binary and multinomial sparse models are solvable in time almost linear in problem size, representing a very small extra relative cost compared to the classical naive Bayes. Numerical experiments on text data show that the naive Bayes feature selection method is as statistically effective as state-of-the-art feature selection methods such as recursive feature elimination, l1-penalized logistic regression and LASSO, while being orders of magnitude faster. For a large data set, having more than with 1.6 million training points and about 12 million features, and with a non-optimized CPU implementation, our sparse naive Bayes model can be trained in less than 15 seconds.

7.3. Regularity as Regularization: Smooth and Strongly Convex Brenier Potentials in Optimal Transport.

The problem of estimating Wasserstein distances in high-dimensional spaces suffers from the curse of dimensionality: Indeed, ones needs an exponential (w.r.t. dimension) number of samples for the distance between the two samples to be comparable to that between the two measures. Therefore, regularizing the optimal transport (OT) problem is crucial when using Wasserstein distances in machine learning. One of the greatest achievements of the OT literature in recent years lies in regularity theory: one can prove under suitable hypothesis that the OT map between two measures is Lipschitz, or, equivalently when studying 2-Wasserstein distances, that the Brenier convex potential (whose gradient yields an optimal map) is a smooth function. We propose in this work to go backwards, to adopt instead regularity as a regularization tool. We propose algorithms working on discrete measures that can recover nearly optimal transport maps that have small distortion, or, equivalently, nearly optimal Brenier potential that are strongly convex and smooth. For univariate measures, we show that computing these potentials is equivalent to solving an isotonic regression problem under Lipschitz and strong monotonicity constraints. For multivariate measures the problem boils down to a non-convex QCQP problem. We show that this QCQP can be lifted a semidefinite program. Most importantly, these potentials and their gradient can be evaluated on the measures themselves, but can more generally be evaluated on any new point by solving each time a QP. Building on these two formulations we propose practical algorithms to estimate and evaluate transport maps, and illustrate their performance statistically as well as visually on a color transfer task.

7.4. Ranking and synchronization from pairwise measurements via SVD.

Given a measurement graph $G = (V, E)$ and an unknown signal $r \in R^n$, we investigate algorithms for recovering r from pairwise measurements of the form $r_i - r_j; i, j$ in E. This problem arises in a variety of applications, such as ranking teams in sports data and time synchronization of distributed networks. Framed in the context of ranking, the task is to recover the ranking of n teams (induced by r) given a small subset of noisy pairwise rank offsets. We propose a simple SVD-based algorithmic pipeline for both the problem of time synchronization and ranking. We provide a detailed theoretical analysis in terms of robustness against both sampling sparsity and noise perturbations with outliers, using results from matrix perturbation and random matrix theory. Our theoretical findings are complemented by a detailed set of numerical experiments on both synthetic and real data, showcasing the competitiveness of our proposed algorithms with other state-of-the-art methods.

7.5. Polyak Steps for Adaptive Fast Gradient Methods.

Accelerated algorithms for minimizing smooth strongly convex functions usually require knowledge of the strong convexity parameter mu. In the case of an unknown mu, current adaptive techniques are based on restart schemes. When the optimal value f^* is known, these strategies recover the accelerated linear convergence bound without additional grid search. In this paper we propose a new approach that has the same bound without any restart, using an online estimation of strong convexity parameter. We show the robustness of the Fast Gradient Method when using a sequence of upper bounds on mu. We also present a good candidate for this estimate sequence and detail consistent empirical results.

7.6. An Accelerated Decentralized Stochastic Proximal Algorithm for Finite Sums.

Modern large-scale finite-sum optimization relies on two key aspects: distribution and stochastic updates. For smooth and strongly convex problems, existing decentralized algorithms are slower than modern accelerated variance-reduced stochastic algorithms when run on a single machine, and are therefore not efficient. Centralized algorithms are fast, but their scaling is limited by global aggregation steps that result in communication bottlenecks. In this work, we propose an efficient Accelerated Decentralized stochastic algorithm for Finite Sums named ADFS, which uses local stochastic proximal updates and randomized pairwise communications between nodes. On n machines, ADFS learns from nm samples in the same time it takes optimal algorithms to learn from m samples on one machine. This scaling holds until a critical network size is reached, which depends on communication delays, on the number of samples m, and on the network topology. We provide a theoretical analysis based on a novel augmented graph approach combined with a precise evaluation of synchronization times and an extension of the accelerated proximal coordinate gradient algorithm to arbitrary sampling. We illustrate the improvement of ADFS over state-of-the-art decentralized approaches with experiments.

7.7. On Lazy Training in Differentiable Programming.

In a series of recent theoretical works, it was shown that strongly overparameterized neural networks trained with gradient-based methods could converge exponentially fast to zero training loss, with their parameters hardly varying. In this work, we show that this “lazy training” phenomenon is not specific to overparameterized neural networks, and is due to a choice of scaling, often implicit, that makes the model behave as its linearization around the initialization, thus yielding a model equivalent to learning with positive-definite kernels. Through a theoretical analysis, we exhibit various situations where this phenomenon arises in non-convex optimization and we provide bounds on the distance between the lazy and linearized optimization paths. Our numerical experiments bring a critical note, as we observe that the performance of commonly used non-linear deep convolutional neural networks in computer vision degrades when trained in the lazy regime. This makes it unlikely that “lazy training” is behind the many successes of neural networks in difficult high dimensional tasks.

7.8. Implicit Regularization of Discrete Gradient Dynamics in Linear Neural Networks.

When optimizing over-parameterized models, such as deep neural networks, a large set of parameters can achieve zero training error. In such cases, the choice of the optimization algorithm and its respective hyper-parameters introduces biases that will lead to convergence to specific minimizers of the objective. Consequently, this choice can be considered as an implicit regularization for the training of over-parametrized models. In this work, we push this idea further by studying the discrete gradient dynamics of the training of a two-layer linear network with the least-squares loss. Using a time rescaling, we show that, with a vanishing initialization and a small enough step size, this dynamics sequentially learns the solutions of a reduced-rank regression with a gradually increasing rank.

7.9. Efficient Primal-Dual Algorithms for Large-Scale Multiclass Classification.

We develop efficient algorithms to train ℓ_1-regularized linear classifiers with large dimensionality d of the feature space, number of classes k, and sample size n. Our focus is on a special class of losses that includes, in particular, the multiclass hinge and logistic losses. Our approach combines several ideas: (i) passing to the equivalent saddle-point problem with a quasi-bilinear objective; (ii) applying stochastic mirror descent with a proper choice of geometry which guarantees a favorable accuracy bound; (iii) devising non-uniform sampling schemes to approximate the matrix products. In particular, for the multiclass hinge loss we propose a sublinear algorithm with iterations performed in $O(d+n+k)$ arithmetic operations.

7.10. Fast and Faster Convergence of SGD for Over-Parameterized Models (and an Accelerated Perceptron).

Modern machine learning focuses on highly expressive models that are able to fit or interpolate the data completely, resulting in zero training loss. For such models, we show that the stochastic gradients of common loss functions satisfy a strong growth condition. Under this condition, we prove that constant step-size stochastic gradient descent (SGD) with Nesterov acceleration matches the convergence rate of the deterministic accelerated method for both convex and strongly-convex functions. We also show that this condition implies that SGD can find a first-order stationary point as efficiently as full gradient descent in non-convex settings. Under interpolation, we further show that all smooth loss functions with a finite-sum structure satisfy a weaker growth condition. Given this weaker condition, we prove that SGD with a constant step-size attains the deterministic convergence rate in both the strongly-convex and convex settings. Under additional assumptions, the above results enable us to prove an $O(1/k^2)$ mistake bound for k iterations of a stochastic perceptron algorithm using the squared-hinge loss. Finally, we validate our theoretical findings with experiments on synthetic and real datasets.

7.11. Globally Convergent Newton Methods for Ill-conditioned Generalized Self-concordant Losses.

In this project, we study large-scale convex optimization algorithms based on the Newton method applied to regularized generalized self-concordant losses, which include logistic regression and softmax regression. We first prove that our new simple scheme based on a sequence of problems with decreasing regularization parameters is provably globally convergent, that this convergence is linear with a constant factor which scales only logarithmically with the condition number. In the parametric setting, we obtain an algorithm with the same scaling than regular first-order methods but with an improved behavior, in particular in ill-conditioned problems. Second, in the non parametric machine learning setting, we provide an explicit algorithm combining the previous scheme with Nyström projection techniques, and prove that it achieves optimal generalization bounds with a time complexity of order $O(nd_{eff}(\lambda))$, a memory complexity of order $O(d_{eff}(\lambda)^2)$ and no dependence on the condition number, generalizing the results known for least-squares regression. Here n is the number of observations and df λ is the associated degrees of freedom. In particular, this is the first large-scale algorithm to solve logistic and softmax regressions in the non-parametric setting with large condition numbers and theoretical guarantees.

7.12. Efficient online learning with kernels for adversarial large scale problems

We are interested in a framework of online learning with kernels for low-dimensional but large-scale and potentially adversarial datasets. We study the computational and theoretical performance of online variations of kernel Ridge regression. Despite its simplicity, the algorithm we study is the first to achieve the optimal regret for a wide range of kernels with a per-round complexity of order n^α with $\alpha < 2$. The algorithm we consider is based on approximating the kernel with the linear span of basis functions. Our contributions is two-fold: 1) For the Gaussian kernel, we propose to build the basis beforehand (independently of the data) through Taylor expansion. For d-dimensional inputs, we provide a (close to) optimal regret of order $O((logn)^{d+1})$ with per-round time complexity and space complexity $O((logn)^{2d})$. This makes the algorithm a suitable choice as soon as $n \geq e^d$ which is likely to happen in a scenario with small dimensional and large-scale dataset; 2) For general kernels with low effective dimension, the basis functions are updated sequentially in a data-adaptive fashion by sampling Nyström points. In this case, our algorithm improves the computational trade-off known for online kernel regression

7.13. Affine Invariant Covariance Estimation for Heavy-Tailed Distributions

In this work we provide an estimator for the covariance matrix of a heavy-tailed multivariate distribution. We prove that the proposed estimator $\widehat{S}$ admits an *affine-invariant* bound of the form

$$(1-\epsilon)S \leq \widehat{S}(1+\epsilon)S$$

in high probability, where S is the unknown covariance matrix, and $\leq$ is the positive semidefinite order on symmetric matrices. The result only requires the existence of fourth-order moments, and allows for $\epsilon = O(\sqrt{k^4 dlog(d/\delta)/n})$ where k^4 is a measure of kurtosis of the distribution, d is the dimensionality of the space, n is the sample size, and $1-\delta$ is the desired confidence level. More generally, we can allow for regularization with level λ, then d gets replaced with the degrees of freedom number. Denoting $cond(S)$ the condition number of S, the computational cost of the novel estimator is $O(d^2 n + d^3 log(cond(S)))$, which is comparable to the cost of the sample covariance estimator in the statistically interesting regime $n \geq d$. We consider applications of our estimator to eigenvalue estimation with relative error, and to ridge regression with heavy-tailed random design.

7.14. Beyond Least-Squares: Fast Rates for Regularized Empirical Risk Minimization through Self-Concordance

We consider learning methods based on the regularization of a convex empirical risk by a squared Hilbertian norm, a setting that includes linear predictors and non-linear predictors through positive-definite kernels. In order to go beyond the generic analysis leading to convergence rates of the excess risk as $O(\sqrt{1/n})$ from n observations, we assume that the individual losses are self-concordant, that is, their third-order derivatives are bounded by their second-order derivatives. This setting includes least-squares, as well as all generalized linear models such as logistic and softmax regression. For this class of losses, we provide a bias-variance decomposition and show that the assumptions commonly made in least-squares regression, such as the source and capacity conditions, can be adapted to obtain fast non-asymptotic rates of convergence by improving the bias terms, the variance terms or both.

7.15. Statistical Estimation of the Poincaré constant and Application to Sampling Multimodal Distributions

Poincaré inequalities are ubiquitous in probability and analysis and have various applications in statistics (concentration of measure, rate of convergence of Markov chains). The Poincaré constant, for which the inequality is tight, is related to the typical convergence rate of diffusions to their equilibrium measure. In this paper, we show both theoretically and experimentally that, given sufficiently many samples of a measure, we can estimate its Poincaré constant. As a by-product of the estimation of the Poincaré constant, we derive an algorithm that captures a low dimensional representation of the data by finding directions which are difficult to sample. These directions are of crucial importance for sampling or in fields like molecular dynamics, where they are called reaction coordinates. Their knowledge can leverage, with a simple conditioning step, computational bottlenecks by using importance sampling techniques.

7.16. Stochastic first-order methods: non-asymptotic and computer-aided analyses via potential functions

We provide a novel computer-assisted technique for systematically analyzing first-order methods for optimization. In contrast with previous works, the approach is particularly suited for handling sublinear convergence rates and stochastic oracles. The technique relies on semidefinite programming and potential functions. It allows simultaneously obtaining worst-case guarantees on the behavior of those algorithms, and assisting in choosing appropriate parameters for tuning their worst-case performances. The technique also benefits from comfortable tightness guarantees, meaning that unsatisfactory results can be improved only by changing the setting. We use the approach for analyzing deterministic and stochastic first-order methods under different assumptions on the nature of the stochastic noise. Among others, we treat unstructured noise with bounded variance, different noise models arising in over-parametrized expectation minimization problems, and randomized block-coordinate descent schemes.

7.17. Optimal Complexity and Certification of Bregman First-Order Methods

We provide a lower bound showing that the O(1/k) convergence rate of the NoLips method (a.k.a. Bregman Gradient) is optimal for the class of functions satisfying the h-smoothness assumption. This assumption, also known as relative smoothness, appeared in the recent developments around the Bregman Gradient method, where acceleration remained an open issue. On the way, we show how to constructively obtain the corresponding worst-case functions by extending the computer-assisted performance estimation framework of Drori and Teboulle (Mathematical Programming, 2014) to Bregman first-order methods, and to handle the classes of differentiable and strictly convex functions.

7.18. Efficient First-order Methods for Convex Minimization: a Constructive Approach

We describe a novel constructive technique for devising efficient first-order methods for a wide range of large-scale convex minimization settings, including smooth, non-smooth, and strongly convex minimization. The technique builds upon a certain variant of the conjugate gradient method to construct a family of methods such that a) all methods in the family share the same worst-case guarantee as the base conjugate gradient method, and b) the family includes a fixed-step first-order method. We demonstrate the effectiveness of the approach by deriving optimal methods for the smooth and non-smooth cases, including new methods that forego knowledge of the problem parameters at the cost of a one-dimensional line search per iteration, and a universal method for the union of these classes that requires a three-dimensional search per iteration. In the strongly convex case, we show how numerical tools can be used to perform the construction, and show that the resulting method offers an improved worst-case bound compared to Nesterov's celebrated fast gradient method.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Microsoft Research: "Structured Large-Scale Machine Learning". Machine learning is now ubiquitous in industry, science, engineering, and personal life. While early successes were obtained by applying off-the-shelf techniques, there are two main challenges faced by machine learning in the "big data" era: structure and scale. The project proposes to explore three axes, from theoretical, algorithmic and practical perspectives: (1) large-scale convex optimization, (2) large-scale combinatorial optimization and (3) sequential decision making for structured data. The project involves two Inria sites (Paris and Grenoble) and four MSR sites (Cambridge, New England, Redmond, New York). Project website: http://www.msr-inria.fr/projects/structured-large-scale-machine-learning/.

8.2. Bilateral Grants with Industry

- Alexandre d'Aspremont, Francis Bach, Martin Jaggi (EPFL): Google Focused award.
- Francis Bach: Gift from Facebook AI Research.
- Alexandre d'Aspremont: fondation AXA, "Mécénat scientifique", optimisation & machine learning.

9. Partnerships and Cooperations

9.1. National Initiatives

Alexandre d'Aspremont: IRIS, PSL "Science des données, données de la science".

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

ERC Sequoia Title: Robust algorithms for learning from modern data

Programm: H2020

Type: ERC

Duration: 2017-2022

Coordinator: Inria

Inria contact: Francis Bach

Abstract: Machine learning is needed and used everywhere, from science to industry, with a growing impact on many disciplines. While first successes were due at least in part to simple supervised learning algorithms used primarily as black boxes on medium-scale problems, modern data pose new challenges. Scalability is an important issue of course: with large amounts of data, many current problems far exceed the capabilities of existing algorithms despite sophisticated computing architectures. But beyond this, the core classical model of supervised machine learning, with the usual assumptions of independent and identically distributed data, or well-defined features, outputs and loss functions, has reached its theoretical and practical limits. Given this new setting, existing optimization-based algorithms are not adapted. The main objective of this project is to push the frontiers of supervised machine learning, in terms of (a) scalability to data with massive numbers of observations, features, and tasks, (b) adaptability to modern computing environments, in particular for parallel and distributed processing, (c) provable adaptivity and robustness to problem and hardware specifications, and (d) robustness to non-convexities inherent in machine learning problems. To achieve the expected breakthroughs, we will design a novel generation of learning algorithms amenable to a tight convergence analysis with realistic assumptions and efficient implementations. They will help transition machine learning algorithms towards the same wide-spread robust use as numerical linear algebra libraries. Outcomes of the research described in this proposal will include algorithms that come with strong convergence guarantees and are well-tested on real-life benchmarks coming from computer vision, bioinformatics, audio processing and natural language processing. For both distributed and non-distributed settings, we will release open-source software, adapted to widely available computing platforms.

9.3. International Research Visitors

9.3.1. Visits of International Scientists

- Sebastian Pokutta from TU & Zuse Institute, Berlin, December 2019.
- Critobal Guzman from Universidad Católica de Chile, July 2019.

- Quentin Berthet from University of Cambridge, from Feb 2019 until Apr 2019.
- Eduard Gorbunov from Moscow Institute of Physics and Technology, Oct 2019.
- Song Mei, from Stanford University, from Sep 2019 until Oct 2019.
- Anant Raj, from M.P.I. Tubingen, from Oct 2019.
- Aadirupa Saha, from Indian Institute of Technology, Bangalore, from Nov 2019

10. Dissemination

10.1. Promoting Scientific Activities

- Adrien Taylor, ICCOPT session organizer: *Performance Estimation of First-Order Methods* (with F. Glineur), *Splitting Methods and Applications (Part I)* (with P. Giselsson and E. Ryu), *Splitting Methods and Applications (Part III)* (with P. Giselsson and E. Ryu).

10.1.1. Scientific Events: Organisation

10.1.1.1. Member of the Organizing Committees

- Alexandre d'Aspremont, co-organizer, Les Houches Workshop on Optimization and Machine Learning, March 2019.

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

- Senior Area Chair, NeurIPS conference 2019 (Francis Bach).

10.1.2.2. Reviewer

Most of the team referees for the major machine learning conferences such as NIPS, AISTATS, ICML.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Alexandre d'Aspremont: SIAM Journal on Optimization, Associate Editor
- Alexandre d'Aspremont: SIAM Journal on the Mathematics of Data Science, Associate Editor
- Alexandre d'Aspremont: Mathematical Programming B, Associate Editor
- Alexandre d'Aspremont: Mathematics of Operations Research, Associate Editor
- Francis Bach: Journal of Machine Learning Research, co-editor-in-chief
- Francis Bach: Information and Inference, Associate Editor
- Francis Bach: Electronic Journal of Statistics, Associate Editor.
- Francis Bach: Mathematical Programming, Associate Editor
- Francis Bach: Foundations of Computational Mathematics, Associate Editor

10.1.3.2. Reviewer - Reviewing Activities

1. Adrien Taylor, reviewer for Automatica.
2. Adrien Taylor, reviewer for SIAM Journal on Numerical Analysis (SINUM).
3. Adrien Taylor, reviewer for SIAM Journal on Optimization (SIOPT).
4. Adrien Taylor, reviewer for Mathematical Programming (MPA).
5. Adrien Taylor, reviewer for International Conference on Machine Learning 2019 (ICML19).

10.1.4. Invited Talks

- Alexandre d'Aspremont, ICCOPT 2019, Berlin, August 2019.

- Alexandre d'Aspremont, CIMI Workshop, Toulouse, October 2019.
- Alexandre d'Aspremont, France-Germany-Swiss Optimization conference, Nice, September 2019.
- Alexandre d'Aspremont, BIRS Workshop, Oaxaca, Oct. 2019.
- Alexandre d'Aspremont, SAMPTA, Bordeaux, July 2019.
- Francis Bach,Optimization workshop, Les Houches, March 2019.
- Francis Bach, Oberwolfach, Germany, May 2019.
- Francis Bach, AI Global summit, Geneva, Switzerland, May 2019.
- Francis Bach, ETH Data science seminar, Zurich, Switzerland, June 2019.
- Francis Bach, ETH Imaging workshop, Zurich, Switzerland, June 2019.
- Francis BACH, ICIAM invited session, Valencia, Spain, July 2019.
- Francis Bach, Workshop on covariance operators, Germany, Berlin, September 2019.
- Francis Bach, GAMM Workshop COMinDS2019, Berlin, Germany, October 2019.
- Francis Bach, DIMS workshop, Leipzig, Germany, November 2019.
- Francis Bach, Conference on Decision and Control, Nice, December 2019.
- Adrien Taylor, SPOT optimization seminar, Toulouse, February 2019.
- Adrien Taylor, Optimization workshop, Les Houches, March 2019.
- Adrien Taylor, CWI Network & Optimization Seminar, Amsterdam, April 2019.
- Adrien Taylor, Summer school on Optimization (MIPT & HSE), Moscow, June 2019.
- Adrien Taylor, ICCOPT 2019, Berlin, August 2019.
- Alessandro Rudi, *Recent developments on kernel methods*, RKM 2019 (Sept. 2019, London, UK)
- Alessandro Rudi, *Data, Learning and Inference meeting*, DALI 2019 (Sept. 2019, San Sebastian, Spain)
- Alessandro Rudi,*32nd European Meeting of Statisticians*, EMS 2019 (July 2019 , Palermo, Italy)
- Alessandro Rudi,*Applied Inverse Problems Conference*, AIP 2019 (July 2019, Grenoble, France)
- Alessandro Rudi, *Imaging and Machine Learning Conference* (April 2019, Paris, France)
- Alessandro Rudi, *Seminar on the Mathematics of Imaging* (March 2019, Paris, France)

10.1.5. Scientific Expertise

- Alexandre d'Aspremont. Conseil scientifique, Vivienne Investissement.

10.1.6. Research Administration

- Alexandre d'Aspremont. Reponsable scientifique, IRIS PSL, "Sciences des données, données de la science".
- Francis Bach, Deputy Scientific Delegate for Inria Paris research center, member of the Inria Evaluation Committee.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Alexandre d'Aspremont, Optimisation Combinatoire et Convexe, avec Zhentao Li, (2015-Present) cours magistraux 30h, Master M1, ENS Paris.

Master: Alexandre d'Aspremont, Optimisation convexe: modélisation, algorithmes et applications cours magistraux 21h (2011-Present), Master M2 MVA, ENS PS.

Summer school: Francis Bach, optimization for machine learning, 12 heures, AIMS Master, Kigali, Rwanda

Summer school: Francis Bach, optimization for machine learning, 4 heures, D3S summer school, Ecole Polytechnique.

Master : Francis Bach, Optimisation et apprentissage statistique, 20h, Master M2 (Mathématiques de l'aléatoire), Université Paris-Sud, France.

Master: Pierre Gaillard, Alessandro Rudi, Introduction to Machine Learning, 52h, L3, ENS, Paris.

10.2.2. Supervision

PhD in progress : Thomas Kerdreux, New Complexity Bounds for Frank Wolfe, 2017, Alexandre d'Aspremont

PhD in progress : Radu - Dragomir Alexandru, Bregman Gradient Methods, 2018, Alexandre d'Aspremont

PhD in progress : Mathieu Barré, Accelerated Polyak Methods, 2018, Alexandre d'Aspremont

PhD in progress : Grégoire Mialon, Sample Selection Methods, 2018, Alexandre d'Aspremont

PhD in progress : Raphaël Berthier, started September 2017, supervised by Francis Bach and Pierre Gaillard.

PhD in progress: Loucas Pillaud-Vivien, supervised by Francis Bach and Alessandro Rudi.

PhD in progress: Alexandre Défossez, supervised by Francis Bach and Léon Bottou (Facebook AI Research).

PhD in progress: Alex Nowak-Vila, supervised by Francis Bach and Alessandro Rudi.

PhD in progress: Ulysse Marteau Ferey, supervised by Francis Bach and Alessandro Rudi.

PhD in progress: Vivien Cabannes, supervised by Francis Bach and Alessandro Rudi.

PhD in progress: Eloise Berthier, supervised by Francis Bach.

PhD in progress: Theo Ryffel, supervised by Francis Bach and David Pointcheval.

PhD in progress: Margaux Brégère, supervised by Pierre Gaillard and Gilles Stoltz (Université Paris-Sud).

PhD in progress: Rémi Jezequel, supervised by Pierre Gaillard and Alessandro Rudi.

PhD defended: Dmitry Babichev, co-advised by Francis Bach and Anatoly Judistky, defended February 22 2019

PhD defended: Tatiana Shpakova, advised by Francis Bach, defended February 21 2019

10.2.3. Juries

- HdR Pierre Weiss, IMT Toulouse, September 2019 (Alexandre d'Aspremont).
- HDR Rémi Flamary, Université de Nice, November 2019 (Francis Bach).

10.3. Popularization

- Participation to "Fête de la Science" (with the Apprenti Illustrateur).

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] D. BABICHEV.*On Efficient Methods for High-dimensional Statistical Estimation*, PSL Research University, February 2019, https://hal.archives-ouvertes.fr/tel-02433016

[2] T. SHPAKOVA.*On Parameter Learning for Perturb-and-MAP Models*, PSL Research University, February 2019, https://hal.archives-ouvertes.fr/tel-02431640

Articles in International Peer-Reviewed Journal

[3] P. ASKENAZY, F. BACH.*IA et emploi : Une menace artificielle*, in "Pouvoirs - Revue française d'études constitutionnelles et politiques", September 2019, vol. 170, p. 1-7, https://hal.archives-ouvertes.fr/hal-02343633

[4] M. BRÉGÈRE, P. GAILLARD, Y. GOUDE, G. STOLTZ.*Target Tracking for Contextual Bandits: Application to Demand Side Management*, in "Proceedings of Machine Learning Research", June 2019, vol. 97, p. 754-763, https://arxiv.org/abs/1901.09532 , https://hal.archives-ouvertes.fr/hal-01994144

[5] Y. DRORI, A. B. TAYLOR.*Efficient First-order Methods for Convex Minimization: a Constructive Approach*, in "Mathematical Programming, Series A", 2019, https://arxiv.org/abs/1803.05676 - Code available at https://github.com/AdrienTaylor/GreedyMethods, forthcoming [*DOI :* 10.1007/S10107-019-01410-2], https://hal.inria.fr/hal-01902048

[6] P. GAILLARD, S. GERCHINOVITZ, M. HUARD, G. STOLTZ.*Uniform regret bounds over R^d for the sequential linear regression problem with the square loss*, in "Proceedings of Machine Learning Research", 2019, vol. 98, p. 404-432, https://arxiv.org/abs/1805.11386 , https://hal.archives-ouvertes.fr/hal-01802004

[7] J.-L. PEYROT, L. DUVAL, F. PAYAN, L. BOUARD, L. CHIZAT, S. SCHNEIDER, M. ANTONINI.*HexaShrink, an exact scalable framework for hexahedral meshes with attributes and discontinuities: multiresolution rendering and storage of geoscience models*, in "Computational Geosciences", August 2019, vol. 23, n$^{\text{o}}$ 4, p. 723-743, https://arxiv.org/abs/1903.07614 [*DOI :* 10.1007/S10596-019-9816-2], https://hal-ifp.archives-ouvertes.fr/hal-01857997

[8] V. ROULET, N. BOUMAL, A. D'ASPREMONT.*Computational complexity versus statistical performance on sparse recovery problems*, in "Information and Inference", January 2019, https://arxiv.org/abs/1506.03295 [*DOI :* 10.1093/IMAIAI/IAY020], https://hal.archives-ouvertes.fr/hal-02340337

International Conferences with Proceedings

[9] A. BIETTI, G. MIALON, D. CHEN, J. MAIRAL.*A Kernel Perspective for Regularizing Deep Neural Networks*, in "ICML 2019 - 36th International Conference on Machine Learning", Long Beach, United States, Proceedings of Machine Learning Research, June 2019, vol. 97, p. 664-674, https://arxiv.org/abs/1810.00363 , https://hal.inria.fr/hal-01884632

[10] R. BOLLAPRAGADA, D. SCIEUR, A. D'ASPREMONT.*Nonlinear Acceleration of Momentum and Primal-Dual Algorithms*, in "AISTATS 2019 - 22nd International Conference on Artificial Intelligence and Statistics", Naha, Japan, The 22nd International Conference on Artificial Intelligence and Statistics,, April 2019, vol. 89, https://arxiv.org/abs/1810.04539 [*DOI :* 10.04539], https://hal.archives-ouvertes.fr/hal-01893921

[11] T. KERDREUX, A. D'ASPREMONT, S. POKUTTA.*Restarting Frank-Wolfe*, in "AISTATS 2019 - 22nd International Conference on Artificial Intelligence and Statistics", Naha, Japan, Proceedings of Machine Learning Research, April 2019, vol. 89, https://arxiv.org/abs/1810.02429 [*DOI :* 10.02429], https://hal.archives-ouvertes.fr/hal-01893922

[12] U. MARTEAU-FEREY, F. BACH, A. RUDI.*Globally Convergent Newton Methods for Ill-conditioned Generalized Self-concordant Losses*, in "NeurIPS 2019 - Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, https://arxiv.org/abs/1907.01771 , https://hal.inria.fr/hal-02169626

[13] T. RYFFEL, E. DUFOUR-SANS, R. GAY, F. BACH, D. POINTCHEVAL.*Partially Encrypted Machine Learning using Functional Encryption*, in "NeurIPS 2019 - Thirty-third Conference on Neural Information Processing Systems", Vancouver, Canada, Advances in Neural Information Processing Systems, December 2019, https://arxiv.org/abs/1905.10214 , https://hal.inria.fr/hal-02357181

[14] A. TAYLOR, F. BACH.*Stochastic first-order methods: non-asymptotic and computer-aided analyses via potential functions*, in "COLT 2019 - Conference on Learning Theory", Phoenix, United States, June 2019, https://arxiv.org/abs/1902.00947 - 12 pages + appendix; code available at https://github.com/AdrienTaylor/Potential-functions-for-first-order-methods, https://hal.inria.fr/hal-02009309

Conferences without Proceedings

[15] P. ABLIN, A. GRAMFORT, J.-F. CARDOSO, F. BACH.*Stochastic algorithms with descent guarantees for ICA*, in "AISTATS 2019 - 22nd International Conference on Artificial Intelligence and Statistics", Naha, Japan, April 2019, https://hal.archives-ouvertes.fr/hal-02372092

[16] L. CHIZAT, E. OYALLON, F. BACH.*On Lazy Training in Differentiable Programming*, in "NeurIPS 2019 - 33rd Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, https://arxiv.org/abs/1812.07956 , https://hal.inria.fr/hal-01945578

[17] A. GENEVAY, L. CHIZAT, F. BACH, M. CUTURI, G. PEYRÉ.*Sample Complexity of Sinkhorn divergences*, in "AISTATS'19 - 22nd International Conference on Artificial Intelligence and Statistics", Okinawa, Japan, K. CHAUDHURI, M. SUGIYAMA (editors), April 2019, https://arxiv.org/abs/1810.02733 [*DOI :* 10.02733], https://hal.archives-ouvertes.fr/hal-02411822

[18] R. M. GOWER, N. LOIZOU, X. QIAN, A. SAILANBAYEV, E. SHULGIN, P. RICHTÁRIK.*SGD: General Analysis and Improved Rates*, in "International Conference on Machine Learning", Los Angeles, United States, June 2019, https://hal.telecom-paristech.fr/hal-02365318

[19] D. M. OSTROVSKII, A. RUDI.*Affine Invariant Covariance Estimation for Heavy-Tailed Distributions*, in "COLT 2019 - 32nd Annual Conference on Learning Theory", Phoenix, United States, June 2019, https://arxiv.org/abs/1902.03086 , https://hal.archives-ouvertes.fr/hal-02011464

[20] A. PODOSINNIKOVA, A. PERRY, A. WEIN, F. BACH, A. D'ASPREMONT, D. SONTAG.*Overcomplete Independent Component Analysis via SDP*, in "AISTATS 2019 - 22nd International Conference on Artificial Intelligence and Statistics", Naha, Japan, April 2019, https://arxiv.org/abs/1901.08334 - Appears in: Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS 2019). 21 pages, https://hal.archives-ouvertes.fr/hal-02340366

Other Publications

[21] A. ASKARI, A. D'ASPREMONT, L. E. GHAOUI.*Naive Feature Selection: Sparsity in Naive Bayes*, October 2019, https://arxiv.org/abs/1905.09884 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02340374

[22] D. BABICHEV, D. M. OSTROVSKII, F. BACH.*Efficient Primal-Dual Algorithms for Large-Scale Multi-class Classification*, February 2019, https://arxiv.org/abs/1902.03755 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02012569

[23] F. BACH.*Max-Plus Matching Pursuit for Deterministic Markov Decision Processes*, June 2019, https://arxiv.org/abs/1906.08524 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02155865

[24] M. BARRÉ, A. D'ASPREMONT.*Polyak Steps for Adaptive Fast Gradient Method*, October 2019, https://arxiv.org/abs/1906.03056 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02340373

[25] R. BERTHIER, F. BACH, P. GAILLARD.*Accelerated Gossip in Networks of Given Dimension using Jacobi Polynomial Iterations*, February 2019, https://arxiv.org/abs/1805.08531 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01797016

[26] A. D'ASPREMONT, M. CUCURINGU, H. TYAGI.*Ranking and synchronization from pairwise measurements via SVD*, October 2019, https://arxiv.org/abs/1906.02746 - 42 pages, 9 figures, https://hal.archives-ouvertes.fr/hal-02340372

[27] R.-A. DRAGOMIR, A. D'ASPREMONT, J. BOLTE.*Quartic First-Order Methods for Low Rank Minimization*, October 2019, https://arxiv.org/abs/1901.10791 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02340369

[28] R.-A. DRAGOMIR, A. TAYLOR, A. D'ASPREMONT, J. BOLTE.*Optimal Complexity and Certification of Bregman First-Order Methods*, November 2019, https://arxiv.org/abs/1911.08510 - working paper or preprint, https://hal.inria.fr/hal-02384167

[29] A. DÉFOSSEZ, N. USUNIER, L. BOTTOU, F. BACH.*Demucs: Deep Extractor for Music Sources with extra unlabeled data remixed*, September 2019, https://arxiv.org/abs/1909.01174 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02277338

[30] A. DÉFOSSEZ, N. USUNIER, L. BOTTOU, F. BACH.*Music Source Separation in the Waveform Domain*, November 2019, https://arxiv.org/abs/1911.13254 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02379796

[31] H. HENDRIKX, F. BACH, L. MASSOULIÉ.*An Accelerated Decentralized Stochastic Proximal Algorithm for Finite Sums*, September 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02280763

[32] R. JÉZÉQUEL, P. GAILLARD, A. RUDI.*Efficient online learning with kernels for adversarial large scale problems*, May 2019, https://arxiv.org/abs/1902.09917 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02019402

[33] U. MARTEAU-FEREY, D. M. OSTROVSKII, F. BACH, A. RUDI.*Beyond Least-Squares: Fast Rates for Regularized Empirical Risk Minimization through Self-Concordance*, June 2019, https://arxiv.org/abs/1902.03046 - working paper or preprint, https://hal.inria.fr/hal-02011895

[34] G. MIALON, A. D'ASPREMONT, J. MAIRAL.*Screening Data Points in Empirical Risk Minimization via Ellipsoidal Regions and Safe Loss Functions*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02395624

[35] F.-P. PATY, A. D'ASPREMONT, M. CUTURI.*Regularity as Regularization: Smooth and Strongly Convex Brenier Potentials in Optimal Transport*, October 2019, https://arxiv.org/abs/1905.10812 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02340371

[36] L. PILLAUD-VIVIEN, F. BACH, T. LELIÈVRE, A. RUDI, G. STOLTZ.*Statistical Estimation of the Poincaré constant and Application to Sampling Multimodal Distributions*, November 2019, https://arxiv.org/abs/1910.14564 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02327453

[37] T.-H. VU, A. OSOKIN, I. LAPTEV.*Tube-CNN: Modeling temporal evolution of appearance for object detection in video*, January 2019, https://arxiv.org/abs/1812.02619 - 13 pages, 8 figures, technical report, https://hal.archives-ouvertes.fr/hal-01980339

Project-Team VALDA

Value from Data

IN COLLABORATION WITH: Département d'Informatique de l'Ecole Normale Supérieure

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Paris

RESEARCH CENTER
Paris

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team VALDA

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

Keywords:

Computer Science and Digital Science:

A3.1. - Data
A3.1.1. - Modeling, representation
A3.1.2. - Data management, quering and storage
A3.1.3. - Distributed data
A3.1.4. - Uncertain data
A3.1.5. - Control access, privacy
A3.1.6. - Query optimization
A3.1.7. - Open data
A3.1.8. - Big data (production, storage, transfer)
A3.1.9. - Database
A3.1.10. - Heterogeneous data
A3.1.11. - Structured data
A3.2. - Knowledge
A3.2.1. - Knowledge bases
A3.2.2. - Knowledge extraction, cleaning
A3.2.3. - Inference
A3.2.4. - Semantic Web
A3.2.5. - Ontologies
A3.2.6. - Linked data
A3.3.2. - Data mining
A3.4.3. - Reinforcement learning
A3.4.5. - Bayesian methods
A3.5.1. - Analysis of large graphs
A4.7. - Access control
A7.2. - Logic in Computer Science
A7.3. - Calculability and computability
A9.1. - Knowledge
A9.8. - Reasoning

Other Research Topics and Application Domains:

B6.3.1. - Web
B6.3.4. - Social Networks
B6.5. - Information systems
B9.5.6. - Data science
B9.6.5. - Sociology
B9.6.10. - Digital humanities
B9.7.2. - Open data
B9.9. - Ethics
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

Serge Abiteboul [Inria, Emeritus, from Oct 2019; also ARCEP, Board Member, HDR]
Camille Bourgaux [CNRS, Researcher]
Olivier Cappé [CNRS, Senior Researcher, HDR]
Luc Segoufin [Inria, Senior Researcher, HDR]
Michaël Thomazo [Inria, Researcher]

Faculty Members

Pierre Senellart [Team leader, École normale supérieure, Professor, HDR]
Leonid Libkin [École normale supérieure, Professor, from Sep 2019]
Silviu Maniu [Univ Paris-Saclay, Associate Professor, from Sep 2019]

PhD Students

Juliette Achddou [Numberly, PhD Student, from Mar 2019]
Julien Grange [Université Paris Diderot, PhD Student]
Karima Rafes [BorderCloud, PhD Student, until Feb 2019]
Yann Ramusat [École normale supérieure, PhD Student]
Yoan Russac [École normale supérieure, PhD Student]

Post-Doctoral Fellows

Ashish Deepak Dandekar [École normale supérieure, Post-Doctoral Fellow, from Jul 2019]
Nathan Grosshans [École normale supérieure, Post-Doctoral Fellow]

Visiting Scientist

Victor Vianu [UCSD, Université Paris Diderot, & École normale supérieure, from Jun 2018]

Administrative Assistant

Meriem Guemair [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Objectives

> Valda's focus is on both *foundational and systems aspects of* complex *data management*, especially *human-centric data*. The data we are interested in is typically heterogeneous, massively distributed, rapidly evolving, intensional, and often subjective, possibly erroneous, imprecise, incomplete. In this setting, Valda is in particular concerned with the optimization of complex resources such as computer time and space, communication, monetary, and privacy budgets. The goal is to extract *value from data*, beyond simple query answering.

Data management [37], [46] is now an old, well-established field, for which many scientific results and techniques have been accumulated since the sixties. Originally, most works dealt with static, homogeneous, and precise data. Later, works were devoted to heterogeneous data [35] [38], and possibly distributed [76] but at a small scale.

However, these classical techniques are poorly adapted to handle the new challenges of data management. Consider human-centric data, which is either produced by humans, e.g., emails, chats, recommendations, or produced by systems when dealing with humans, e.g., geolocation, business transactions, results of data analysis. When dealing with such data, and to accomplish any task to extract value from such data, we rapidly encounter the following facets:

- *Heterogeneity*: data may come in many different structures such as unstructured text, graphs, data streams, complex aggregates, etc., using many different schemas or ontologies.
- *Massive distribution*: data may come from a large number of autonomous sources distributed over the web, with complex access patterns.
- *Rapid evolution*: many sources may be producing data in real time, even if little of it is perhaps relevant to the specific application. Typically, recent data is of particular interest and changes have to be monitored.
- *Intensionality*[0]: in a classical database, all the data is available. In modern applications, the data is more and more available only intensionally, possibly at some cost, with the difficulty to discover which source can contribute towards a particular goal, and this with some uncertainty.
- *Confidentiality and security*: some personal data is critical and need to remain confidential. Applications manipulating personal data must take this into account and must be secure against linking.
- *Uncertainty*: modern data, and in particular human-centric data, typically includes errors, contradictions, imprecision, incompleteness, which complicates reasoning. Furthermore, the subjective nature of the data, with opinions, sentiments, or biases, also makes reasoning harder since one has, for instance, to consider different agents with distinct, possibly contradicting knowledge.

These problems have already been studied individually and have led to techniques such as *query rewriting* [59] or *distributed query optimization* [64].

Among all these aspects, intensionality is perhaps the one that has least been studied, so we pay particular attention to it. Consider a user's query, taken in a very broad sense: it may be a classical database query, some information retrieval search, a clustering or classification task, or some more advanced knowledge extraction request. Because of intensionality of data, solving such a query is a typically dynamic task: each time new data is obtained, the partial knowledge a system has of the world is revised, and query plans need to be updated, as in adaptive query processing [52] or aggregated search [75]. The system then needs to decide, based on this partial knowledge, of the best next access to perform. This is reminiscent of the central problem of reinforcement learning [73] (train an agent to accomplish a task in a partially known world based on rewards obtained) and of active learning [70] (decide which action to perform next in order to optimize a learning strategy) and we intend to explore this connection further.

Uncertainty of the data interacts with its intensionality: efforts are required to obtain more precise, more complete, sounder results, which yields a trade-off between *processing cost* and *data quality*.

Other aspects, such as heterogeneity and massive distribution, are of major importance as well. A standard data management task, such as query answering, information retrieval, or clustering, may become much more challenging when taking into account the fact that data is not available in a central location, or in a common format. We aim to take these aspects into account, to be able to apply our research to real-world applications.

2.2. The Issues

We intend to tackle hard technical issues such as query answering, data integration, data monitoring, verification of data-centric systems, truth finding, knowledge extraction, data analytics, that take a different flavor in this modern context. In particular, we are interested in designing strategies to *minimize data access cost towards a specific goal, possibly a massive data analysis task*. That cost may be in terms of communication (accessing data in distributed systems, on the Web), of computational resources (when data is produced

[0] We use the spelling *intensional*, as in mathematical logic and philosophy, to describe something that is neither available nor defined in *extension*; *intensional* is derived from *intension*, while *intentional* is derived from *intent*.

by complex tools such as information extraction, machine learning systems, or complex query processing), of monetary budget (paid-for application programming interfaces, crowdsourcing platforms), or of a privacy budget (as in the standard framework of differential privacy).

A number of data management tasks in Valda are inherently intractable. In addition to properly characterizing this intractability in terms of complexity theory, we intend to develop solutions for solving these tasks in practice, based on approximation strategies, randomized algorithms, enumeration algorithms with constant delay, or identification of restricted forms of data instances lowering the complexity of the task.

3. Research Program

3.1. Scientific Foundations

We now detail some of the scientific foundations of our research on complex data management. This is the occasion to review connections between data management, especially on complex data as is the focus of Valda, with related research areas.

3.1.1. Complexity & Logic

Data management has been connected to logic since the advent of the relational model as main representation system for real-world data, and of first-order logic as the logical core of database querying languages [37]. Since these early developments, logic has also been successfully used to capture a large variety of query modes, such as data aggregation [63], recursive queries (Datalog), or querying of XML databases [46]. Logical formalisms facilitate reasoning about the expressiveness of a query language or about its complexity.

The main problem of interest in data management is that of query evaluation, i.e., computing the results of a query over a database. The complexity of this problem has far-reaching consequences. For example, it is because first-order logic is in the AC_0 complexity class that evaluation of SQL queries can be parallelized efficiently. It is usual [74] in data management to distinguish *data complexity*, where the query is considered to be fixed, from *combined complexity*, where both the query and the data are considered to be part of the input. Thus, though conjunctive queries, corresponding to a simple SELECT-FROM-WHERE fragment of SQL, have PTIME data complexity, they are NP-hard in combined complexity. Making this distinction is important, because data is often far larger (up to the order of terabytes) than queries (rarely more than a few hundred bytes). Beyond simple query evaluation, a central question in data management remains that of complexity; tools from algorithm analysis, and complexity theory can be used to pinpoint the tractability frontier of data management tasks.

3.1.2. Automata Theory

Automata theory and formal languages arise as important components of the study of many data management tasks: in temporal databases [36], queries, expressed in temporal logics, can often by compiled to automata; in graph databases [42], queries are naturally given as automata; typical query and schema languages for XML databases such as XPath and XML Schema can be compiled to tree automata [67], or for more complex languages to data tree automata[4]. Another reason of the importance of automata theory, and tree automata in particular, comes from Courcelle's results [50] that show that very expressive queries (from the language of monadic second-order language) can be evaluated as tree automata over *tree decompositions* of the original databases, yielding linear-time algorithms (in data complexity) for a wide variety of applications.

3.1.3. Verification

Complex data management also has connections to verification and static analysis. Besides query evaluation, a central problem in data management is that of deciding whether two queries are *equivalent* [37]. This is critical for query optimization, in order to determine if the rewriting of a query, maybe cheaper to evaluate, will return the same result as the original query. Equivalence can easily be seen to be an instance of the problem of (non-)satisfiability: $q \equiv q'$ if and only if $(q \wedge \neg q') \vee (\neg q \wedge q')$ is not satisfiable. In other words, some aspects of query optimization are static analysis issues. Verification is also a critical part of any database application where it is important to ensure that some property will never (or always) arise [48].

3.1.4. Workflows

The orchestration of distributed activities (under the responsibility of a conductor) and their choreography (when they are fully autonomous) are complex issues that are essential for a wide range of data management applications including notably, e-commerce systems, business processes, health-care and scientific workflows. The difficulty is to guarantee consistency or more generally, quality of service, and to statically verify critical properties of the system. Different approaches to workflow specifications exist: automata-based, logic-based, or predicate-based control of function calls [34].

3.1.5. Probability & Provenance

To deal with the uncertainty attached to data, proper models need to be used (such as attaching *provenance* information to data items and viewing the whole database as being *probabilistic*) and practical methods and systems need to be developed to both reliably estimate the uncertainty in data items and properly manage provenance and uncertainty information throughout a long, complex system.

The simplest model of data uncertainty is the NULLs of SQL databases, also called Codd tables [37]. This representation system is too basic for any complex task, and has the major inconvenient of not being closed under even simple queries or updates. A solution to this has been proposed in the form of *conditional tables* [61] where every tuple is annotated with a Boolean formula over independent Boolean random events. This model has been recognized as foundational and extended in two different directions: to more expressive models of *provenance* than what Boolean functions capture, through a semiring formalism [57], and to a probabilistic formalism by assigning independent probabilities to the Boolean events [58]. These two extensions form the basis of modern provenance and probability management, subsuming in a large way previous works [49], [43]. Research in the past ten years has focused on a better understanding of the tractability of query answering with provenance and probabilistic annotations, in a variety of specializations of this framework [72] [62], [40].

3.1.6. Machine Learning

Statistical machine learning, and its applications to data mining and data analytics, is a major foundation of data management research. A large variety of research areas in complex data management, such as wrapper induction [68], crowdsourcing [41], focused crawling [56], or automatic database tuning [44] critically rely on machine learning techniques, such as classification [60], probabilistic models [55], or reinforcement learning [73].

Machine learning is also a rich source of complex data management problems: thus, the probabilities produced by a conditional random field [65] system result in probabilistic annotations that need to be properly modeled, stored, and queried.

Finally, complex data management also brings new twists to some classical machine learning problems. Consider for instance the area of *active learning* [70], a subfield of machine learning concerned with how to optimally use a (costly) oracle, in an interactive manner, to label training data that will be used to build a learning model, e.g., a classifier. In most of the active learning literature, the cost model is very basic (uniform or fixed-value costs), though some works [69] consider more realistic costs. Also, oracles are usually assumed to be perfect with only a few exceptions [53]. These assumptions usually break when applied to complex data management problems on real-world data, such as crowdsourcing.

3.2. Research Directions

At the beginning of the Valda team, the project was to focus on the following directions:

- foundational aspects of data management, in particular related to query enumeration and reasoning on data, especially regarding security issues;
- implementation of provenance and uncertainty management, real-world applications, other aspects of uncertainty and incompleteness, in particular dynamic;
- development of personal information management systems, integration of machine learning techniques.

We believe the first two directions have been followed in a satisfactory manner. The focus on personal information management has not been kept for various organizational reasons, however, but the third axis of the project is reoriented to more general aspects of Web data management.

New permanent arrivals in the group since its creation have impacted its research directions in the following manner:

- Camille BOURGAUX and Michaël THOMAZO are both specialists of knowledge representation and formal aspects of knowledge bases, which is an expertise that did not exist in the group. They are also both interested in, and have started working on aspects related to connecting their research with database theory, and investigating aspects of uncertainty and incompleteness in their research. This will lead to more work on knowledge representation and symbolic AI aspects, while keeping the focus of Valda on foundations of data management and uncertainty.
- Olivier CAPPÉ is a specialist in statistics and machine learning, in particular multi-armed bandits and reinforcement learning. He is also interested in applications of these learning techniques to data management problems. His arrival in the group therefore complements the expertise of other researchers, and will lead to more work on machine learning issues.
- Leonid LIBKIN is a specialist of database theory, of incomplete data management, and has a line of current research on graph data management. His profile fits very well with the original orientation of the Valda project.

We intend to keep producing leading research on the foundations of data management. Generally speaking, the goal is to investigate the borders of feasibility of various tasks. For instance, what are the assumptions on data that allow for computable problems? When is it not possible at all? When can we hope for efficient query answering, when is it hopeless? This is a problem of theoretical nature which is necessary for understanding the limit of the methods and driving research towards the scenarios where positive results may be obtainable. Only when we have understood the limitation of different methods and have many examples where this is possible, we can hope to design a solid foundation that allowing for a good trade-off between what can be done (needs from the users) and what can be achieved (limitation from the system).

Similarly, we will continue our work, both foundational and practical, on various aspects of provenance and uncertainty management. One overall long-term goal is to reach a full understanding of the interactions between query evaluation or other broader data management tasks and uncertain and annotated data models. We would in particular want to go towards a full classification of tractable (typically polynomial-time) and intractable (typically NP-hard for decision problems, or #P-hard for probability evaluation) tasks, extending and connecting the query-based dichotomy [51] on probabilistic query evaluation with the instance-based one of [39], [40]. Another long-term goal is to consider more dynamic scenarios than what has been considered so far in the uncertain data management literature: when following a workflow, or when interacting with intensional data sources, how to properly represent and update uncertainty annotations that are associated with data. This is critical for many complex data management scenarios where one has to maintain a probabilistic current knowledge of the world, while obtaining new knowledge by posing queries and accessing data sources. Such intensional tasks requires minimizing jointly data uncertainty and cost to data access.

As application area, in addition to the historical focus on personal information management which is now less stressed, we target Web data (Web pages, the semantic Web, social networks, the deep Web, crowdsourcing platforms, etc.).

We aim at keeping a delicate balance between theoretical, foundational research, and systems research, including development and implementation. This is a difficult balance to find, especially since most Valda researchers have a tendency to favor theoretical work, but we believe it is also one of the strengths of the team.

4. Application Domains

4.1. Personal Information Management Systems

We recall that Valda's focus is on human-centric data, i.e., data produced by humans, explicitly or implicitly, or more generally containing information about humans. Quite naturally, we have used as a privileged application area to validate Valda's results that of personal information management systems (Pims for short) [33].

A Pims is a system that allows a user to integrate her own data, e.g., emails and other kinds of messages, calendar, contacts, web search, social network, travel information, work projects, etc. Such information is commonly spread across different services. The goal is to give back to a user the control on her information, allowing her to formulate queries such as "What kind of interaction did I have recently with Alice B.?", "Where were my last ten business trips, and who helped me plan them?". The system has to orchestrate queries to the various services (which means knowing the existence of these services, and how to interact with them), integrate information from them (which means having data models for this information and its representation in the services), e.g., align a GPS location of the user to a business address or place mentioned in an email, or an event in a calendar to some event in a Web search. This information must be accessed intensionally: for instance, costly information extraction tools should only be run on emails which seem relevant, perhaps identified by a less costly cursory analysis (this means, in turn, obtaining a cost model for access to the different services). Impacted people can be found by examining events in the user's calendar and determining who is likely to attend them, perhaps based on email exchanges or former events' participant lists. Of course, uncertainty has to be maintained along the entire process, and provenance information is needed to explain query results to the user (e.g., indicate which meetings and trips are relevant to each person of the output). Knowledge about services, their data models, their costs, need either to be provided by the system designer, or to be automatically learned from interaction with these services, as in [68].

One motivation for that choice is that Pims concentrate many of the problems we intend to investigate: heterogeneity (various sources, each with a different structure), massive distribution (information spread out over the Web, in numerous sources), rapid evolution (new data regularly added), intensionality (knowledge from Wikidata, OpenStreetMap...), confidentiality and security (mostly private data), and uncertainty (very variable quality). Though the data is distributed, its size is relatively modest; other applications may be considered for works focusing on processing data at large scale, which is a potential research direction within Valda, though not our main focus. Another strong motivation for the choice of Pims as application domain is the importance of this application from a societal viewpoint.

A Pims is essentially a system built on top of a user's *personal knowledge base*; such knowledge bases are reminiscent of those found in the Semantic Web, e.g., linked open data. Some issues, such as ontology alignment [71] exist in both scenarios. However, there are some fundamental differences in building personal knowledge bases vs collecting information from the Semantic Web: first, the scope is quite smaller, as one is only interested in knowledge related to a given individual; second, a small proportion of the data is already present in the form of semantic information, most needs to be extracted and annotated through appropriate wrappers and enrichers; third, though the linked open data is meant to be read-only, the only update possible to a user being adding new triples, a personal knowledge base is very much something that a user needs to be able to edit, and propagating updates from the knowledge base to original data sources is a challenge in itself.

4.2. Web Data

The choice of Pims is not exclusive. We also consider other application areas as well. In particular, we have worked in the past and have a strong expertise on Web data [38] in a broad sense: semi-structured, structured, or unstructured content extracted from Web databases [68]; knowledge bases from the Semantic Web [71]; social networks [66]; Web archives and Web crawls [54]; Web applications and deep Web databases [47]; crowdsourcing platforms [41]. We intend to continue using Web data as a natural application domain for the research within Valda when relevant. For instance [45], deep Web databases are a natural application scenario for intensional data management issues: determining if a deep Web database contains some information requires optimizing the number of costly requests to that database.

A common aspect of both personal information and Web data is that their exploitation raises ethical considerations. Thus, a user needs to remain fully in control of the usage that is made of her personal information; a search engine or recommender system that ranks Web content for display to a specific user needs to do so in an unbiased, justifiable, manner. These ethical constraints sometimes forbid some technically solutions that may be technically useful, such as sharing a model learned from the personal data of a user to another user, or using blackboxes to rank query result. We fully intend to consider these ethical considerations within Valda. One of the main goals of a Pims is indeed to empower the user with a full control on the use of this data.

5. Highlights of the Year

5.1. Highlights of the Year

Leonid Libkin, formerly Professor at the University of Edinburgh, was recruited as a senior member of the group in 2019, first (from September to November 2019), with a *Chaire d'Excellence* from FSMP (Fédération des Sciences Mathématiques de Paris), and then as a Professor at ENS.

5.1.1. Awards

Mikaël Monet received the 2019 PhD award of the French database community (BDA) for his PhD prepared within Valda and defended in 2018.

6. New Software and Platforms

6.1. ProvSQL

KEYWORDS: Databases - Provenance - Probability

FUNCTIONAL DESCRIPTION: The goal of the ProvSQL project is to add support for (m-)semiring provenance and uncertainty management to PostgreSQL databases, in the form of a PostgreSQL extension/module/plugin.

NEWS OF THE YEAR: Miscellaneous enhancements and bug fixes. Addition of a tutorial.

- Participants: Pierre Senellart and Yann Ramusat
- Contact: Pierre Senellart
- Publications: Provenance and Probabilities in Relational Databases: From Theory to Practice - ProvSQL: Provenance and Probability Management in PostgreSQL
- URL: https://github.com/PierreSenellart/provsql

6.2. apxproof

KEYWORD: LaTeX

FUNCTIONAL DESCRIPTION: apxproof is a LaTeX package facilitating the typesetting of research articles with proofs in appendix, a common practice in database theory and theoretical computer science in general. The appendix material is written in the LaTeX code along with the main text which it naturally complements, and it is automatically deferred. The package can automatically send proofs to the appendix, can repeat in the appendix the theorem environments stated in the main text, can section the appendix automatically based on the sectioning of the main text, and supports a separate bibliography for the appendix material.

RELEASE FUNCTIONAL DESCRIPTION: Fix formatting of theorems (and proof sketches) to be faithful to the way they are formatted in the base document class (this will change some difference in the appearance of documents typset with earlier versions of apxproof), Configurable mainbodyrepeatedtheorem command to add some styling to repeated theorems, Allow using apxproof without bibunits (e.g., for biblatex compatibility), Restore predefined theorem counters, allowing more robust use of apxproof when the base document class predefines theorems.

NEWS OF THE YEAR: Major 1.2.0 release with a much more faithful rendering of theorems compared to the original base classes, bug fixes, compatibility enhancements (in particular, with respect to the use of biblatex or of fancyvrb).

- Participant: Pierre Senellart
- Contact: Pierre Senellart
- URL: https://github.com/PierreSenellart/apxproof

7. New Results

7.1. Foundations of data management

We obtained a number of results on the foundations of data management, i.e., in database theory.

We worked on **knowledge bases**. In our work a knowledge base consists of an incomplete database together with a set of existential rules. We investigated the problem of query answering: computing the answers that are logically entailed from the knowledge base. This brings to light the fundamental chase tool, and its different variants that have been proposed in the literature. We studied the problem of chase termination, which has applications beyond query answering, and studied its complexity for restricted but useful classes of existential rules [27].

We worked on **data integration**. In our scenario a user can access data sitting in multiple sources by means of queries over a global schema, related to the sources via mappings. Data sources often contain sensitive information, and thus an analysis is needed to verify that a schema satisfies a privacy policy, given as a set of queries whose answers should not be accessible to users. We show that source constraints can have a dramatic impact on disclosure analysis [22]. Another work related to data integration is [16], where we connect the problem of answering queries under limited accesses (e.g., using Web forms) to two foundational issues: containment of Monadic datalog (MDL) programs, and containment problems involving regular tree languages. In particular, we establish a 2EXPTIME lower bound on the problem of containment of a MDL program into a conjunctive query, resolving an open problem from the early 1990s.

We also considered some other foundational topics, further from core database topics. In [18], we establish bounds on the height of maximal finite towers (a *tower* is a sequence of words alternating between two languages in such a way that every word is a subsequence of the following word) between two regular languages. In [17], we present an online $O(\sigma|y|)$-time algorithm for finding approximate occurrences of a word x within a word y, where σ is the alphabet size.

Note that two other works in this theme will be described in the 2020 activity report, as they are published in 2020 conferences [25], [26].

7.2. Uncertainty and provenance of data

We have a strong focus on the uncertainty and provenance in databases. See [20] for a high-level introduction to the area.

In [15], we investigate the use of knowledge compilation, i.e., obtaining compact circuit-based representations of functions, for (Boolean) provenance. Some width parameters of the circuit, such as bounded treewidth or pathwidth, can be leveraged to convert the circuit to structured classes, e.g., deterministic structured NNFs (d-SDNNFs) or OBDDs. In [14], we investigate parameterizations of both database instances and queries that make query evaluation fixed-parameter tractable in combined complexity. We show that clique-frontier-guarded Datalog with stratified negation (CFG-Datalog) enjoys bilinear-time evaluation on structures of bounded treewidth for programs of bounded rule size. Such programs capture in particular conjunctive queries with simplicial decompositions of bounded width, guarded negation fragment queries of bounded CQ-rank, or two-way regular path queries. Our result is shown by translating to alternating two-way automata, whose semantics is defined via cyclic provenance circuits (cycluits) that can be tractably evaluated.

In previous work [39], [40]. we have shown that the only restrictions to database instances that make probabilistic query evaluation tractable for a large class of queries is that of having a small treewidth. In [28], [32], we provide the first large-scale experimental study of treewidth and tree decompositions of real-world database instances (25 datasets from 8 different domains, with sizes ranging from a few thousand to a few million vertices). The goal is to determine which data, if any, has reasonably low treewidth. We also show that, even when treewidth is high, using partial tree decompositions can result in data structures that can assist algorithms.

To conclude on provenance management, in [23], [24], after investigating the complexity of satisfiability and query answering for attributed DL-LiteR ontologies, we propose a new semantics, based on provenance semirings, for integrating provenance information with query answering. Finally, we establish complexity results for satisfiability and query answering under this semantics.

We also consider **other notions of incompleteness**, such as in [13], where we study the complexity of query evaluation for databases whose relations are partially ordered; the problem commonly arises when combining or transforming ordered data from multiple sources. We focus on queries in a useful fragment of SQL, namely positive relational algebra with aggregates, whose bag semantics we extend to the partially ordered setting. Our semantics leads to the study of two main computational problems: the possibility and certainty of query answers. We show that these problems are respectively NP-complete and coNP-complete, but identify tractable cases depending on the query operators or input partial orders.

Finally, we also consider uncertainty through another angle, that of learning in a dynamic environment, using techniques from **reinforcement learning** and the **multi-armed bandit** field.

In [19], we tackle the problem of *influence maximization*: finding influential users, or nodes, in a graph so as to maximize the spread of information. We study a highly generic version of influence maximization, one of optimizing influence campaigns by sequentially selecting “spread seeds” from a set of influencers, a small subset of the node population, under the hypothesis that, in a given campaign, previously activated nodes remain persistently active. We introduce an estimator on the influencers’ remaining potential – the expected number of nodes that can still be reached from a given influencer – and justify its strength to rapidly estimate the desired value, relying on real data gathered from Twitter. We then describe a novel algorithm, GT-UCB, relying on probabilistic upper confidence bounds on the remaining potential.

In [21], we propose a Bayesian information-geometric approach to the exploration-exploitation trade-off in stochastic multi-armed bandits. The uncertainty on reward generation and belief is represented using the manifold of joint distributions of rewards and beliefs. Accumulated information is summarised by the barycentre of joint distributions, the pseudobelief-reward. While the pseudobelief-reward facilitates information accumulation through exploration, another mechanism is needed to increase exploitation by gradually focusing on higher rewards, the pseudobelief-focal-reward. Our resulting algorithm, BelMan, alternates between projection of the pseudobelief-focal-reward onto belief-reward distributions to choose the arm to play, and projection of the updated belief-reward distributions onto the pseudobelief-focal-reward.

In [29], we consider another form of bandits, *linear bandits*, in which the available actions correspond to arbitrary context vectors whose associated rewards follow a non-stationary linear regression model. In this setting, the unknown regression parameter is al- lowed to vary in time. To address this problem, we propose D-LinUCB, a novel optimistic algorithm based on discounted linear regression, where exponential weights are used to smoothly forget the past.

7.3. Web data management

We finally describe research more oriented towards applications.

The PhD of Karima Rafes [11] dealt with **semantic knowledge bases** and their applications to the management of scientific data, through the development of the LinkedWiki platform. Another practical work on semantic knowledge bases is [30], where we show how the edit history of of a knowledge base can help correct constraint violations.

Finally, we investigate **transparency and bias** in data management and artificial intelligence. [12] presents to the data management community the challenges raised by new regulatory frameworks in this area. In [31], we discuss the possibility for artificial intelligence systems to be used in the practice of law.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Juliette Achddou's PhD research is set up as a CIFRE contract and supervision agreement between her employer, the Numberly company, and École normale supérieure.

We are in the process of finalizing a contract with Neo4j, the leading company in the field of graph databases, to work towards the creation of a new standard for graph languages called GQL, building on Neo4j's Cypher query language. On this, we do not start from scratch. In a joint effort between the Neo4j's Cypher group and the Edinburgh database group led by Leonid Libkin, a formal specification of the core querying and update features of Cypher was produced. Starting in 2020, Libkin will chair a working group on the formal semantics of GQL. In addition to Valda, it will involve researchers from Edinburgh, Santiago, Warsaw, and other universities in Paris (Marne-la-Vallee and Paris-Diderot). The project is supported by a grant from Neo4j.

9. Partnerships and Cooperations

9.1. Regional Initiatives

The ISORE project from the Île-de-France region (6k€ grant, DIM RFSI), which started in 2019, was completed in 2020.

Leonid Libkin received funding from FSMP through his *Chaire d'Excellence*, in the fall of 2019.

Pierre Senellart is a recipient of a Chair of the PaRis Artificial Intelligence Research InstitutE, PRAIRIE, sarting in the fall of 2019.

9.2. National Initiatives

9.2.1. ANR

Valda has been part of four ANR projects in 2019:

HEADWORK (2016–2021; 38 k€ for Valda, budget managed by Inria), together with IRISA (Druid, coordinator), Inria Lille (Links & Spirals), and Inria Rennes (Sumo), and two application partners: MNHN (Cesco) and FouleFactory. The topic is workflows for crowdsourcing. See http://headwork.gforge.inria.fr/.

BioQOP (2017–2020; 66 k€ for Valda, budget managed by ENS), with Idemia (coordinator) and GREYC, on the optimization of queries for privacy-aware biometric data management. See http://bioqop.di.ens.fr/.

CQFD (2018–2022; 19 k€ for Valda, budget managed by Inria), with Inria Sophia (GraphIK, coordinator), LaBRI, LIG, Inria Saclay (Cedar), IRISA, Inria Lille (Spirals), and Télécom ParisTech, on complex ontological queries over federated and heterogeneous data. See http://www.lirmm.fr/cqfd/.

QUID (2018–2022; 49 k€ for Valda, budget managed by Inria), LIGM (coordinator), IRIF, and LaBRI, on incomplete and inconsistent data. See https://quid.labri.fr/home.html.

Camille Bourgaux is participating in the AI Chair of Meghyn Bienvenu on *INTENDED (Intelligent handling of imperfect data)* to start in 2020.

9.3. European Initiatives

9.3.1. Collaborations in European Programs, Except FP7 & H2020

A bilateral French–German ANR project, entitled *EQUUS – Efficient Query answering Under UpdateS* was accepted in 2019. It will start in 2020. It involves CNRS (CRIL, CRIStAL, IMJ), Télécom Paris, HU Berlin, and Bayreuth University, in addition to Inria Valda.

9.4. International Initiatives

9.4.1. Informal International Partners

Valda has strong collaborations with the following international groups:

Univ. Edinburgh, United Kingdom: Paolo Guagliardo, Andreas Pieris

Univ. Oxford, United Kingdom: Michael Benedikt, Dan Olteanu, and Georg Gottlob

TU Dresden, Germany: Markus Krötzsch and Sebastian Rudolph

Dortmund University, Germany: Thomas Schwentick

Free Univ. Bozen-Bolzano, Italy: Ana Ozaki

Warsaw University, Poland: Mikołaj Bojańczyk and Szymon Toruńczyk

Tel Aviv University, Israel: Daniel Deutch and Tova Milo

Drexel University, USA: Julia Stoyanovich

Univ. California San Diego, USA: Victor Vianu

Pontifical Catholic University of Chile: Marcelo Arenas, Pablo Barceló

National University of Singapore: Stéphane Bressan

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Victor Vianu, Professor at UC San Diego and former holder of an Inria international chair, spent 6 months within Valda, as a University Paris-Diderot and ENS invited professor.

Thomas Schwentick, Professor at TU Dortmund, spend 1 month within Valda in May–June.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- Camille Bourgaux, organizer of the yearly meeting of the national working group on *Automata, Logic, Games, and Algebra* (ALGA, GDR IM) in 2019
- Leonid Libkin was appointed general chair of PODS
- Luc Segoufin, chair of the steering committee of the conference series *Highlights of Logic, Games and Automata*
- Luc Segoufin and Pierre Senellart, co-organizers of École de Printemps en Informatique Théorique (EPIT) 2019
- Pierre Senellart, co-organizer and chief judge of the ICPC (International Collegiate Programming Contest) Southwestern Europe 2019-2020 competition

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- Leonid Libkin, LICS 2021 (in 2019: constitution of the program committee)

10.1.2.2. Member of the Conference Program Committees

- Camille Bourgaux, IJCAI 2019, AAAI 2020, ECAI 2020
- Olivier Cappé, ALT 2019
- Leonid Libkin, IJCAI 2019, AAAI 2020, ECAI 2020, FOSSACS 2020, ICDT 2020
- Pierre Senellart, SUM 2019, PODS 2019, STACS 2020
- Michaël Thomazo, IJCAI 2019

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Olivier Cappé, *Annals of the Institute of Statistical Mathematics*
- Leonid Libkin, *Acta Informatica*
- Leonid Libkin, *Bulletin of Symbolic Logic*
- Leonid Libkin, *Journal of Applied Logic*
- Leonid Libkin, *SN Computer Science*

10.1.3.2. Reviewer - Reviewing Activities

- Pierre Senellart, *Future Generation Computer Systems*

10.1.4. Invited Talks

- Leonid Libkin, keynote at BDA 2019
- Pierre Senellart, invited talk at Singaporean-French workshop on Artificial Intelligence (SinFra), Singapore
- Pierre Senellart, invited lecture at *Reasoning Web* summer school, Bolzano, Italy

10.1.5. Leadership within the Scientific Community

- Serge Abiteboul is a member of the French Academy of Sciences, of the Academia Europaea, of the scientific council of the Société Informatique de France, and an ACM Fellow.
- Leonid Libkin is a Fellow of the Royal Society of Edinburgh, a member of the Academia Europaea, of the UK Computing research committee, and an ACM Fellow.
- Pierre Senellart is a member of the steering committee of BDA, the French scientific community on data management.

10.1.6. Scientific Expertise

- Pierre Senellart has performed a confidential audit of a company for the French government (*direction interministérielle du numérique et du système d'information et de communication de l'État*)
- Pierre Senellart, ANR

10.1.7. Research Administration

- Olivier Cappé is a scientific deputy director of CNRS division of Information Sciences and Technologies (INS2I).
- Luc Segoufin is a member of the CNHSCT of Inria.
- Pierre Senellart is a member of the board of section 6 of the National Committee for Scientific Research.

- Pierre Senellart is deputy director of the DI ENS laboratory, joint between ENS, CNRS, and Inria.
- Pierre Senellart is a member of the board of the DIM RFSI (Réseau Francilien en Sciences Informatiques).
- Pierre Senellart was a member of the scientific council of PGMO (Programme Gaspard Monge) until mid-2019.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence: Serge Abiteboul, *Scientific reading group*, 15 heqTD, L3, École normale supérieure

Licence: Pierre Senellart, Nathan Grosshans, Michaël Thomazo, *Databases*, 74 heqTD, L3, École normale supérieure

Licence: Pierre Senellart, *Algorithms*, 18 heqTD, L3, École normale supérieure

Licence: Nathan Grosshans, *Formal languages*, 22 heqTD, L3, Télécom ParisTech

Master: Pierre Senellart, *Web data management*, 36 heqTD, M2, MPRI

Pierre Senellart has various teaching responsibilities (L3 internships, M1 projects, M2 administration, entrance competition) at ENS. Nathan Grosshans is the secretary of the entrance competition at ENS for computer science. Most members of the group are also involved in tutoring ENS students, advising them on their curriculum, their internships, etc. They are also occasionally involved with reviewing internship reports, supervising student projects, etc.

10.2.2. Supervision

PhD: Karima Rafes, *Le Linked Data à l'université : la plateforme LinkedWiki*, Université Paris-Saclay, 25 January 2019, Serge Abiteboul & Sarah Cohen-Boulakia

PhD in progess: Juliette Achddou, *Application of reinforcement learning strategies to the context of Real-Time Bidding*, started in September 2018, Olivier Cappé & Aurélien Garivier

PhD in progress: Julien Grange, *Graph properties: order and arithmetic in predicate logics*, started in September 2017, Luc Segoufin

PhD in progress: Yann Ramusat, *Provenance-based routing in probabilistic graphs*, started in September 2018, Silviu Maniu & Pierre Senellart

PhD in progess: Yoan Russac, *Sequential methods for robust decision making*, started in December 2018, Olivier Cappé

10.2.3. Juries

- HdR Paolo Papotti, April 2019, Université de Nice – Sophia-Antipolis, Pierre Senellart (reviewer)
- PhD Ugo Comignani, September 2019, Université Claude Bernard Lyon 1, Pierre Senellart (reviewer)

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

Serge Abiteboul is the president of the strategic committee of the Blaise Pascal foundation for scientific mediation.

Pierre SENELLART is a research fellow within the CERRE (Centre on Regulation in Europe), a European think tank that produces policy papers and organize events about the regulation of network industries. He contributes in particular to reflections on the use of artificial intelligence techniques and on the interoperability of software platforms.

10.3.2. Articles and contents

Serge Abiteboul writes regular columns on popularization of computer science in La Recherche and Le Monde (Économie). He is a founding editor of the *binaire* blog for popularizing computer science. See https://www.lemonde.fr/blog/binaire/.

11. Bibliography

Major publications by the team in recent years

[1] S. ABITEBOUL, P. BOURHIS, V. VIANU.*Explanations and Transparency in Collaborative Workflows*, in "PODS 2018 - 37th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles Of Database Systems", Houston, Texas, United States, June 2018, https://hal.inria.fr/hal-01744978

[2] A. AMARILLI, F. CAPELLI, M. MONET, P. SENELLART.*Connecting Knowledge Compilation Classes and Width Parameters*, in "Theory of Computing Systems", June 2019, https://arxiv.org/abs/1811.02944 [*DOI :* 10.1007/s00224-019-09930-2], https://hal.inria.fr/hal-02163749

[3] C. BOURGAUX, A. OZAKI.*Querying Attributed DL-Lite Ontologies Using Provenance Semirings*, in "Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)", Honolulu, United States, January 2019, https://hal.inria.fr/hal-02109645

[4] F. JACQUEMARD, L. SEGOUFIN, J. DIMINO.*FO2(<, +1, ~) on data trees, data tree automata and branching vector addition systems*, in "Logical Methods in Computer Science", 2016, vol. 12, n° 2, https://doi.org/10.2168/LMCS-12(2:3)2016

[5] P. LAGRÉE, O. CAPPÉ, B. CAUTIS, S. MANIU.*Algorithms for Online Influencer Marketing*, in "ACM Transactions on Knowledge Discovery from Data (TKDD)", January 2019, vol. 13, n° 1, p. 1-30 [*DOI :* 10.1145/3274670], https://hal.inria.fr/hal-01478788

[6] M. LECLÈRE, M.-L. MUGNIER, M. THOMAZO, F. ULLIANA.*A Single Approach to Decide Chase Termination on Linear Existential Rules*, in "ICDT 2019 - International Conference on Database Theory", Lisbonne, Portugal, 2019 [*DOI :* 10.4230/LIPIcs.ICDT.2019.15], https://hal-lirmm.ccsd.cnrs.fr/lirmm-02148200

[7] S. MANIU, R. CHENG, P. SENELLART.*An Indexing Framework for Queries on Probabilistic Graphs*, in "ACM Trans. Datab. Syst", 2017, https://hal.inria.fr/hal-01437580

[8] Y. RUSSAC, C. VERNADE, O. CAPPÉ.*Weighted Linear Bandits for Non-Stationary Environments*, in "NeurIPS 2019 - 33rd Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, https://arxiv.org/abs/1909.09146 , https://hal.inria.fr/hal-02291460

[9] N. SCHWEIKARDT, L. SEGOUFIN, A. VIGNY.*Enumeration for FO Queries over Nowhere Dense Graphs*, in "PODS 2018 - Principles Of Database Systems", Houston, United States, June 2018, https://hal.inria.fr/hal-01895786

[10] P. SENELLART, L. JACHIET, S. MANIU, Y. RAMUSAT.*ProvSQL: Provenance and Probability Management in PostgreSQL*, in "Proceedings of the VLDB Endowment (PVLDB)", August 2018, vol. 11, n° 12, p. 2034-2037 [*DOI :* 10.14778/3229863.3236253], https://hal.inria.fr/hal-01851538

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] K. RAFES.*Linked Data at university : the LinkedWiki platform*, Université Paris-Saclay, January 2019, https://tel.archives-ouvertes.fr/tel-02003672

Articles in International Peer-Reviewed Journal

[12] S. ABITEBOUL, J. STOYANOVICH.*Transparency, Fairness, Data Protection, Neutrality: Data Management Challenges in the Face of New Regulation*, in "Journal of data and information quality", 2019 [*DOI :* 10.1145/3310231], https://hal.inria.fr/hal-02066516

[13] A. AMARILLI, M. L. BA, D. DEUTCH, P. SENELLART.*Computing Possible and Certain Answers over Order-Incomplete Data*, in "Theoretical Computer Science", 2019, vol. 797, p. 42-76, https://arxiv.org/abs/1801.06396 [*DOI :* 10.1016/J.TCS.2019.05.013], https://hal.inria.fr/hal-01891814

[14] A. AMARILLI, P. BOURHIS, M. MONET, P. SENELLART.*Evaluating Datalog via Tree Automata and Cycluits*, in "Theory of Computing Systems", 2019, vol. 63, n^{o} 7, p. 1620-1678, https://arxiv.org/abs/1808.04663 [*DOI :* 10.1007/S00224-018-9901-2], https://hal.inria.fr/hal-01891811

[15] A. AMARILLI, F. CAPELLI, M. MONET, P. SENELLART.*Connecting Knowledge Compilation Classes and Width Parameters*, in "Theory of Computing Systems", June 2019, https://arxiv.org/abs/1811.02944 [*DOI :* 10.1007/S00224-019-09930-2], https://hal.inria.fr/hal-02163749

[16] M. BENEDIKT, P. BOURHIS, G. GOTTLOB, P. SENELLART.*Monadic Datalog, Tree Validity, and Limited Access Containment*, in "ACM Transactions on Computational Logic", October 2019, vol. 21, n^{o} 1, p. 6:1-6:45 [*DOI :* 10.1145/3344514], https://hal.inria.fr/hal-02307999

[17] M. CROCHEMORE, A. HÉLIOU, G. KUCHEROV, L. MOUCHARD, S. PISSIS, Y. RAMUSAT.*Absent words in a sliding window with applications*, in "Information and Computation", September 2019, 104461 [*DOI :* 10.1016/J.IC.2019.104461], https://hal.archives-ouvertes.fr/hal-02414839

[18] S. HOLUB, T. MASOPUST, M. THOMAZO.*On the Height of Towers of Subsequences and Prefixes*, in "Information and Computation", April 2019 [*DOI :* 10.1016/J.IC.2019.01.004], https://hal.inria.fr/hal-02269576

[19] P. LAGRÉE, O. CAPPÉ, B. CAUTIS, S. MANIU.*Algorithms for Online Influencer Marketing*, in "ACM Transactions on Knowledge Discovery from Data (TKDD)", January 2019, vol. 13, n^{o} 1, p. 1-30 [*DOI :* 10.1145/3274670], https://hal.inria.fr/hal-01478788

Invited Conferences

[20] P. SENELLART.*Provenance in Databases: Principles and Applications*, in "RW 2019 : Reasoning Web Summer School", Bolzano, Italy, September 2019, p. 104-109 [*DOI :* 10.1007/978-3-030-31423-1_3], https://hal.inria.fr/hal-02293688

International Conferences with Proceedings

[21] D. BASU, P. SENELLART, S. BRESSAN.*BelMan: An Information-Geometric Approach to Stochastic Bandits*, in "ECML/PKDD - The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases", Würzburg, Germany, September 2019, https://hal.inria.fr/hal-02195539

[22] M. BENEDIKT, P. BOURHIS, L. JACHIET, M. THOMAZO.*Reasoning about disclosure in data integration in the presence of source constraints*, in "IJCAI 2019 - 28th International Joint Conference on Artificial Intelligence", Macao, China, August 2019, https://arxiv.org/abs/1906.00624 , https://hal.inria.fr/hal-02145369

[23] C. BOURGAUX, A. OZAKI.*Querying Attributed DL-Lite Ontologies Using Provenance Semirings (Extended Abstract)*, in "DL 2019 - 32nd International Workshop on Description Logics", Oslo, Norway, June 2019, https://hal.inria.fr/hal-02152064

[24] C. BOURGAUX, A. OZAKI.*Querying Attributed DL-Lite Ontologies Using Provenance Semirings*, in "Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)", Honolulu, United States, January 2019, https://hal.inria.fr/hal-02109645

[25] J. GRANGE, L. SEGOUFIN.*Order-Invariant First-Order Logic over Hollow Trees*, in "CSL 2020 - 28th annual conference of the European Association for Computer Science Logic", Barcelona, Spain, January 2020, vol. 23, p. 1-23 [*DOI :* 10.4230/LIPICS.CSL.2020.23], https://hal.inria.fr/hal-02310749

[26] N. GROSSHANS.*The Power of Programs over Monoids in J*, in "14th International Conference on Language and Automata Theory and Applications (LATA 2020)", Milan, Italy, March 2020, https://arxiv.org/abs/1912.07992 , https://hal.archives-ouvertes.fr/hal-02414771

[27] M. LECLÈRE, M.-L. MUGNIER, M. THOMAZO, F. ULLIANA.*A Single Approach to Decide Chase Termination on Linear Existential Rules*, in "ICDT 2019 - International Conference on Database Theory", Lisbonne, Portugal, 2019 [*DOI :* 10.4230/LIPICS.ICDT.2019.15], https://hal-lirmm.ccsd.cnrs.fr/lirmm-02148200

[28] S. MANIU, P. SENELLART, S. JOG.*An Experimental Study of the Treewidth of Real-World Graph Data*, in "ICDT 2019 – 22nd International Conference on Database Theory", Lisbon, Portugal, March 2019, 18 [*DOI :* 10.4230/LIPICS.ICDT.2019.12], https://hal.inria.fr/hal-02087763

[29] Y. RUSSAC, C. VERNADE, O. CAPPÉ.*Weighted Linear Bandits for Non-Stationary Environments*, in "NeurIPS 2019 - 33rd Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, https://arxiv.org/abs/1909.09146 , https://hal.inria.fr/hal-02291460

[30] T. P. TANON, C. BOURGAUX, F. M. SUCHANEK.*Learning How to Correct a Knowledge Base from the Edit History*, in "World Wide Web Conference", San Francisco, United States, Proceedings of the 2019 World Wide Web Conference (WWW '19), May 2019 [*DOI :* 10.1145/3308558.3313584], https://hal-imt.archives-ouvertes.fr/hal-02066041

Scientific Books (or Scientific Book chapters)

[31] S. ABITEBOUL, F. G'SELL.*Les algorithmes pourraient-ils remplacer les juges ?*, in "Le Big Data et le droit", Thèmes et Commentaire, Dalloz, 2019, https://hal.inria.fr/hal-02304016

Other Publications

[32] S. MANIU, P. SENELLART, S. JOG.*An Experimental Study of the Treewidth of Real-World Graph Data (Extended Version)*, April 2019, https://arxiv.org/abs/1901.06862 - Extended version of an article published in the proceedings of ICDT 2019, https://hal.inria.fr/hal-02087770

References in notes

[33] S. ABITEBOUL, B. ANDRÉ, D. KAPLAN.*Managing your digital life*, in "Commun. ACM", 2015, vol. 58, n^o 5, p. 32-35, http://doi.acm.org/10.1145/2670528

[34] S. ABITEBOUL, P. BOURHIS, V. VIANU.*Comparing workflow specification languages: A matter of views*, in "ACM Trans. Database Syst.", 2012, vol. 37, n^o 2, p. 10:1-10:59, http://doi.acm.org/10.1145/2188349.2188352

[35] S. ABITEBOUL, P. BUNEMAN, D. SUCIU.*Data on the Web: From Relations to Semistructured Data and XML*, Morgan Kaufmann, 1999

[36] S. ABITEBOUL, L. HERR, J. VAN DEN BUSSCHE.*Temporal Versus First-Order Logic to Query Temporal Databases*, in "Proceedings of the Fifteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, June 3-5, 1996, Montreal, Canada", R. HULL (editor), ACM Press, 1996, p. 49-57, http://doi.acm.org/10.1145/237661.237674

[37] S. ABITEBOUL, R. HULL, V. VIANU.*Foundations of Databases*, Addison-Wesley, 1995, http://webdam.inria.fr/Alice/

[38] S. ABITEBOUL, I. MANOLESCU, P. RIGAUX, M. ROUSSET, P. SENELLART.*Web Data Management*, Cambridge University Press, 2011, http://webdam.inria.fr/Jorge

[39] A. AMARILLI, P. BOURHIS, P. SENELLART.*Provenance Circuits for Trees and Treelike Instances*, in "Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part II", 2015, p. 56-68, https://doi.org/10.1007/978-3-662-47666-6_5

[40] A. AMARILLI, P. BOURHIS, P. SENELLART.*Tractable Lineages on Treelike Instances: Limits and Extensions*, in "Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS 2016, San Francisco, CA, USA, June 26 - July 01, 2016", T. MILO, W. TAN (editors), ACM, 2016, p. 355-370, http://doi.acm.org/10.1145/2902251.2902301

[41] Y. AMSTERDAMER, Y. GROSSMAN, T. MILO, P. SENELLART.*CrowdMiner: Mining association rules from the crowd*, in "PVLDB", 2013, vol. 6, n^o 12, p. 1250-1253, http://www.vldb.org/pvldb/vol6/p1250-amsterdamer.pdf

[42] P. B. BAEZA.*Querying graph databases*, in "Proceedings of the 32nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS 2013, New York, NY, USA - June 22 - 27, 2013", R. HULL, W. FAN (editors), ACM, 2013, p. 175-188, http://doi.acm.org/10.1145/2463664.2465216

[43] D. BARBARÁ, H. GARCIA-MOLINA, D. PORTER.*The Management of Probabilistic Data*, in "IEEE Trans. Knowl. Data Eng.", 1992, vol. 4, n^o 5, p. 487-502, https://doi.org/10.1109/69.166990

[44] D. BASU, Q. LIN, W. CHEN, H. T. VO, Z. YUAN, P. SENELLART, S. BRESSAN.*Regularized Cost-Model Oblivious Database Tuning with Reinforcement Learning*, in "T. Large-Scale Data- and Knowledge-Centered Systems", 2016, vol. 28, p. 96-132, https://doi.org/10.1007/978-3-662-53455-7_5

[45] M. BENEDIKT, G. GOTTLOB, P. SENELLART.*Determining relevance of accesses at runtime*, in "Proceedings of the 30th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS 2011, June 12-16, 2011, Athens, Greece", M. LENZERINI, T. SCHWENTICK (editors), ACM, 2011, p. 211-222, http://doi.acm.org/10.1145/1989284.1989309

[46] M. BENEDIKT, P. SENELLART.*Databases*, in "Computer Science, The Hardware, Software and Heart of It", Springer, 2011, p. 169-229, https://doi.org/10.1007/978-1-4614-1168-0_10

[47] M. BIENVENU, D. DEUTCH, D. MARTINENGHI, P. SENELLART, F. M. SUCHANEK.*Dealing with the Deep Web and all its Quirks*, in "Proceedings of the Second International Workshop on Searching and Integrating New Web Data Sources, Istanbul, Turkey, August 31, 2012", M. BRAMBILLA, S. CERI, T. FURCHE, G. GOTTLOB (editors), CEUR Workshop Proceedings, CEUR-WS.org, 2012, vol. 884, p. 21-24, http://ceur-ws.org/Vol-884/VLDS2012_p21_Bienvenu.pdf

[48] M. BOJAŃCZYK, L. SEGOUFIN, S. TORUŃCZYK.*Verification of database-driven systems via amalgamation*, in "Proceedings of the 32nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS 2013, New York, NY, USA - June 22 - 27, 2013", R. HULL, W. FAN (editors), ACM, 2013, p. 63-74, http://doi.acm.org/10.1145/2463664.2465228

[49] P. BUNEMAN, S. KHANNA, W.-C. TAN.*Why and Where: A Characterization of Data Provenance*, in "Database Theory - ICDT 2001, 8th International Conference, London, UK, January 4-6, 2001, Proceedings.", J. VAN DEN BUSSCHE, V. VIANU (editors), Lecture Notes in Computer Science, Springer, 2001, vol. 1973, p. 316-330, https://doi.org/10.1007/3-540-44503-X_20

[50] B. COURCELLE.*The Monadic Second-Order Logic of Graphs. I. Recognizable Sets of Finite Graphs*, in "Inf. Comput.", 1990, vol. 85, n^{o} 1, p. 12-75, https://doi.org/10.1016/0890-5401(90)90043-H

[51] N. N. DALVI, D. SUCIU.*The dichotomy of probabilistic inference for unions of conjunctive queries*, in "J. ACM", 2012, vol. 59, n^{o} 6, p. 30:1-30:87, http://doi.acm.org/10.1145/2395116.2395119

[52] A. DESHPANDE, Z. G. IVES, V. RAMAN.*Adaptive Query Processing*, in "Foundations and Trends in Databases", 2007, vol. 1, n^{o} 1, p. 1-140, https://doi.org/10.1561/1900000001

[53] P. DONMEZ, J. G. CARBONELL.*Proactive learning: cost-sensitive active learning with multiple imperfect oracles*, in "Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM 2008, Napa Valley, California, USA, October 26-30, 2008", J. G. SHANAHAN, S. AMER-YAHIA, I. MANOLESCU, Y. ZHANG, D. A. EVANS, A. KOLCZ, K. CHOI, A. CHOWDHURY (editors), ACM, 2008, p. 619-628, http://doi.acm.org/10.1145/1458082.1458165

[54] M. FAHEEM, P. SENELLART.*Adaptive Web Crawling Through Structure-Based Link Classification*, in "Digital Libraries: Providing Quality Information - 17th International Conference on Asia-Pacific Digital Libraries, ICADL 2015, Seoul, Korea, December 9-12, 2015, Proceedings", R. B. ALLEN, J. HUNTER, M. L. ZENG (editors), Lecture Notes in Computer Science, Springer, 2015, vol. 9469, p. 39-51, https://doi.org/10.1007/978-3-319-27974-9_5

[55] L. GETOOR.*Introduction to statistical relational learning*, MIT Press, 2007

[56] G. GOURITEN, S. MANIU, P. SENELLART.*Scalable, generic, and adaptive systems for focused crawling*, in "25th ACM Conference on Hypertext and Social Media, HT '14, Santiago, Chile, September 1-4, 2014", L. FERRES, G. ROSSI, V. A. F. ALMEIDA, E. HERDER (editors), ACM, 2014, p. 35-45, http://doi.acm.org/10.1145/2631775.2631795

[57] T. J. GREEN, G. KARVOUNARAKIS, V. TANNEN.*Provenance semirings*, in "Proceedings of the Twenty-Sixth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, June 11-13, 2007, Beijing, China", L. LIBKIN (editor), ACM, 2007, p. 31-40, http://doi.acm.org/10.1145/1265530.1265535

[58] T. J. GREEN, V. TANNEN.*Models for Incomplete and Probabilistic Information*, in "IEEE Data Eng. Bull.", 2006, vol. 29, n^o 1, p. 17-24, http://sites.computer.org/debull/A06mar/green.ps

[59] A. Y. HALEVY.*Answering queries using views: A survey*, in "VLDB J.", 2001, vol. 10, n^o 4, p. 270-294, https://doi.org/10.1007/s007780100054

[60] M. A. HEARST, S. T. DUMAIS, E. OSUNA, J. PLATT, B. SCHOLKOPF.*Support vector machines*, in "IEEE Intelligent Systems", 1998, vol. 13, n^o 4, p. 18-28, https://doi.org/10.1109/5254.708428

[61] T. IMIELINSKI, W. LIPSKI JR..*Incomplete Information in Relational Databases*, in "J. ACM", 1984, vol. 31, n^o 4, p. 761-791, http://doi.acm.org/10.1145/1634.1886

[62] B. KIMELFELD, P. SENELLART.*Probabilistic XML: Models and Complexity*, in "Advances in Probabilistic Databases for Uncertain Information Management", Z. MA, L. YAN (editors), Studies in Fuzziness and Soft Computing, Springer, 2013, vol. 304, p. 39-66, https://doi.org/10.1007/978-3-642-37509-5_3

[63] A. C. KLUG.*Equivalence of Relational Algebra and Relational Calculus Query Languages Having Aggregate Functions*, in "J. ACM", 1982, vol. 29, n^o 3, p. 699-717, http://doi.acm.org/10.1145/322326.322332

[64] D. KOSSMANN.*The State of the art in distributed query processing*, in "ACM Comput. Surv.", 2000, vol. 32, n^o 4, p. 422-469, http://doi.acm.org/10.1145/371578.371598

[65] J. D. LAFFERTY, A. MCCALLUM, F. C. N. PEREIRA.*Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data*, in "Proceedings of the Eighteenth International Conference on Machine Learning (ICML 2001), Williams College, Williamstown, MA, USA, June 28 - July 1, 2001", C. E. BRODLEY, A. P. DANYLUK (editors), Morgan Kaufmann, 2001, p. 282-289

[66] S. LEI, S. MANIU, L. MO, R. CHENG, P. SENELLART.*Online Influence Maximization*, in "Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, August 10-13, 2015", 2015, p. 645-654, http://doi.acm.org/10.1145/2783258.2783271

[67] F. NEVEN.*Automata Theory for XML Researchers*, in "SIGMOD Record", 2002, vol. 31, n^o 3, p. 39-46, http://doi.acm.org/10.1145/601858.601869

[68] P. SENELLART, A. MITTAL, D. MUSCHICK, R. GILLERON, M. TOMMASI.*Automatic wrapper induction from hidden-web sources with domain knowledge*, in "10th ACM International Workshop on Web Information

and Data Management (WIDM 2008), Napa Valley, California, USA, October 30, 2008", C. Y. CHAN, N. POLYZOTIS (editors), ACM, 2008, p. 9-16, http://doi.acm.org/10.1145/1458502.1458505

[69] B. SETTLES, M. CRAVEN, L. FRIEDLAND.*Active learning with real annotation costs*, in "NIPS 2008 Workshop on Cost-Sensitive Learning", 2008, http://burrsettles.com/pub/settles.nips08ws.pdf

[70] B. SETTLES.*Active Learning*, Synthesis Lectures on Artificial Intelligence and Machine Learning, Morgan & Claypool Publishers, 2012, https://doi.org/10.2200/S00429ED1V01Y201207AIM018

[71] F. M. SUCHANEK, S. ABITEBOUL, P. SENELLART.*PARIS: Probabilistic Alignment of Relations, Instances, and Schema*, in "PVLDB", 2011, vol. 5, n^{o} 3, p. 157-168, http://www.vldb.org/pvldb/vol5/p157_fabianmsuchanek_vldb2012.pdf

[72] D. SUCIU, D. OLTEANU, C. RÉ, C. KOCH.*Probabilistic Databases*, Synthesis Lectures on Data Management, Morgan & Claypool Publishers, 2011, https://doi.org/10.2200/S00362ED1V01Y201105DTM016

[73] R. S. SUTTON, A. G. BARTO.*Reinforcement learning - an introduction*, Adaptive computation and machine learning, MIT Press, 1998, http://www.worldcat.org/oclc/37293240

[74] M. Y. VARDI.*The Complexity of Relational Query Languages (Extended Abstract)*, in "Proceedings of the 14th Annual ACM Symposium on Theory of Computing, May 5-7, 1982, San Francisco, California, USA", H. R. LEWIS, B. B. SIMONS, W. A. BURKHARD, L. H. LANDWEBER (editors), ACM, 1982, p. 137-146, http://doi.acm.org/10.1145/800070.802186

[75] K. ZHOU, M. LALMAS, T. SAKAI, R. CUMMINS, J. M. JOSE.*On the reliability and intuitiveness of aggregated search metrics*, in "22nd ACM International Conference on Information and Knowledge Management, CIKM'13, San Francisco, CA, USA, October 27 - November 1, 2013", Q. HE, A. IYENGAR, W. NEJDL, J. PEI, R. RASTOGI (editors), ACM, 2013, p. 689-698, http://doi.acm.org/10.1145/2505515.2505691

[76] M. T. ÖZSU, P. VALDURIEZ.*Principles of Distributed Database Systems, Third Edition*, Springer, 2011, https://doi.org/10.1007/978-1-4419-8834-8

Project-Team WHISPER

Well Honed Infrastructure Software for Programming Environments and Runtimes

IN COLLABORATION WITH: Laboratoire d'informatique de Paris 6 (LIP6)

IN PARTNERSHIP WITH:
CNRS

Sorbonne Université (UPMC)

RESEARCH CENTER
Paris

THEME
Distributed Systems and middleware

Table of contents

Project-Team WHISPER

Creation of the Team: 2014 May 15, updated into Project-Team: 2015 December 01

Keywords:

Computer Science and Digital Science:

A1. - Architectures, systems and networks
A1.1.1. - Multicore, Manycore
A1.1.13. - Virtualization
A2.1.6. - Concurrent programming
A2.1.10. - Domain-specific languages
A2.1.11. - Proof languages
A2.2.1. - Static analysis
A2.2.5. - Run-time systems
A2.2.8. - Code generation
A2.3.1. - Embedded systems
A2.3.3. - Real-time systems
A2.4. - Formal method for verification, reliability, certification
A2.4.3. - Proofs
A2.5. - Software engineering
A2.6.1. - Operating systems
A2.6.2. - Middleware
A2.6.3. - Virtual machines

Other Research Topics and Application Domains:

B5. - Industry of the future
B5.2.1. - Road vehicles
B5.2.3. - Aviation
B5.2.4. - Aerospace
B6.1. - Software industry
B6.1.1. - Software engineering
B6.1.2. - Software evolution, maintenance
B6.3.3. - Network Management
B6.5. - Information systems
B6.6. - Embedded systems

1. Team, Visitors, External Collaborators

Research Scientists

Gilles Muller [Team leader, Inria, Senior Researcher, HDR]
Pierre-Évariste Dagand [CNRS, Researcher]
Julia Lawall [Inria, Senior Researcher]

Faculty Members

Bertil Folliot [Univ Pierre et Marie Curie, Professor, HDR]
Maria Virginia Aponte Garcia [CNAM, Associate Professor, from Oct 2019]

Technical Staff

Michael Damien Carver [Univ Pierre et Marie Curie, Engineer, from Mar 2019]

PhD Students

Cédric Courtaud [Thales, PhD Student, until Jun 2019, granted by CIFRE]
Yoann Ghigoff [Orange Labs, PhD Student, from Jun 2019]
Redha Gouicem [Univ Pierre et Marie Curie, PhD Student]
Darius Mercadier [Univ Pierre et Marie Curie, PhD Student]
Pierre Nigron [Inria, PhD Student, from Nov 2019]
Lucas Serrano [Univ Pierre et Marie Curie, PhD Student]

Administrative Assistants

Nelly Maloisel [Inria, Administrative Assistant]
Eugène Kamdem [SU, Assistant]

2. Overall Objectives

2.1. Overall Objectives

The focus of Whisper is on how to develop (new) and improve (existing) infrastructure software. Infrastructure software (also called systems software) is the software that underlies all computing. Such software allows applications to access resources and provides essential services such as memory management, synchronization and inter-process interactions. Starting bottom-up from the hardware, examples include virtual machine hypervisors, operating systems, managed runtime environments, standard libraries, and browsers, which amount to the new operating system layer for Internet applications. For such software, efficiency and correctness are fundamental. Any overhead will impact the performance of all supported applications. Any failure will prevent the supported applications from running correctly. Since computing now pervades our society, with few paper backup solutions, correctness of software at all levels is critical. Formal methods are increasingly being applied to operating systems code in the research community [44], [50], [79]. Still, such efforts require a huge amount of manpower and a high degree of expertise which makes this work difficult to replicate in standard infrastructure-software development.

In terms of methodology, Whisper is at the interface of the domains of operating systems, software engineering and programming languages. Our approach is to combine the study of problems in the development of real-world infrastructure software with concepts in programming language design and implementation, *e.g.*, of domain-specific languages, and knowledge of low-level system behavior. A focus of our work is on providing support for legacy code, while taking the needs and competences of ordinary system developers into account.

We aim at providing solutions that can be easily learned and adopted by system developers in the short term. Such solutions can be tools, such as Coccinelle [2], [9], [10] for transforming C programs, or domain-specific languages such as Devil [8] and Bossa [7] for designing drivers and kernel schedulers. Due to the small size of the team, Whisper mainly targets operating system kernels and runtimes for programming languages. We put an emphasis on achieving measurable improvements in performance and safety in practice, and on feeding these improvements back to the infrastructure software developer community.

3. Research Program

3.1. Scientific Foundations

3.1.1. Program analysis

A fundamental goal of the research in the Whisper team is to elicit and exploit the knowledge found in existing code. To do this in a way that scales to a large code base, systematic methods are needed to infer code properties. We may build on either static [36], [38], [40] or dynamic analysis [58], [60], [65]. Static analysis consists of approximating the behavior of the source code from the source code alone, while dynamic analysis draws conclusions from observations of sample executions, typically of test cases. While dynamic analysis can be more accurate, because it has access to information about actual program behavior, obtaining adequate test cases is difficult. This difficulty is compounded for infrastructure software, where many, often obscure, cases must be handled, and external effects such as timing can have a significant impact. Thus, we expect to primarily use static analyses. Static analyses come in a range of flavors, varying in the extent to which the analysis is *sound*, *i.e.*, the extent to which the results are guaranteed to reflect possible run-time behaviors.

One form of sound static analysis is *abstract interpretation* [38]. In abstract interpretation, atomic terms are interpreted as sound abstractions of their values, and operators are interpreted as functions that soundly manipulate these abstract values. The analysis is then performed by interpreting the program in a compositional manner using these abstracted values and operators. Alternatively, *dataflow analysis* [49] iteratively infers connections between variable definitions and uses, in terms of local transition rules that describe how various kinds of program constructs may impact variable values. Schmidt has explored the relationship between abstract interpretation and dataflow analysis [73]. More recently, more general forms of symbolic execution [36] have emerged as a means of understanding complex code. In symbolic execution, concrete values are used when available, and these are complemented by constraints that are inferred from terms for which only partial information is available. Reasoning about these constraints is then used to prune infeasible paths, and obtain more precise results. A number of works apply symbolic execution to operating systems code [33], [34].

While sound approaches are guaranteed to give correct results, they typically do not scale to the very diverse code bases that are prevalent in infrastructure software. An important insight of Engler et al. [42] was that valuable information could be obtained even when sacrificing soundness, and that sacrificing soundness could make it possible to treat software at the scales of the kernels of the Linux or BSD operating systems. Indeed, for certain types of problems, on certain code bases, that may mostly follow certain coding conventions, it may mostly be safe to *e.g.*, ignore the effects of aliases, assume that variable values are unchanged by calls to unanalyzed functions, etc. Real code has to be understood by developers and thus cannot be too complicated, so such simplifying assumptions are likely to hold in practice. Nevertheless, approaches that sacrifice soundness also require the user to manually validate the results. Still, it is likely to be much more efficient for the user to perform a potentially complex manual analysis in a specific case, rather than to implement all possible required analyses and apply them everywhere in the code base. A refinement of unsound analysis is the CEGAR approach [37], in which a highly approximate analysis is complemented by a sound analysis that checks the individual reports of the approximate analysis, and then any errors in reasoning detected by the sound analysis are used to refine the approximate analysis. The CEGAR approach has been applied effectively on device driver code in tools developed at Microsoft [25]. The environment in which the driver executes, however, is still represented by possibly unsound approximations.

Going further in the direction of sacrificing soundness for scalability, the software engineering community has recently explored a number of approaches to code understanding based on techniques developed in the areas of natural language understanding, data mining, and information retrieval. These approaches view code, as well as other software-reated artifacts, such as documentation and postings on mailing lists, as bags of words structured in various ways. Statistical methods are then used to collect words or phrases that seem to be highly correlated, independently of the semantics of the program constructs that connect them. The obliviousness to program semantics can lead to many false positives (invalid conclusions) [55], but can also highlight trends that are not apparent at the low level of individual program statements. We have previously explored combining such statistical methods with more traditional static analysis in identifying faults in the usage of constants in Linux kernel code [53].

3.1.2. Domain Specific Languages

Writing low-level infrastructure code is tedious and difficult, and verifying it is even more so. To produce non-trivial programs, we could benefit from moving up the abstraction stack to enable both programming and proving as quickly as possible. Domain-specific languages (DSLs), also known as *little languages*, are a means to that end [6] [61].

3.1.2.1. Traditional approach.

Using little languages to aid in software development is a tried-and-trusted technique [75] by which programmers can express high-level ideas about the system at hand and avoid writing large quantities of formulaic C boilerplate.

This approach is typified by the Devil language for hardware access [8]. An OS programmer describes the register set of a hardware device in the high-level Devil language, which is then compiled into a library providing C functions to read and write values from the device registers. In doing so, Devil frees the programmer from having to write extensive bit-manipulation macros or inline functions to map between the values the OS code deals with, and the bit-representation used by the hardware: Devil generates code to do this automatically.

However, DSLs are not restricted to being “stub” compilers from declarative specifications. The Bossa language [7] is a prime example of a DSL involving imperative code (syntactically close to C) while offering a high-level of abstraction. This design of Bossa enables the developer to implement new process scheduling policies at a level of abstraction tailored to the application domain.

Conceptually, a DSL both abstracts away low-level details and justifies the abstraction by its semantics. In principle, it reduces development time by allowing the programmer to focus on high-level abstractions. The programmer needs to write less code, in a language with syntax and type checks adapted to the problem at hand, thus reducing the likelihood of errors.

3.1.2.2. Embedding DSLs.

The idea of a DSL has yet to realize its full potential in the OS community. Indeed, with the notable exception of interface definition languages for remote procedure call (RPC) stubs, most OS code is still written in a low-level language, such as C. Where DSL code generators are used in an OS, they tend to be extremely simple in both syntax and semantics. We conjecture that the effort to implement a given DSL usually outweighs its benefit. We identify several serious obstacles to using DSLs to build a modern OS: specifying what the generated code will look like, evolving the DSL over time, debugging generated code, implementing a bug-free code generator, and testing the DSL compiler.

Filet-o-Fish (FoF) [39] addresses these issues by providing a framework in which to build correct code generators from semantic specifications. This framework is presented as a Haskell library, enabling DSL writers to *embed* their languages within Haskell. DSL compilers built using FoF are quick to write, simple, and compact, but encode rigorous semantics for the generated code. They allow formal proofs of the run-time behavior of generated code, and automated testing of the code generator based on randomized inputs, providing greater test coverage than is usually feasible in a DSL. The use of FoF results in DSL compilers that OS developers can quickly implement and evolve, and that generate provably correct code. FoF has been used to build a number of domain-specific languages used in Barrelfish, [26] an OS for heterogeneous multicore systems developed at ETH Zurich.

The development of an embedded DSL requires a few supporting abstractions in the host programming language. FoF was developed in the purely functional language Haskell, thus benefiting from the type class mechanism for overloading, a flexible parser offering convenient syntactic sugar, and purity enabling a more algebraic approach based on small, composable combinators. Object-oriented languages – such as Smalltalk [43] and its descendant Pharo [30] – or multi-paradigm languages – such as the Scala programming language [63] – also offer a wide range of mechanisms enabling the development of embedded DSLs. Perhaps suprisingly, a low-level imperative language – such as C – can also be extended so as to enable the development of embedded compilers [27].

3.1.2.3. Certifying DSLs.

Whilst automated and interactive software verification tools are progressively being applied to larger and larger programs, we have not yet reached the point where large-scale, legacy software – such as the Linux kernel –

could formally be proved “correct”. DSLs enable a pragmatic approach, by which one could realistically strengthen a large legacy software by first narrowing down its critical component(s) and then focus our verification efforts onto these components.

Dependently-typed languages, such as Coq or Idris, offer an ideal environment for embedding DSLs [35], [31] in a unified framework enabling verification. Dependent types support the type-safe embedding of object languages and Coq’s mixfix notation system enables reasonably idiomatic domain-specific concrete syntax. Coq’s powerful abstraction facilities provide a flexible framework in which to not only implement and verify a range of domain-specific compilers [39], but also to combine them, and reason about their combination.

Working with many DSLs optimizes the “horizontal” compositionality of systems, and favors reuse of building blocks, by contrast with the “vertical” composition of the traditional compiler pipeline, involving a stack of comparatively large intermediate languages that are harder to reuse the higher one goes. The idea of building compilers from reusable building blocks is a common one, of course. But the interface contracts of such blocks tend to be complex, so combinations are hard to get right. We believe that being able to write and verify formal specifications for the pieces will make it possible to know when components can be combined, and should help in designing good interfaces.

Furthermore, the fact that Coq is also a system for formalizing mathematics enables one to establish a close, formal connection between embedded DSLs and non-trivial domain-specific models. The possibility of developing software in a truly “model-driven” way is an exciting one. Following this methodology, we have implemented a certified compiler from regular expressions to x86 machine code [48]. Interestingly, our development crucially relied on an existing Coq formalization, due to Braibant and Pous, [32] of the theory of Kleene algebras.

While these individual experiments seem to converge toward embedding domain-specific languages in rich type theories, further experimental validation is required. Indeed, Barrelfish is an extremely small software compared to the Linux kernel. The challenge lies in scaling this methodology up to large software systems. Doing so calls for a unified platform enabling the development of a myriad of DSLs, supporting code reuse across DSLs as well as providing support for mechanically-verified proofs.

3.2. Research direction: Tools for improving legacy infrastructure software

A cornerstone of our work on legacy infrastructure software is the Coccinelle program matching and transformation tool for C code. Coccinelle has been in continuous development since 2005. Today, Coccinelle is extensively used in the context of Linux kernel development, as well as in the development of other software, such as wine, python, kvm, and systemd. Currently, Coccinelle is a mature software project, and no research is being conducted on Coccinelle itself. Instead, we leverage Coccinelle in other research projects [28], [29], [64], [66], [70], [72], [74], [59], [54], both for code exploration, to better understand at a large scale problems in Linux development, and as an essential component in tools that require program matching and transformation. The continuing development and use of Coccinelle is also a source of visibility in the Linux kernel developer community. We submitted the first patches to the Linux kernel based on Coccinelle in 2007. Since then, over 5500 patches have been accepted into the Linux kernel based on the use of Coccinelle, including around 3000 by over 500 developers from outside our research group.

Our recent work has focused on driver porting. Specifically, we have considered the problem of porting a Linux device driver across versions, particularly backporting, in which a modern driver needs to be used by a client who, typically for reasons of stability, is not able to update their Linux kernel to the most recent version. When multiple drivers need to be backported, they typically need many common changes, suggesting that Coccinelle could be applicable. Using Coccinelle, however, requires writing backporting transformation rules. In order to more fully automate the backporting (or symmetrically forward porting) process, these rules should be generated automatically. We have carried out a preliminary study in this direction with David Lo of Singapore Management University; this work, published at ICSME 2016 [77], is limited to a port from one version to the next one, in the case where the amount of change required is limited to a single line of code. Whisper has been awarded an ANR PRCI grant to collaborate with the group of David Lo on scaling up the rule inference process and proposing a fully automatic porting solution.

3.3. Research direction: developing infrastructure software using Domain Specific Languages

We wish to pursue a *declarative* approach to developing infrastructure software. Indeed, there exists a significant gap between the high-level objectives of these systems and their implementation in low-level, imperative programming languages. To bridge that gap, we propose an approach based on domain-specific languages (DSLs). By abstracting away boilerplate code, DSLs increase the productivity of systems programmers. By providing a more declarative language, DSLs reduce the complexity of code, thus the likelihood of bugs.

Traditionally, systems are built by accretion of several, independent DSLs. For example, one might use Devil [8] to interact with devices, Bossa [7] to implement the scheduling policies. However, much effort is duplicated in implementing the back-ends of the individual DSLs. Our long term goal is to design a unified framework for developing and composing DSLs, following our work on Filet-o-Fish [39]. By providing a single conceptual framework, we hope to amortize the development cost of a myriad of DSLs through a principled approach to reusing and composing them.

Beyond the software engineering aspects, a unified platform brings us closer to the implementation of mechanically-verified DSLs. Using the Coq proof assistant as an x86 macro-assembler [48] is a step in that direction, which belongs to a larger trend of hosting DSLs in dependent type theories [31], [35], [62]. A key benefit of those approaches is to provide – by construction – a formal, mechanized semantics to the DSLs thus developed. This semantics offers a foundation on which to base further verification efforts, whilst allowing interaction with non-verified code. We advocate a methodology based on incremental, piece-wise verification. Whilst building fully-certified systems from the top-down is a worthwhile endeavor [50], we wish to explore a bottom-up approach by which one focuses first and foremost on crucial subsystems and their associated properties.

Our current work on DSLs has two complementary goals: (i) the design of a unified framework for developing and composing DSLs, following our work on Filet-o-Fish, and (ii) the design of domain-specific languages for domains where there is a critical need for code correctness, and corresponding methodologies for proving properties of the run-time behavior of the system.

4. Application Domains

4.1. Linux

Linux is an open-source operating system that is used in settings ranging from embedded systems to supercomputers. The most recent release of the Linux kernel, v4.14, comprises over 16 million lines of code, and supports 30 different families of CPU architectures, around 50 file systems, and thousands of device drivers. Linux is also in a rapid stage of development, with new versions being released roughly every 2.5 months. Recent versions have each incorporated around 13,500 commits, from around 1500 developers. These developers have a wide range of expertise, with some providing hundreds of patches per release, while others have contributed only one. Overall, the Linux kernel is critical software, but software in which the quality of the developed source code is highly variable. These features, combined with the fact that the Linux community is open to contributions and to the use of tools, make the Linux kernel an attractive target for software researchers. Tools that result from research can be directly integrated into the development of real software, where it can have a high, visible impact.

Starting from the work of Engler et al. [41], numerous research tools have been applied to the Linux kernel, typically for finding bugs [40], [57], [67], [76] or for computing software metrics [46], [78]. In our work, we have studied generic C bugs in Linux code [10], bugs in function protocol usage [51], [52], issues related to the processing of bug reports [71] and crash dumps [45], and the problem of backporting [66], [77], illustrating the variety of issues that can be explored on this code base. Unique among research groups working in this area, we have furthermore developed numerous contacts in the Linux developer community. These contacts provide insights into the problems actually faced by developers and serve as a means of validating the practical relevance of our work.

4.2. Device Drivers

Device drivers are essential to modern computing, to provide applications with access, via the operating system, to physical devices such as keyboards, disks, networks, and cameras. Development of new computing paradigms, such as the internet of things, is hampered because device driver development is challenging and error-prone, requiring a high level of expertise in both the targeted OS and the specific device. Furthermore, implementing just one driver is often not sufficient; today's computing landscape is characterized by a number of OSes, *e.g.*, Linux, Windows, MacOS, BSD and many real time OSes, and each is found in a wide range of variants and versions. All of these factors make the development, porting, backporting, and maintenance of device drivers a critical problem for device manufacturers, industry that requires specific devices, and even for ordinary users.

The last fifteen years have seen a number of approaches directed towards easing device driver development. Réveillère, who was supervised by G. Muller, proposes Devil [8], a domain-specific language for describing the low-level interface of a device. Chipounov *et al.* propose RevNic, [34] a template-based approach for porting device drivers from one OS to another. Ryzhyk *et al.* propose Termite, [68], [69] an approach for synthesizing device driver code from a specification of an OS and a device. Currently, these approaches have been successfully applied to only a small number of toy drivers. Indeed, Kadav and Swift [47] observe that these approaches make assumptions that are not satisfied by many drivers; for example, the assumption that a driver involves little computation other than the direct interaction between the OS and the device. At the same time, a number of tools have been developed for finding bugs in driver code. These tools include SDV [25], Coverity [41], CP-Miner, [56] PR-Miner [57], and Coccinelle [9]. These approaches, however, focus on analyzing existing code, and do not provide guidelines on structuring drivers.

In summary, there is still a need for a methodology that first helps the developer understand the software architecture of drivers for commonly used operating systems, and then provides tools for the maintenance of existing drivers.

5. Highlights of the Year

5.1. Highlights of the Year

The Whisper team published one paper at USENIX ATC, one at ASPLOS, one at EuroSys, one at RTSS, and one at PLDI which are five of the major conferences in the scope of our team:

- Effective Static Analysis of Concurrency Use-After-FreeBugs in Linux Device Drivers. USENIX 2019. [13]
- DCNS: Automated Detection of Conservative Non-SleepDefects in the Linux Kernel. ASPLOS 2019. [14]
- When eXtended Para-Virtualization (XPV) meets NUMA. EuroSys 2019. [15]
- Improving Prediction Accuracy of Memory Interferences for Multicore Platforms, RTSS 2019 [18]
- Usuba: high-throughput and constant-time ciphers, by construction. PLDI 2019. [23]

Julia Lawall was co-PC chair of the ASE 2019 research paper track.

6. New Software and Platforms

6.1. Coccinelle

KEYWORDS: Code quality - Evolution - Infrastructure software

FUNCTIONAL DESCRIPTION: Coccinelle is a tool for code search and transformation for C programs. It has been extensively used for bug finding and evolutions in Linux kernel code.

- Participants: Gilles Muller, Julia Lawall, Nicolas Palix, Rene Rydhof Hansen and Thierry Martinez
- Partners: LIP6 - IRILL
- Contact: Julia Lawall
- URL: http://coccinelle.lip6.fr

6.2. Prequel

KEYWORDS: Code search - Git

SCIENTIFIC DESCRIPTION: The commit history of a code base such as the Linux kernel is a gold mine of information on how evolutions should be made, how bugs should be fixed, etc. Nevertheless, the high volume of commits available and the rudimentary filtering tools provided mean that it is often necessary to wade through a lot of irrelevant information before finding example commits that can help with a specific software development problem. To address this issue, we propose Prequel (Patch Query Language), which brings the descriptive power of code matching to the problem of querying a commit history.

FUNCTIONAL DESCRIPTION: Prequel is a tool for searching for complex patterns in the commits of software managed using git.

- Participants: Gilles Muller and Julia Lawall
- Partners: LIP6 - IRILL
- Contact: Julia Lawall
- URL: http://prequel-pql.gforge.inria.fr/

6.3. Usuba

KEYWORDS: Cryptography - Optimizing compiler - Synchronous Language

FUNCTIONAL DESCRIPTION: Usuba is a programming language for specifying block ciphers as well as a bitslicing compiler, for producing high-throughput and secure code.

- Contact: Pierre-Evariste Dagand
- Publication: Usuba, Optimizing & Trustworthy Bitslicing Compiler
- URL: https://github.com/DadaIsCrazy/usuba/

6.4. SchedDisplay

KEYWORDS: Linux kernel - Scheduling - Multicore

FUNCTIONAL DESCRIPTION: SchedDisplay is a visualization tool for SchedLog, a custom ring buffer collecting scheduling events in the Linux kernel. SchedDisplay allows kernel developers to analyze the behavior of the Linux scheduler while running a multicore application.

RELEASE FUNCTIONAL DESCRIPTION: First version released as part of a Demo made during the 10th PLOS workshop: https://ess.cs.uni-osnabrueck.de/workshops/plos/2019/program.php

- Partner: Oracle Labs
- Contact: Gilles Muller
- URL: https://gitlab.inria.fr/gmuller/scheddisplay

7. New Results

7.1. Software engineering for infrastructure software

Data races are often hard to detect in device drivers, due to the non-determinism of concurrent execution. With colleagues from Tsinghua University, we have addressed this issue using dynamic analysis. According to our study of Linux driver patches that fix data races, more than 38% of patches involve a pattern that we call inconsistent lock protection. Specifically, if a variable is accessed within two concurrently executed functions, the sets of locks held aroundeach access are disjoint, at least one of the locksets is non-empty, and at least one of the involved accesses is a write, then a datarace may occur. In a paper published at SANER 2019 [17], we present a runtime analysis approach, named DILP, to detect data races caused by inconsistent lock protection in device drivers. By monitoring driver execution, DILP collects the information about runtime variable accesses and executed functions. Then after driver execution, DILP analyzes the collected information to detect and report data races caused by inconsistent lock protection. We evaluate DILP on 12 device drivers in Linux 4.16.9, and find 25 real data races.

For waiting, the Linux kernel offers both sleep-able and non-sleep operations. However, only non-sleep operations can be used in atomic context. Detecting the possibility of execution in atomic context requires a complete inter-procedural flow analysis, often involving function pointers. Developers may thus conservatively use non-sleep operations even outside of atomic context, which may damage system performance, as such operations unproductively monopolize the CPU. Until now, no systematic approach has been proposed to detect such conservative non-sleep (CNS) defects. In a paper published at ASPLOS 2019 [14] with colleagues from Tsinghua University, we propose a practical static approach, named DCNS, to automatically detect conservative non-sleep defects in the Linux kernel. DCNS uses a summary-based analysis to effectively identify the code in atomic context and a novel file-connection-based alias analysis to correctly identify the set of functions referenced by a function pointer. We evaluate DCNS on Linux 4.16, and in total find 1629 defects. We manually check 943 defects whose call paths are not so difficult to follow, and find that 890 are real. We have randomly selected 300 of the real defects and sent them to kernel developers, and 251 have been confirmed.

In Linux device drivers, use-after-free (UAF) bugs can cause system crashes and serious security problems. We have addressed this issue in work with colleagues at Tsinghua University. According to our study of Linux kernel commits, 42% of the driver commits fixing use-after-free bugs involve driver concurrency. We refer to these use-after-free bugs as concurrency use-after-free bugs. Due to the non-determinism of concurrent execution, concurrency use-after-free bugs are often more difficult to reproduce and detect than sequential use-after-free bugs. In a paper published at USENIX ATC 2019 [13], we propose a practical static analysis approach named DCUAF, to effectively detect concurrency use-after-free bugs in Linux device drivers. DCUAF combines a local analysis analyzing the source code of each driver with a global analysis statistically analyzing the local results of all drivers, forming a local-global analysis, to extract the pairs of driver interface functions that may be concurrently executed. Then, with these pairs, DCUAF performs a summary-based lockset analysis to detect concurrency use-after-free bugs. We have evaluated DCUAF on the driver code of Linux 4.19, and found 640 real concurrency use-after-free bugs. We have randomly selected 130 of the real bugs and reported them to Linux kernel developers, and 95 have been confirmed.

Linux kernel stable versions serve the needs of users who value stability of the kernel over new features. The quality of such stableversions depends on the initiative of kernel developers and maintainers to propagate bug fixing patches to the stable versions. Thus, it is desirable to consider to what extent this process can be automated. A previous approach relies on words from commit messages and a small set of manually constructed code features. This approach, however, shows only moderate accuracy. In a tool paper published ICSE 2019 [11], in the context of the ANR-NRF ITrans project with colleagues from Singapore Management University, paper, we investigate whether deep learning can provide a more accurate solution. We propose PatchNet, a hierarchical deep learning-based approach capable of automatically extracting features from commit messages and commit code and usingthem to identify stable patches. PatchNet contains a deep hierarchical structure that mirrors the hierarchical and sequential structure of commit code, making it distinctive from the existing deep learning models on source code. Experiments on 82,403 recent Linux patches confirm the superiority of PatchNet against various state-of-the-art baselines, including the one recently-adopted by Linux kernel maintainers.

Developing software often requires code changes that are widespread and applied to multiple locations. Previously, the Whisper team has addressed this problem with the tool Coccinelle. In a recent experience paper, published at ECOOP 2019 [21], in the context of the ANR-NRF ITrans project with colleagues from Singapore Management University, we have considered the benefits of extending Coccinelle to Java code. There are tools for Java that allow developers to specify patterns for program matching and source-to-source transformation. However, to our knowledge, none allows for transforming code based on its control-flow context. We prototype Coccinelle4J, an extension to Coccinelle, which is a program transformation tool designed for widespread changes in C code, in order to work on Java source code. We adapt Coccinelle to be able to apply scripts written in the Semantic Patch Language (SmPL), a language provided by Coccinelle, to Java source files. As a case study, we demonstrate the utility of Coccinelle4J with the task of API migration. We show 6 semantic patches to migrate from deprecated Android API methods on several open source Android projects. We describe how SmPL can be used to express several API migrations and justify several of our design decisions. This paper was accompanied by a tool demo.

A challenge in designing cooperative distributed systems is to develop feasible and cost-effective mechanisms to foster cooperation among selfish nodes, i.e., nodes that strategically deviate from the intended specification to increase their individual utility. Finding a satisfactory solution to this challenge may be complicated by the intrinsic characteristics of each system, as well as by the particular objectives set by the system designer. In a previous work we addressed this challenge by proposing RACOON, a general and semi-automatic framework for designing selfishness-resilient cooperative systems. RACOON relies on classical game theory and a custom built simulator to predict the impact of a fixed set of selfish behaviours on the designer's objectives. In a paper published in IEEE Transactions on Dependable and Secure Computing [12], we present RACOON++, which extends the previous framework with a declarative model for defining the utility function and the static behaviour of selfish nodes, along with a new model for reasoning on the dynamic interactions of nodes, based on evolutionary game theory. We illustrate the benefits of using RACOON++ by designing three cooperative systems: a peer-to-peer live streaming system, a load balancing protocol, and an anonymous communication system. Extensive experimental results using the state-of-the-art PeerSim simulator verify that the systems designed using RACOON++ achieve both selfishness-resilience and high performance.

7.2. Programming after the end of Moore's law

The end of Moore's law is a wake-up call that resonates across Computer Science at large. We are now firmly in an era of custom hardware design, as witnessed by the diversity of system-on-chip (SoC) and specialized processing units – such as graphics processing units (GPUs), tensor processing unit (TPUs) or programmable network adapters, to name but a few. This trend is justified by the existence of niche application domains (graphic processing, linear algebra, packet processing, etc.) that greatly benefit from specialized hardware. Faced with the imminent explosion of the number of niche applications and niche architectures, we are still grasping for a programming model that would accommodate this diversity.

The Usuba project is an exploratory effort in that direction. We chose a niche application domain (symmetric cryptographic algorithms), a specialized execution platform (Single Instruction Multiple Data, SIMD) processors and we set out to design a programming language faithfully describing our application domain as well as an optimizing compiler efficiently exploiting our target execution platform.

Indeed, cryptographic primitives are subject to diverging imperatives. Functional correctness and auditability pushes for the use of a high-level programming language. Performance and the threat of timing attacks push for directly programming in assembler to exploit (or avoid!) the micro-architectural features of a given machine. In a paper published at PLDI 2019 [23], we have demonstrated that a suitable programming language could reconcile both views and actually improve on the state of the art of both.

USUBA is a dataflow programming language in which block ciphers become so simple as to be "obviously correct" and whose types document and enforce valid parallelization strategies at the granularity of individual bits. Its optimizing compiler, USUBAC, produces high-throughput, constant-time implementations performing on par with hand-tuned reference implementations. The cornerstone of our approach is a systematization and generalization of *bitslicing*, an implementation trick frequently used by cryptographers. We have shown that USUBA can produce code that executes between 5% slower to 22% faster than hand-tuned reference implementations while gracefully scaling across a wide range of architectures and automatically exploiting Single Instruction Multiple Data (SIMD) instructions whenever the cipher's structure allows it.

7.3. Support for multicore machines

The complexity of computer architectures has risen since the early years of the Linux kernel: Simultaneous Multi-Threading (SMT), multicore processing, and frequency scaling with complex algorithms such as Intel Turbo Boost have all become omnipresent. In order to keep up with hardware innovations, the Linux scheduler has been rewritten several times, and many hardware-related heuristics have been added. Despite this, we have shown in a PLOS paper [16] that a fundamental problem was never identified: the POSIX process creation model, i.e., fork/wait, can behave inefficiently on current multicore architectures due to frequency scaling. We investigate this issue through a simple case study: the compilation of the Linux kernel source tree. To do this, we have developed SchedLog, a low-overhead scheduler tracing tool, and SchedDisplay, a scriptable tool to graphically analyze SchedLog's traces efficiently. We implement two solutions to the problem at the scheduler level which improve the speed of compiling part of the Linux kernel by up to 26%, and the whole kernel by up to 10%.

In an Eurosys paper [15], we address the problem of efficiently virtualizing NUMA architectures. The major challenge comes from the fact that the hypervisor regularly reconfigures the placement of a virtual machine (VM) over the NUMA topology. However, neither guest operating systems (OSes) nor system runtime libraries (e.g., Hotspot) are designed to consider NUMA topology changes at runtime, leading end user applications to unpredictable performance. We present eXtended Para-Virtualization (XPV), a new principle to efficiently virtualize a NUMA architecture. XPV consists in revisiting the interface between the hypervisor and the guest OS, and between the guest OS and system runtime libraries (SRL) so that they can dynamically take into account NUMA topology changes. We introduce a methodology for systematically adapting legacy hypervisors, OSes, and SRLs. We have applied our approach with less than 2k line of codes in two legacy hypervisors (Xen and KVM), two legacy guest OSes (Linux and FreeBSD), and three legacy SRLs (Hotspot, TCMalloc, and jemalloc). The evaluation results showed that XPV outperforms all existing solutions by up to 304%.

Memory interferences may introduce important slowdowns in applications running on COTS multi-core processors. They are caused by concurrent accesses to shared hardware resources of the memory system. The induced delays are difficult to predict, making memory interferences a major obstacle to the adoption of COTS multi-core processors in real-time systems. In an RTSS paper[18], we propose an experimental characterization of applications' memory consumption to determine their sensitivity to memory interferences. Thanks to a new set of microbenchmarks, we show the lack of precision of a purely quantitative characterization. To improve accuracy, we define new metrics quantifying qualitative aspects of memory consumption and implement a profiling tool using the VALGRIND framework. In addition, our profiling tool produces high resolution profiles allowing us to clearly distinguish the various phases in applications' behavior. Using our microbenchmarks and our new characterization, we train a state-of-the-art regressor. The validation on applications from the MIBENCH and the PARSEC suites indicates significant gain in prediction accuracy compared to a purely quantitative characterization.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Orange Labs, 2019-2021, 30 000 euros. The purpose of this contract is to design application-specific proxies so as to speed up network services. The PhD of Yoann Ghigoff is supported by a CIFRE fellowship as part of this contract.
- Thales Research, 2016-2019, 45 000 euros. The purpose of this contract is to enable the usage of multicore architectures in avionics systems. The PhD of Cédric Courtaud is supported by a CIFRE fellowship as part of this contract.
- DGA-Inria, 2019-2021, 60 000 euros. The purpose of this PhD grant is to develop a high-performance, certified packet processing system. The PhD of Pierre Nigron is supported by this grant.

8.2. Bilateral Grants with Industry

- Oracle, 2018-2019, 100 000 dollars.

Operating system schedulers are often a performance bottleneck on multicore architectures because in order to scale, schedulers cannot make optimal decisions and instead have to rely on heuristics. Detecting that performance degradation comes from the scheduler level is extremely difficult because the issue has not been recognized until recently, and with traditional profilers, both the application and the scheduler affect the monitored metrics in the same way.

The first objective of this project is to produce a profiler that makes it possible to find out whether a bottleneck during application runtime is caused by the application itself, by suboptimal OS scheduler behavior, or by a combination of the two. It will require understanding, analyzing and classifying performance bottlenecks that are caused by schedulers, and devising ways to detect them and to provide enough information for the user to understand the root cause of the issue. Following this, the second objective of this project is to use the profiler to better understand which kinds of workloads suffer from poor scheduling, and to propose new algorithms, heuristics and/or a new scheduler design that will improve the situation. Finally, the third contribution will be a methodology that makes it possible to track scheduling bottlenecks in a specific workload using the profiler, to understand them, and to fix them either at the application or at the scheduler level. We believe that the combination of these three contributions will make it possible to fully harness the power of multicore architectures for any workload.

As part of this project, we have already identified frequency scaling and the “fork/wait” paradigm as a source of inefficiency in modern multicore machines. These first results were published in the PLOS workshop that is held together with SOSP. The diagnosis of the problem was possible thanks to the *SchedLog and* SchedDisplay tools that we developed as part of this project.

9. Partnerships and Cooperations

9.1. Regional Initiatives

- City of Paris, 2016-2019, 100 000 euros. As part of the “Émergence - young team” program the city of Paris is supporting part of our work on domain-specific languages and trustworthy domain-specific compilers.

9.2. National Initiatives

9.2.1. ANR

ITrans - awarded in 2016, duration 2017 - 2020

Members: LIP6 (Whisper), David Lo (Singapore Management University)

Coordinator: Julia Lawall

Whisper members: Julia Lawall, Gilles Muller, Lucas Serrano, Van-Anh Nguyen

Funding: ANR PRCI, 287,820 euros.

Objectives:

Large, real-world software must continually change, to keep up with evolving requirements, fix bugs, and improve performance, maintainability, and security. This rate of change can pose difficulties for clients, whose code cannot always evolve at the same rate. This project will target the problems of *forward porting*, where one software component has to catch up to a code base with which it needs to interact, and *back porting*, in which it is desired to use a more modern component in a context where it is necessary to continue to use a legacy code base, focusing on the context of Linux device drivers. In this project, we will take a *history-guided source-code transformation-based* approach, which automatically traverses the history of the changes made to a software system, to find where changes in the code to be ported are required, gathers examples of the required changes, and generates change rules to incrementally back port or forward port the code. Our approach will be a success if it is able to automatically back and forward port a large number of drivers for the Linux operating system to various earlier and later versions of the Linux kernel with high accuracy while requiring minimal developer effort. This objective is not achievable by existing techniques.

VeriAmos - awarded in 2018, duration 2018 - 2021

Members: Inria (Antique, Whisper), UGA (Erods)

Coordinator: Xavier Rival

Whisper members: Julia Lawall, Gilles Muller

Funding: ANR, 121,739 euros.

Objectives:

General-purpose Operating Systems, such as Linux, are increasingly used to support high-level functionalities in the safety-critical embedded systems industry with usage in automotive, medical and cyber-physical systems. However, it is well known that general purpose OSes suffer from bugs. In the embedded systems context, bugs may have critical consequences, even affecting human life. Recently, some major advances have been done in verifying OS kernels, mostly employing interactive theorem-proving techniques. These works rely on the formalization of the programming language semantics, and of the implementation of a software component, but require significant human intervention to supply the main proof arguments. The VeriAmos project will attack this problem by building on recent advances in the design of domain-specific languages and static analyzers for systems code. We will investigate whether the restricted expressiveness and the higher level of abstraction provided by the use of a DSL will make it possible to design static analyzers that can statically and fully automatically verify important classes of semantic properties on OS code, while retaining adequate performance of the OS service. As a specific use-case, the project will target I/O scheduling components.

9.3. International Initiatives

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

9.3.1.1. CSG

Title: Proving Concurrent Multi-Core Operating Systems

International Partner (Institution - Laboratory - Researcher):

University of Sydney (Australia) - Willy Zwaenepoel

Start year: 2019

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

The initial topic of this cooperation is the development of proved multicore schedulers. Over the last two years, we have explored a novel approach based on the identification of key scheduling abstractions and the realization of these abstractions as a Domain-Specific Language (DSL), Ipanema. We have introduced a concurrency model that relies on execution of scheduling events in mutual execution locally on a core, but that still permits reading the state of other cores without requiring locks.

In the three next years, we will leverage on our existing results towards the following directions: (i) Better understanding of what should be the best scheduler for a given multicore application, (ii) Proving the correctness of the C code generated from the DSL policy and of the Ipanema abstract machine, (iii) Extend the Ipanema DSL to the domain of I/O request scheduling, (iv) Design of a provable complete concurrent kernel.

9.3.2. Inria International Partners

9.3.2.1. Informal International Partners

Julia Lawall and Gilles Muller collaborate with David Lo and Lingxiao Jiang of Singapore Management University in the context of the ANR-NRF funded project ITrans. This project supports the PhD of Lucas Serrano. In 2019, this collaboration led to an experience paper at ECOOP on a transformation tool (a variant of Coccinelle) for Java [21] and a tool paper at ICSE on using machine learning for identifying bug-fixing patches for the Linux kernel [19]. The latter has been extended to a journal article published in the IEEE Transactions on Software Engineering [11]. Lawall and Serrano spent two weeks visiting Lo and Jiang at Singapore Management University in December 2019.

Julia Lawall collaborates with Jia-Ju Bai at Tsinghua University on bug finding for the Linux kernel. In 2019, this collaboration led to a paper at SANER on detecting data races in device drivers [17], a paper at ISSRE on extending Linux kernel fuzzing to be able to detect bugs in error-handling code [20], a paper at ASPLOS on detection of unnecessary spinning in the Linux kernel [14], a paper at USENIX ATC on detection of use-after free concurrency bugs in the Linux kernel [13]. Bai visited the Whisper team for 2 months starting in January 2019. Lawall visited Bai at Tsinghua University for one week in August.

Michele Martone of the Leibniz Supercomputing Centre in Munich, Germany has been using Coccinelle in an HPC context and giving workshops on Coccinelle in the HPC research engineer community. Martone has contributed some patches to Coccinelle and we keep in touch with him about possible improvements to Coccinelle that may have an impact on its use in the HPC community.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

Jia-Ju Bai visited the Whisper team for 2 months starting in January 2019. During this time, he and Julia Lawall worked on a prototype of an interprocedural program analysis tool for C code.

Victor Miraldo (Ultrech University) visited the Whisper team for 2 weeks in June 2019, where he worked with Pierre-Évariste Dagand on data structures for efficient differencing of data structures.

9.4.1.1. Internships

Pierre-Évariste Dagand has supervised the Master 2 research internship of Pierre Nigron (University Paris Diderot), from April to August 2019, on the topic of “Effectful programs and their proofs in a dependently-typed setting”. Pierre Nigron was awarded a DGA-Inria grant to pursue a PhD under Julia Lawall’s supervision, co-supervised by Pierre-Évariste Dagand.

Pierre-Évariste Dagand has supervised the Bachelor research internship of Quentin Corradi (École Normale Supérieure de Lyon), for 6 weeks starting in June 2019, on the topic of “A Formal Semantics of SIMD Instruction Sets”.

Pierre-Évariste Dagand has supervised a pre-doctoral internship of Rémi Oudin (École Normale Supérieure de Cachan), from April to August 2019, on the topic of “Hardware interfaces for transiently-powered systems”. Rémi Oudin was awarded a “Contrat Doctoral Spécifique pour Normaliens”.

9.4.2. Visits to International Teams

Gilles Muller spent two weeks in November 2019 visiting the University of Sydney as part of our associated team.

Julia Lawall spent one week at Tsinghua University visiting the group of Jia-Ju Bai in August 2019. Julia Lawall and Lucas Serrano spent two weeks at Singapore Management University visiting the group of David Lo and Lingxiao Jiang in December 2019. During the latter visit, Lawall and Serrano also visited National University of Singapore and Lawall also visited Yale-NUS.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

Gilles Muller is the elected representative member of the system community in the steering committee of the COMPAS conference and also the head of the committee.

10.1.1.2. Member of the Organizing Committees

Gilles Muller was one the co-organizers of the 10th PLOS Workshop that was associated with SOSP.

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

Julia Lawall was program chair of the 34th IEEE/ACM International Conference on Automated Software Engineering (ASE 2019) with Darko Marinov (University of Illinois).

10.1.2.2. Member of the Conference Program Committees

- Julia Lawall: ICSE 2020 New Ideas and Emerging Results, 2019 USENIX Annual Technical Conference (USENIX ATC 2019, ERC), Innovations in Software Engineering Conference (ISEC 2020), The 18th Belgium-Netherlands Software Evolution, Workshop (BENEVOL 2019), 45th International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM 2019), ACM SIGPLAN Workshop on Partial Evaluation and Program Manipulation (PEPM)
- Gilles Muller, Usenix ATC 2019, SOSP 2019, Supercomputing - system track 2019
- Pierre-Évariste Dagand: ICFP 2019, PLOS 2019

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Julia Lawall: Science of Computer Programming

10.1.3.2. Reviewer - Reviewing Activities

- Julia Lawall: IEEE Transactions on Software Engineering
- Pierre-Évariste Dagand: Journal of Functional Programming

10.1.4. Invited Talks

- Julia Lawall: Automation in the Maintenance of the Linux Kernel The Coccinelle Experience. Collège de France, February 6.

- Julia Lawall: Coccinelle: 10 Years of Automated Evolution in the Linux Kernel, keynote Linaro Connect conference, September 27.
- Julia Lawall: 10 Years of Automated Evolution in the Linux Kernel, keynote MASCOTS conference, October 23.
- Julia Lawall: Fast and Precise Retrieval of Forward and Back Porting Information for Linux Device Drivers. Lund University, Tsinghua University, University of Illinois, National University of Singapore.
- Julia Lawall: Julia's adventures with Why3. CNAM.
- Lucas Serrano: SPINFER: Semantic patch inference for the Linux kernel. Singapore Management University.
- Gilles Muller: *Research problems in thread scheduling for multicore general purpose operating systems*, WOS Workshop (Rennes), University of Utah, University of New South Wales, University of Sydney, ENS Lyon.
- Pierre-Évariste Dagand: *Programming with dependent types*, Collège de France.

10.1.5. Scientific Expertise

Julia Lawall, Gilles Muller: Discussion with Starburst as part of their survey of sovereign operating systems (March 19, 2019).

10.1.6. Research Administration

- Gilles Muller: Member of a hiring committee for a Professor Position at ENS Lyon.
- Gilles Muller: Elected member of IFIP WG 10.4 (Dependability),
- Bertil Folliot: Elected member of the IFIP WG10.3 working group (Concurrent systems)
- Julia Lawall: Founding member of IFIP WG 2.11 (Program Generation),
- Pierre-Évariste Dagand: Member of a hiring committee for a "Professeur Assistant d'exercice incomplet en Informatique" at École Polytechnique

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Professional Licence: Bertil Folliot, Programmation C, L2, SU, France
- Professional Licence: Bertil Folliot, Lab projects, L2, SU, France
- Licence: Pierre-Évariste Dagand, INF311: Introduction to Programming (40h), L1, École Polytechnique, France
- Licence: Pierre-Évariste Dagand, CSE205: Computer Architecture and Operating Systems (27h), Bachelor, École Polytechnique, France
- Master: Pierre-Évariste Dagand, MPRI 2-4.4: Dependently-typed programming (12h), M2, MPRI, France
- Master: Pierre-Évariste Dagand, INF559: Computer Architecture and Operating Systems (18h), M1, École Polytechnique, France

10.2.2. Supervision

- PhD in progress : Cédric Courtaud, CIFRE Thalès, 2016-2019 (defense January 28 2020), Gilles Muller, Julien Sopéna (Delys).
- PhD in progress : Redha Gouicem, 2016-2020, Gilles Muller, Julien Sopéna (Delys).
- PhD in progress : Yoann Ghigoff, 2019-2022, Gilles Muller, Julien Sopéna (Delys), Kahina Lazri (Orange Labs).

- PhD in progress : Darius Mercadier, 2017-2020, Pierre-Évariste Dagand, Gilles Muller.
- PhD in progress : Lucas Serrano, 2017-2020, Julia Lawall.
- PhD in progress : Pierre Nigron, 2019-2021, Pierre-Évariste Dagand, Julia Lawall.

10.2.3. Juries

- Julia Lawall: PhD jury of Jesper Öqvists (University of Lund, reporter), January 18, 2019.
- Julia Lawall: PhD jury of David Come (Onera, reporter), January 25, 2019.
- Julia Lawall: PhD jury of Jason Lecerf (University of Lille, reporter), November 26, 2019.
- Julia Lawall: HDR jury of Nic Volanschi (University of Bordeaux, president), November 29, 2019.
- Gilles Muller: PhD jury of Damien Carver (Sorbonne University, president), May 17, 2019.
- Pierre-Évariste Dagand: PhD jury of Kenji Maillard (University PSL, examiner), November 25, 2019.

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- Julia Lawall: Chair of the Commission des Emplois Scientifiques, with Mazyar Mirrahimi.

10.3.2. Education

- Julia Lawall: Coccinelle: Practical program transformation for the Linux kernel. Ecole de Jeunes Chercheurs en Programmation.

10.3.3. Interventions

- Julia Lawall: Systèmes d'exploitation. Girls Can Code.

11. Bibliography

Major publications by the team in recent years

[1] T. Bourke, L. Brun, P. Dagand, X. Leroy, M. Pouzet, L. Rieg.*A formally verified compiler for Lustre*, in "PLDI", 2017, p. 586–601

[2] J. Brunel, D. Doligez, R. R. Hansen, J. L. Lawall, G. Muller.*A foundation for flow-based program matching using temporal logic and model checking*, in "POPL", Savannah, GA, USA, ACM, January 2009, p. 114–126

[3] L. Burgy, L. Réveillère, J. L. Lawall, G. Muller.*Zebu: A Language-Based Approach for Network Protocol Message Processing*, in "IEEE Trans. Software Eng.", 2011, vol. 37, n^{o} 4, p. 575-591

[4] P. Dagand, N. Tabareau, É. Tanter.*Partial type equivalences for verified dependent interoperability*, in "ICFP", 2016, p. 298–310

[5] D. Mercadier, P. Dagand.*Usuba: high-throughput and constant-time ciphers, by construction*, in "PLDI", 2019, p. 157–173

[6] G. Muller, C. Consel, R. Marlet, L. P. Barreto, F. Mérillon, L. Réveillère.*Towards Robust OSes for Appliances: A New Approach Based on Domain-specific Languages*, in "Proceedings of the 9th Workshop on ACM SIGOPS European Workshop: Beyond the PC: New Challenges for the Operating System", Kolding, Denmark, 2000, p. 19–24

[7] G. MULLER, J. L. LAWALL, H. DUCHESNE.*A Framework for Simplifying the Development of Kernel Schedulers: Design and Performance Evaluation*, in "HASE - High Assurance Systems Engineering Conference", Heidelberg, Germany, IEEE, October 2005, p. 56–65

[8] F. MÉRILLON, L. RÉVEILLÈRE, C. CONSEL, R. MARLET, G. MULLER.*Devil: An IDL for hardware programming*, in "Proceedings of the Fourth Symposium on Operating Systems Design and Implementation (OSDI)", San Diego, California, USENIX Association, October 2000, p. 17–30

[9] Y. PADIOLEAU, J. L. LAWALL, R. R. HANSEN, G. MULLER.*Documenting and Automating Collateral Evolutions in Linux Device Drivers*, in "EuroSys", Glasgow, Scotland, March 2008, p. 247–260

[10] N. PALIX, G. THOMAS, S. SAHA, C. CALVÈS, J. L. LAWALL, G. MULLER.*Faults in Linux 2.6*, in "ACM Transactions on Computer Systems", June 2014, vol. 32, n^o 2, p. 4:1–4:40

Publications of the year

Articles in International Peer-Reviewed Journal

[11] T. HOANG, J. L. LAWALL, Y. TIAN, R. J. OENTARYO, D. LO.*PatchNet: Hierarchical Deep Learning-Based Stable Patch Identification for the Linux Kernel*, in "IEEE Transactions on Software Engineering", November 2019, 17 [*DOI :* 10.1109/TSE.2019.2952614], https://hal.inria.fr/hal-02373994

[12] G. LENA COTA, S. BEN MOKHTAR, G. GIANINI, E. DAMIANI, J. L. LAWALL, G. MULLER, L. BRUNIE.*RACOON++: A Semi-Automatic Framework for the Selfishness-Aware Design of Cooperative Systems*, in "IEEE Transactions on Dependable and Secure Computing", July 2019, vol. 16, n^o 4, p. 635-650 [*DOI :* 10.1109/TDSC.2017.2706286], https://hal.archives-ouvertes.fr/hal-02196805

International Conferences with Proceedings

[13] J.-J. BAI, J. L. LAWALL, Q.-L. CHEN, S.-M. HU.*Effective Static Analysis of Concurrency Use-After-Free Bugs in Linux Device Drivers Effective Static Analysis of Concurrency Use-After-Free Bugs in Linux Device Drivers*, in "2019 USENIX Annual Technical Conference", Renton, Washington, United States, July 2019, https://hal.inria.fr/hal-02182516

[14] J.-J. BAI, J. L. LAWALL, W. TAN, S.-M. HU.*DCNS: Automated Detection of Conservative Non-Sleep Defects in the Linux Kernel*, in "ASPLOS 2019 - The 24th ACM International Conference on Architectural Support for Programming Languages and Operating Systems", Providence, Rhode Island, United States, ACM, April 2019, p. 287-299 [*DOI :* 10.1145/3297858.3304065], https://hal.inria.fr/hal-02389543

[15] B. BUI, D. MVONDO, B. TEABE, K. JIOKENG, L. WAPET, A. TCHANA, G. THOMAS, D. HAGIMONT, G. MULLER, N. DEPALMA.*When eXtended Para-Virtualization (XPV) meets NUMA*, in "EUROSYS 2019: 14th European Conference on Computer Systems", Dresde, Germany, ACM Press, 2019, 7 [*DOI :* 10.1145/3302424.3303960], https://hal.archives-ouvertes.fr/hal-02333640

[16] D. CARVER, R. GOUICEM, J.-P. LOZI, J. SOPENA, B. LEPERS, W. ZWAENEPOEL, N. PALIX, J. L. LAWALL, G. MULLER.*Fork/Wait and Multicore Frequency Scaling: a Generational Clash*, in "10th Workshop on Programming Languages and Operating Systems", Huntsville, Canada, ACM Press, October 2019, p. 53-59 [*DOI :* 10.1145/3365137.3365400], https://hal.inria.fr/hal-02349987

[17] Q.-L. CHEN, J.-J. BAI, Z.-M. JIANG, J. L. LAWALL, S.-M. HU.*Detecting Data Races Caused by Inconsistent Lock Protection in Device Drivers*, in "SANER 2019 - 26th IEEE International Conference on Software Analysis, Evolution and Reengineering", Hangzhou, China, February 2019, https://hal.inria.fr/hal-02014196

[18] C. COURTAUD, J. SOPENA, G. MULLER, D. GRACIA.*Improving Prediction Accuracy of Memory Interferences for Multicore Platforms*, in "RTSS 2019 - 40th IEEE Real-Time Systems Symposium", Hong-Kong, China, IEEE, December 2019, https://hal.inria.fr/hal-02401625

[19] T. HOANG, J. L. LAWALL, R. J. OENTARYO, Y. TIAN, D. LO.*PatchNet: A Tool for Deep Patch Classification*, in "ICSE-Companion 2019 - IEEE/ACM 41st International Conference on Software Engineering", Montreal, Canada, IEEE, May 2019, p. 83-86 [*DOI :* 10.1109/ICSE-COMPANION.2019.00044], https://hal.inria.fr/hal-02408347

[20] Z.-M. JIANG, J.-J. BAI, J. L. LAWALL, S.-M. HU.*Fuzzing Error Handling Code in Device Drivers Based on Software Fault Injection*, in "ISSRE 2019 - The 30th International Symposium on Software Reliability Engineering", Berlin, Germany, October 2019 [*DOI :* 10.1109/ISSRE.2019.00022], https://hal.inria.fr/hal-02389293

[21] H. J. KANG, F. THUNG, J. L. LAWALL, G. MULLER, L. JIANG, D. LO.*Semantic Patches for Java Program Transformation*, in "33rd European Conference on Object-Oriented Programming (ECOOP 2019)", London, United Kingdom, Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, July 2019, vol. 134, p. 22:1–22:27 [*DOI :* 10.4230/LIPICS.ECOOP.2019.22], https://hal.inria.fr/hal-02182522

[22] F. LANIEL, D. CARVER, J. SOPENA, F. WAJSBURT, J. LEJEUNE, M. SHAPIRO.*Highlighting the Container Memory Consolidation Problems in Linux*, in "2019 IEEE 18th International Symposium on Network Computing and Applications (NCA)", Cambridge, United States, IEEE, September 2019, p. 1-4 [*DOI :* 10.1109/NCA.2019.8935034], https://hal.archives-ouvertes.fr/hal-02424007

[23] D. MERCADIER, P.-É. DAGAND.*Usuba: high-throughput and constant-time ciphers, by construction*, in "PLDI 2019 - 40th ACM SIGPLAN Conference on Programming Language Design and Implementation", Phoenix, United States, ACM Press, June 2019, p. 157-173 [*DOI :* 10.1145/3314221.3314636], https://hal.inria.fr/hal-02176603

Research Reports

[24] H. J. KANG, F. THUNG, J. L. LAWALL, G. MULLER, L. JIANG, D. LO.*Automating Program Transformation for Java Using Semantic Patches*, Inria Paris, February 2019, n^{o} RR-9256, https://hal.inria.fr/hal-02023368

References in notes

[25] T. BALL, E. BOUNIMOVA, B. COOK, V. LEVIN, J. LICHTENBERG, C. MCGARVEY, B. ONDRUSEK, S. K. RAJAMANI, A. USTUNER.*Thorough Static Analysis of Device Drivers*, in "EuroSys", 2006, p. 73–85

[26] A. BAUMANN, P. BARHAM, P.-É. DAGAND, T. HARRIS, R. ISAACS, S. PETER, T. ROSCOE, A. SCHÜPBACH, A. SINGHANIA.*The multikernel: A new OS architecture for scalable multicore systems*, in "SOSP", 2009, p. 29–44

[27] T. F. BISSYANDÉ, L. RÉVEILLÈRE, J. L. LAWALL, Y.-D. BROMBERG, G. MULLER.*Implementing an embedded compiler using program transformation rules*, in "Software: Practice and Experience", 2013

[28] T. F. BISSYANDÉ, L. RÉVEILLÈRE, J. L. LAWALL, Y.-D. BROMBERG, G. MULLER.*Implementing an Embedded Compiler using Program Transformation Rules*, in "Software: Practice and Experience", February 2015, vol. 45, n^{o} 2, p. 177-196, https://hal.archives-ouvertes.fr/hal-00844536

[29] T. F. BISSYANDÉ, L. RÉVEILLÈRE, J. L. LAWALL, G. MULLER.*Ahead of Time Static Analysis for Automatic Generation of Debugging Interfaces to the Linux Kernel*, in "Automated Software Engineering", May 2014, p. 1-39 [*DOI :* 10.1007/s10515-014-0152-4], https://hal.archives-ouvertes.fr/hal-00992283

[30] A. P. BLACK, S. DUCASSE, O. NIERSTRASZ, D. POLLET.*Pharo by Example*, Square Bracket Associates, 2010

[31] E. BRADY, K. HAMMOND.*Resource-Safe Systems Programming with Embedded Domain Specific Languages*, in "14th International Symposium on Practical Aspects of Declarative Languages (PADL)", LNCS, Springer, 2012, vol. 7149, p. 242–257

[32] T. BRAIBANT, D. POUS.*An Efficient Coq Tactic for Deciding Kleene Algebras*, in "1st International Conference on Interactive Theorem Proving (ITP)", LNCS, Springer, 2010, vol. 6172, p. 163–178

[33] C. CADAR, D. DUNBAR, D. R. ENGLER.*KLEE: Unassisted and Automatic Generation of High-Coverage Tests for Complex Systems Programs*, in "OSDI", 2008, p. 209–224

[34] V. CHIPOUNOV, G. CANDEA.*Reverse Engineering of Binary Device Drivers with RevNIC*, in "EuroSys", 2010, p. 167–180

[35] A. CHLIPALA.*The Bedrock Structured Programming System: Combining Generative Metaprogramming and Hoare Logic in an Extensible Program Verifier*, in "ICFP", 2013, p. 391–402

[36] L. A. CLARKE.*A system to generate test data and symbolically execute programs*, in "IEEE Transactions on Software Engineering", 1976, vol. 2, n^{o} 3, p. 215–222

[37] E. CLARKE, O. GRUMBERG, S. JHA, Y. LU, H. VEITH.*Counterexample-guided abstraction refinement for symbolic model checking*, in "J. ACM", 2003, vol. 50, n^{o} 5, p. 752–794

[38] P. COUSOT, R. COUSOT.*Abstract Interpretation: Past, Present and Future*, in "CSL-LICS", 2014, p. 2:1–2:10

[39] P.-É. DAGAND, A. BAUMANN, T. ROSCOE.*Filet-o-Fish: practical and dependable domain-specific languages for OS development*, in "Programming Languages and Operating Systems (PLOS)", 2009, p. 51–55

[40] I. DILLIG, T. DILLIG, A. AIKEN.*Sound, complete and scalable path-sensitive analysis*, in "PLDI", June 2008, p. 270–280

[41] D. R. ENGLER, B. CHELF, A. CHOU, S. HALLEM.*Checking System Rules Using System-Specific, Programmer-Written Compiler Extensions*, in "OSDI", 2000, p. 1–16

[42] D. R. ENGLER, D. Y. CHEN, A. CHOU, B. CHELF.*Bugs as Deviant Behavior: A General Approach to Inferring Errors in Systems Code*, in "SOSP", 2001, p. 57–72

[43] A. GOLDBERG, D. ROBSON.*Smalltalk-80: The Language and Its Implementation*, Addison-Wesley, 1983

[44] L. GU, A. VAYNBERG, B. FORD, Z. SHAO, D. COSTANZO.*CertiKOS: A Certified Kernel for Secure Cloud Computing*, in "Proceedings of the Second Asia-Pacific Workshop on Systems (APSys)", 2011, p. 3:1–3:5

[45] L. GUO, J. L. LAWALL, G. MULLER.*Oops! Where did that code snippet come from?*, in "11th Working Conference on Mining Software Repositories, MSR", Hyderabad, India, ACM, May 2014, p. 52–61

[46] A. ISRAELI, D. G. FEITELSON.*The Linux kernel as a case study in software evolution*, in "Journal of Systems and Software", 2010, vol. 83, n^o 3, p. 485–501

[47] A. KADAV, M. M. SWIFT.*Understanding modern device drivers*, in "ASPLOS", 2012, p. 87–98

[48] A. KENNEDY, N. BENTON, J. B. JENSEN, P.-É. DAGAND.*Coq: The World's Best Macro Assembler?*, in "PPDP", Madrid, Spain, ACM, 2013, p. 13–24

[49] G. A. KILDALL.*A Unified Approach to Global Program Optimization*, in "POPL", 1973, p. 194–206

[50] G. KLEIN, K. ELPHINSTONE, G. HEISER, J. ANDRONICK, 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 "SOSP", 2009, p. 207–220

[51] J. L. LAWALL, J. BRUNEL, N. PALIX, R. R. HANSEN, H. STUART, G. MULLER.*WYSIWIB: Exploiting fine-grained program structure in a scriptable API-usage protocol-finding process*, in "Software, Practice Experience", 2013, vol. 43, n^o 1, p. 67–92

[52] J. L. LAWALL, B. LAURIE, R. R. HANSEN, N. PALIX, G. MULLER.*Finding Error Handling Bugs in OpenSSL using Coccinelle*, in "Proceeding of the 8th European Dependable Computing Conference (EDCC)", Valencia, Spain, April 2010, p. 191–196

[53] J. L. LAWALL, D. LO.*An automated approach for finding variable-constant pairing bugs*, in "25th IEEE/ACM International Conference on Automated Software Engineering", Antwerp, Belgium, September 2010, p. 103–112

[54] J. L. LAWALL, D. PALINSKI, L. GNIRKE, G. MULLER.*Fast and Precise Retrieval of Forward and Back Porting Information for Linux Device Drivers*, in "2017 USENIX Annual Technical Conference", Santa Clara, CA, United States, July 2017, 12, https://hal.inria.fr/hal-01556589

[55] C. LE GOUES, W. WEIMER.*Specification Mining with Few False Positives*, in "TACAS", York, UK, Lecture Notes in Computer Science, March 2009, vol. 5505, p. 292–306

[56] Z. LI, S. LU, S. MYAGMAR, Y. ZHOU.*CP-Miner: A Tool for Finding Copy-paste and Related Bugs in Operating System Code*, in "OSDI", 2004, p. 289–302

[57] Z. LI, Y. ZHOU.*PR-Miner: automatically extracting implicit programming rules and detecting violations in large software code*, in "Proceedings of the 10th European Software Engineering Conference", 2005, p. 306–315

[58] D. LO, S. KHOO.*SMArTIC: towards building an accurate, robust and scalable specification miner*, in "FSE", 2006, p. 265–275

[59] J.-P. LOZI, F. DAVID, G. THOMAS, J. L. LAWALL, G. MULLER.*Fast and Portable Locking for Multicore Architectures*, in "ACM Transactions on Computer Systems", January 2016 [*DOI :* 10.1145/2845079], https://hal.inria.fr/hal-01252167

[60] S. LU, S. PARK, Y. ZHOU.*Finding Atomicity-Violation Bugs through Unserializable Interleaving Testing*, in "IEEE Transactions on Software Engineering", 2012, vol. 38, n^{o} 4, p. 844–860

[61] M. MERNIK, J. HEERING, A. M. SLOANE.*When and How to Develop Domain-specific Languages*, in "ACM Comput. Surv.", December 2005, vol. 37, n^{o} 4, p. 316–344, http://dx.doi.org/10.1145/1118890.1118892

[62] G. MORRISETT, G. TAN, J. TASSAROTTI, J.-B. TRISTAN, E. GAN.*RockSalt: better, faster, stronger SFI for the x86*, in "PLDI", 2012, p. 395-404

[63] M. ODERSKY, T. ROMPF.*Unifying functional and object-oriented programming with Scala*, in "Commun. ACM", 2014, vol. 57, n^{o} 4, p. 76–86

[64] M. C. OLESEN, R. R. HANSEN, J. L. LAWALL, N. PALIX.*Coccinelle: Tool support for automated CERT C Secure Coding Standard certification*, in "Science of Computer Programming", October 2014, vol. 91, n^{o} B, p. 141–160, https://hal.inria.fr/hal-01096185

[65] T. REPS, T. BALL, M. DAS, J. LARUS.*The Use of Program Profiling for Software Maintenance with Applications to the Year 2000 Problem*, in "ESEC/FSE", 1997, p. 432–449

[66] L. R. RODRIGUEZ, J. L. LAWALL.*Increasing Automation in the Backporting of Linux Drivers Using Coccinelle*, in "11th European Dependable Computing Conference - Dependability in Practice", Paris, France, 11th European Dependable Computing Conference - Dependability in Practice, November 2015, https://hal.inria.fr/hal-01213912

[67] C. RUBIO-GONZÁLEZ, H. S. GUNAWI, B. LIBLIT, R. H. ARPACI-DUSSEAU, A. C. ARPACI-DUSSEAU.*Error propagation analysis for file systems*, in "PLDI", Dublin, Ireland, ACM, June 2009, p. 270–280

[68] L. RYZHYK, P. CHUBB, I. KUZ, E. LE SUEUR, G. HEISER.*Automatic device driver synthesis with Termite*, in "SOSP", 2009, p. 73–86

[69] L. RYZHYK, A. WALKER, J. KEYS, A. LEGG, A. RAGHUNATH, M. STUMM, M. VIJ.*User-Guided Device Driver Synthesis*, in "OSDI", 2014, p. 661–676

[70] R. K. SAHA, J. L. LAWALL, S. KHURSHID, D. E. PERRY.*On the Effectiveness of Information Retrieval Based Bug Localization for C Programs*, in "ICSME 2014 - 30th International Conference on Software Maintenance

and Evolution", Victoria, Canada, IEEE, September 2014, p. 161-170 [*DOI :* 10.1109/ICSME.2014.38], https://hal.inria.fr/hal-01086082

[71] R. SAHA, J. L. LAWALL, S. KHURSHID, D. E. PERRY.*On the Effectiveness of Information Retrieval based Bug Localization for C Programs*, in "International Conference on Software Maintenance and Evolution (ICSME)", Victoria, BC, Canada, September 2014

[72] S. SAHA, J.-P. LOZI, G. THOMAS, J. L. LAWALL, G. MULLER.*Hector: Detecting resource-release omission faults in error-handling code for systems software*, in "DSN 2013 - 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)", Budapest, Hungary, IEEE Computer Society, June 2013, p. 1-12 [*DOI :* 10.1109/DSN.2013.6575307], https://hal.inria.fr/hal-00918079

[73] D. A. SCHMIDT.*Data Flow Analysis is Model Checking of Abstract Interpretations*, in "POPL", 1998, p. 38–48

[74] P. SENNA TSCHUDIN, J. L. LAWALL, G. MULLER.*3L: Learning Linux Logging*, in "BElgian-NEtherlands software eVOLution seminar (BENEVOL 2015)", Lille, France, December 2015, https://hal.inria.fr/hal-01239980

[75] M. SHAPIRO.*Purpose-built languages*, in "Commun. ACM", 2009, vol. 52, n^{o} 4, p. 36–41

[76] R. TARTLER, D. LOHMANN, J. SINCERO, W. SCHRÖDER-PREIKSCHAT.*Feature consistency in compile-time-configurable system software: facing the Linux 10,000 feature problem*, in "EuroSys", 2011, p. 47–60

[77] F. THUNG, D. X. B. LE, D. LO, J. L. LAWALL.*Recommending Code Changes for Automatic Backporting of Linux Device Drivers*, in "32nd IEEE International Conference on Software Maintenance and Evolution (ICSME)", Raleigh, North Carolina, United States, IEEE, October 2016, https://hal.inria.fr/hal-01355859

[78] W. WANG, M. GODFREY.*A Study of Cloning in the Linux SCSI Drivers*, in "Source Code Analysis and Manipulation (SCAM)", IEEE, 2011

[79] J. YANG, C. HAWBLITZEL.*Safe to the Last Instruction: Automated Verification of a Type-safe Operating System*, in "PLDI", 2010, p. 99–110

Team WILLOW

Models of visual object recognition and scene understanding

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
Paris

THEME
Vision, perception and multimedia interpretation

Table of contents

Team WILLOW

Creation of the Project-Team: 2007 June 01

Keywords:

Computer Science and Digital Science:

A3.1.1. - Modeling, representation
A3.4. - Machine learning and statistics
A5.3. - Image processing and analysis
A5.4. - Computer vision
A9. - Artificial intelligence
A9.1. - Knowledge
A9.2. - Machine learning

Other Research Topics and Application Domains:

B9.5.1. - Computer science
B9.5.6. - Data science

1. Team, Visitors, External Collaborators

Research Scientists

Jean Ponce [Team leader, Inria, Senior Researcher, on leave from Ecole Normale Supérieure]
Justin Carpentier [Inria, Researcher, from Oct 2019]
Ivan Laptev [Inria, Senior Researcher, HDR]
Jean-Paul Laumond [CNRS, Senior Researcher, from Feb 2019, HDR]
Josef Sivic [Inria, Senior Researcher, HDR]

External Collaborator

Mathieu Aubry [École Nationale des Ponts et Chaussées]

Technical Staff

Sofiane Allayen [Inria, until Apr 2019]
Yann Dubois de Mont Marin [Inria, from Dec 2019]
Igor Kalevatykh [Inria]
Oumayma Bounou [Inria, from Sep 2019]

PhD Students

Minttu Alakuijala [Inria, from Feb 2019]
Thomas Eboli [École Normale Supérieure de Paris]
Aamr El Kazdadi [Inria, from Oct 2019]
Pierre-Louis Guhur [Univ Paris-Saclay, from Sep 2019]
Yana Hasson [Inria]
Yann Labbe [École Normale Supérieure de Cachan]
Bruno Lecouat [Inria, from Sep 2019]
Zongmian Li [Inria]
Antoine Miech [Inria]
Julia Peyre [Inria, until Aug 2019]
Ronan Riochet [Inria]
Ignacio Rocco Spremolla [Inria]
Robin Strudel [École Normale Supérieure de Paris]
Gul Varol Simsekli [Inria, until Jun 2019]

Van Huy Vo [École Normale Supérieure de Paris]
Dimitri Zhukov [Inria]

Post-Doctoral Fellows
Vijaya Kumar Reddy [Inria]
Makarand Tapaswi [Inria]
Sergey Zagoruyko [Inria, until Feb 2019]

Visiting Scientists
Pierre Yves Masse [Czech Technical University]
Vladimir Petrik [Czech Technical University]
Hazel Doughty [University of Bristol]

Administrative Assistants
Helene Milome [Inria, until Jul 2019]
Mathieu Mourey [Inria, from Jul 2019]

2. Overall Objectives

2.1. Statement

Object recognition —or, in a broader sense, scene understanding— is the ultimate scientific challenge of computer vision: After 40 years of research, robustly identifying the familiar objects (chair, person, pet), scene categories (beach, forest, office), and activity patterns (conversation, dance, picnic) depicted in family pictures, news segments, or feature films is still beyond the capabilities of today's vision systems. On the other hand, truly successful object recognition and scene understanding technology will have a broad impact in application domains as varied as defense, entertainment, health care, human-computer interaction, image retrieval and data mining, industrial and personal robotics, manufacturing, scientific image analysis, surveillance and security, and transportation.

Despite the limitations of today's scene understanding technology, tremendous progress has been accomplished in the past ten years, due in part to the formulation of object recognition as a statistical pattern matching problem. The emphasis is in general on the features defining the patterns and on the algorithms used to learn and recognize them, rather than on the representation of object, scene, and activity categories, or the integrated interpretation of the various scene elements. WILLOW complements this approach with an ambitious research program explicitly addressing the representational issues involved in object recognition and, more generally, scene understanding.

Concretely, our objective is to develop geometric, physical, and statistical models for all components of the image interpretation process, including illumination, materials, objects, scenes, and human activities. These models will be used to tackle fundamental scientific challenges such as three-dimensional (3D) object and scene modeling, analysis, and retrieval; human activity capture and classification; and category-level object and scene recognition. They will also support applications with high scientific, societal, and/or economic impact in domains such as quantitative image analysis in science and humanities; film post-production and special effects; and video annotation, interpretation, and retrieval. Machine learning is a key part of our effort, with a balance of practical work in support of computer vision application and methodological research aimed at developing effective algorithms and architectures.

WILLOW was created in 2007: It was recognized as an Inria team in January 2007, and as an official project-team in June 2007. WILLOW is a joint research team between Inria Paris, Ecole Normale Supérieure (ENS) and Centre National de la Recherche Scientifique (CNRS).

This year we have hired Justin Carpentier as an Inria researcher. Jean-Paul Laumond has joined Willow as a senior CNRS researcher. Both Justin and Jean-Paul have robotics background and will strengthen the team in its new efforts towards learning embodied representations and robotics. Vladimir Petrik and Pierre-Yves Masse have been visiting post-docs within the framework of collaboration with the Intelligent Machine Perception project lead by J. Sivic at the Czech Technical University in Prague. Hazel Doughty has been a visiting PhD student from University of Bristol. Minttu Alakuijala, Aamr El Kazdadi, Pierre-Louis Guhur and Bruno Lecouat have joined Willow as new PhD students.

3. Research Program

3.1. 3D object and scene modeling, analysis, and retrieval

This part of our research focuses on geometric models of specific 3D objects at the local (differential) and global levels, physical and statistical models of materials and illumination patterns, and modeling and retrieval of objects and scenes in large image collections. Our past work in these areas includes research aimed at recognizing rigid 3D objects in cluttered photographs taken from arbitrary viewpoints (Rothganger *et al.*, 2006), segmenting video sequences into parts corresponding to rigid scene components before recognizing these in new video clips (Rothganger *et al.*, 2007), retrieval of particular objects and buildings from images and videos (Sivic and Zisserman, 2003) and (Philbin *et al.*, 2007), and a theoretical study of a general formalism for modeling central and non-central cameras using the formalism and terminology of classical projective geometry (Ponce, 2009 and Batog *et al.*, 2010).

We have also developed multi-view stereopsis algorithms that have proven remarkably effective at recovering intricate details and thin features of compact objects and capturing the overall structure of large-scale, cluttered scenes. We have obtained a US patent 8,331,615 [0] for the corresponding software (PMVS, https://github.com/pmoulon/CMVS-PMVS) which is available under a GPL license and used for film production by ILM and Weta as well as by Google in Google Maps. It is also the basic technology used by Iconem, a start-up founded by Y. Ubelmann, a Willow collaborator. We have also applied our multi-view-stereo approach to model archaeological sites together with developing representations and efficient retrieval techniques to enable matching historical paintings to 3D models of archaeological sites (Russel *et al.*, 2011).

Our current efforts in this area are outlined in detail in Section. 7.1.

3.2. Category-level object and scene recognition

The objective in this core part of our research is to learn and recognize quickly and accurately thousands of visual categories, including materials, objects, scenes, and broad classes of temporal events, such as patterns of human activities in picnics, conversations, etc. The current paradigm in the vision community is to model/learn one object category (read 2D aspect) at a time. If we are to achieve our goal, we have to break away from this paradigm, and develop models that account for the tremendous variability in object and scene appearance due to texture, material, viewpoint, and illumination changes within each object category, as well as the complex and evolving relationships between scene elements during the course of normal human activities.

Our current work in this area is outlined in detail in Section 7.2.

3.3. Image restoration, manipulation and enhancement

The goal of this part of our research is to develop models, and methods for image/video restoration, manipulation and enhancement. The ability to “intelligently" manipulate the content of images and video is just as essential as high-level content interpretation in many applications: This ranges from restoring old films or removing unwanted wires and rigs from new ones in post production, to cleaning up a shot of your

[0] The patent: "Match, Expand, and Filter Technique for Multi-View Stereopsis" was issued December 11, 2012 and assigned patent number 8,331,615.

daughter at her birthday party, which is lovely but noisy and blurry because the lights were out when she blew the candles, or editing out a tourist from your Roman holiday video. Going beyond the modest abilities of current “digital zoom" (bicubic interpolation in general) so you can close in on that birthday cake, “deblock" a football game on TV, or turn your favorite DVD into a blue-ray, is just as important.

In this context, we believe there is a new convergence between computer vision, machine learning, and signal processing. For example: The idea of exploiting self-similarities in image analysis, originally introduced in computer vision for texture synthesis applications (Efros and Leung, 1999), is the basis for non-local means (Buades *et al.*, 2005), one of today’s most successful approaches to image restoration. In turn, by combining a powerful sparse coding approach to non-local means (Dabov *et al.*, 2007) with modern machine learning techniques for dictionary learning (Mairal *et al.*, 2010), we have obtained denoising and demosaicking results that are the state of the art on standard benchmarks (Mairal *et al.*, 2009).

Our current work is outlined in detail in Section 7.3.

3.4. Human activity capture and classification

From a scientific point of view, visual action understanding is a computer vision problem that until recently has received little attention outside of extremely specific contexts such as surveillance or sports. Many of the current approaches to the visual interpretation of human activities are designed for a limited range of operating conditions, such as static cameras, fixed scenes, or restricted actions. The objective of this part of our project is to attack the much more challenging problem of understanding actions and interactions in unconstrained video depicting everyday human activities such as in sitcoms, feature films, or news segments. The recent emergence of automated annotation tools for this type of video data (Everingham, Sivic, Zisserman, 2006; Laptev, Marszałek, Schmid, Rozenfeld, 2008; Duchenne, Laptev, Sivic, Bach, Ponce, 2009) means that massive amounts of labelled data for training and recognizing action models will at long last be available.

Our research agenda in this scientific domain is described below and our recent results are outlined in detail in Section 7.4.

- **Weakly-supervised learning and annotation of human actions in video.** We aim to leverage the huge amount of video data using readily-available annotations in the form of video scripts. Scripts, however, often provide only imprecise and incomplete information about the video. We address this problem with weakly-supervised learning techniques both at the text and image levels.
- **Descriptors for video representation.** Video representation has a crucial role for recognizing human actions and other components of a visual scene. Our work in this domain aims to develop generic methods for representing video data based on realistic assumptions. In particular, we develop deep learning methods and design new trainable representations for various tasks such as human action recognition, person detection, segmentation and tracking.

3.5. Learning embodied representations

Computer vision has come a long way toward understanding images and videos in terms of scene geometry, object labels, locations and poses of people or classes of human actions. This “understanding”, however, remains largely disconnected from reasoning about the physical world. For example, what will happen if removing a tablecloth from a setted table? What actions will be needed to resume an interrupted meal? We believe that a true *embodied* understanding of dynamic scenes from visual observations is the next major research challenge. We plan to address this challenge by developing new models and algorithms with an emphasis on the synergy between vision, learning, robotics and natural language understanding. If successful, this research direction will bring significant advances in high-impact applications such as autonomous driving, home robotics and personal visual assistance.

Learning embodied representations is planned to be a major research axis for the successor of the Willow team. Meanwhile we have already started work in this direction and report our first results in Section 7.5.

4. Application Domains

4.1. Introduction

We believe that foundational modeling work should be grounded in applications. This includes (but is not restricted to) the following high-impact domains.

4.2. Quantitative image analysis in science and humanities

We plan to apply our 3D object and scene modeling and analysis technology to image-based modeling of human skeletons and artifacts in anthropology, and large-scale site indexing, modeling, and retrieval in archaeology and cultural heritage preservation. Most existing work in this domain concentrates on image-based rendering, that is, the synthesis of good-looking pictures of artifacts and digs. We plan to focus instead on quantitative applications. We are engaged in a project involving the archaeology laboratory at ENS and focusing on image-based artifact modeling and decorative pattern retrieval in Pompeii. Application of our 3D reconstruction technology is now being explored in the field of cultural heritage and archeology by the start-up Iconem, founded by Y. Ubelmann, a Willow collaborator.

4.3. Video Annotation, Interpretation, and Retrieval

Both specific and category-level object and scene recognition can be used to annotate, augment, index, and retrieve video segments in the audiovisual domain. The Video Google system developed by Sivic and Zisserman (2005) for retrieving shots containing specific objects is an early success in that area. A sample application, suggested by discussions with Institut National de l'Audiovisuel (INA) staff, is to match set photographs with actual shots in film and video archives, despite the fact that detailed timetables and/or annotations are typically not available for either medium. Automatically annotating the shots is of course also relevant for archives that may record hundreds of thousands of hours of video. Some of these applications will be pursued in our MSR-Inria project.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Four Prairie chairs awarded to I. Laptev, J.-P. Laumond, J. Ponce and J. Sivic by the international selection committee.
- Best Paper Award at FG (Automatic Face and Gesture Recognition - http://fg2019.org/awards/) 2019. (M. Tapaswi)
- Best paper finalist at CVPR 2019 for the work of Z. Li, J. Sedlar, J. Carpentier, I. Laptev, N. Mansard and J. Sivic Estimating 3D Motion and Forces of Person-Object Interactions From Monocular Video, IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2019) [0]
- Best student paper awarded to H Cisneros, J Sivic, T Mikolov, Evolving Structures in Complex Systems at IEEE Symposium Series on Computational Intelligence (2019)

[0] More details at: https://www.ciirc.cvut.cz/vysledek-ciirc-cvut-se-dostal-do-uzsiho-vyberu-nejlepsich-clanku-prestizni-konference-cvpr-v-pocitacovem-videni/ and the list of all shortlisted papers is available at http://cvpr2019.thecvf.com/files/CVPR%202019%20-%20Welcome%20Slides%20Final.pdf

5.1.2. Visibility

- In 2019 we have recruited two excellent researchers with robotics background: Justin Carpentier and Jean-Paul Laumond, who will strengthen the team and will help developing a new research axis on learning embodied representations.
- J. Ponce co-organized the PRAIRIE AI Summer School, Paris, 2019.
- J. Ponce has been a key person in creating the PRAIRIE Institute for AI research in Paris, innagurated in October 2019.

BEST PAPERS AWARDS :

[11]

H. CISNEROS, J. SIVIC, T. MIKOLOV.*Evolving Structures in Complex Systems*, in "SSCI 2019 - IEEE Symposium Series on Computational Intelligence", Xiamen, China, December 2019, https://arxiv.org/abs/1911.01086 - IEEE Symposium Series on Computational Intelligence 2019 (IEEE SSCI 2019), https://hal.inria.fr/hal-02448134

6. New Software and Platforms

6.1. Pinocchio

KEYWORDS: Robotics - Biomechanics - Mechanical multi-body systems

FUNCTIONAL DESCRIPTION: Pinocchio instantiates state-of-the-art Rigid Body Algorithms for poly-articulated systems based on revisited Roy Featherstone's algorithms. In addition, Pinocchio instantiates analytical derivatives of the main Rigid-Body Algorithms like the Recursive Newton-Euler Algorithms or the Articulated-Body Algorithm. Pinocchio is first tailored for legged robotics applications, but it can be used in extra contexts. It is built upon Eigen for linear algebra and FCL for collision detection. Pinocchio comes with a Python interface for fast code prototyping.

- Partner: CNRS
- Contact: Justin Carpentier
- URL: https://github.com/stack-of-tasks/pinocchio

6.2. VRAnalogy

Visual Relations detector using Analogy

KEYWORDS: Computer vision - Machine learning

FUNCTIONAL DESCRIPTION: Implementation of the paper "Detecting Unseen Visual Relations Using Analogies", Peyre et al', ICCV19

- Contact: Julia Peyre

6.3. d2-net

D2-Net: A Trainable CNN for Joint Description and Detection of Local Features

KEYWORD: Feature points

FUNCTIONAL DESCRIPTION: This repository contains the implementation of the following paper:

"D2-Net: A Trainable CNN for Joint Detection and Description of Local Features". M. Dusmanu, I. Rocco, T. Pajdla, M. Pollefeys, J. Sivic, A. Torii, and T. Sattler. CVPR 2019.

- Participants: Mihai Dusmanu, Ignacio Rocco Spremolla, Tomas Pajdla, Marc Pollefeys, Josef Sivic, Akihiko Torii and Torsten Sattler
- Contact: Ignacio Rocco Spremolla
- Publication: D2-Net: A Trainable CNN for Joint Detection and Description of Local Features
- URL: https://github.com/mihaidusmanu/d2-net

6.4. CrossTask

Cross-task weakly supervised learning from instructional videos

KEYWORDS: Videos - Machine learning

FUNCTIONAL DESCRIPTION: Open source release of the software package for the CVPR'19 paper "Cross-task weakly supervised learning from instructional videos" by D. Zhukov, J.-B. Alayrac, R. G. Cinbis, D. Fouhey, I. Laptev and J. Sivic

- Participants: Dimitri Zhukov, Jean-Baptiste Alayrac, Cinbis Gokberk, David Fouhey, Ivan Laptev and Josef Sivic
- Contact: Dimitri Zhukov
- URL: https://github.com/DmZhukov/CrossTask

6.5. MImE

Manipulation Imitation Environments

KEYWORDS: Robotics - Simulator - Computer vision

FUNCTIONAL DESCRIPTION: Simulation environment for learning robotics manipulation policies.

- Contact: Igor Kalevatykh
- URL: https://github.com/ikalevatykh/mime/

6.6. iReal

iReal: Implementing interactive scene understanding for a mixed reality device

KEYWORDS: Computer vision - Augmented reality - Demonstration

FUNCTIONAL DESCRIPTION: The goal of this project is to build a demonstration prototype of a smart personal assistant that sees and understands its surroundings to help the user navigate in unfamiliar environments, recognize new people and operate never seen before devices. The assistant will be implemented on the Microsoft HoloLens mixed reality device and will integrate the latest research software for automatic visual recognition and scene understanding developed in the Inria Willow team

- Contact: Mauricio Diaz

6.7. Sim2RealAugment

Learning to Augment Synthetic Images for Sim2Real Policy Transfer

KEYWORDS: Robotics - Computer vision

FUNCTIONAL DESCRIPTION: Implements the method described in Learning to Augment Synthetic Images for Sim2Real Policy Transfer (2019), A. Pashevich, R. Strudel, I. Kalevatykh, I. Laptev and C. Schmid, in Proc. IROS'19, Macau, China.

- Contact: Ivan Laptev
- URL: http://pascal.inrialpes.fr/data2/sim2real/

6.8. HowTo100M

HowTo100M: Learning a Text-Video Embedding by Watching Hundred Million Narrated Video Clips

KEYWORDS: Computer vision - Video analysis

FUNCTIONAL DESCRIPTION: Implements the method provides the dataset used in the paper HowTo100M: Learning a Text-Video Embedding by Watching Hundred Million Narrated Video Clips (2019), A. Miech, D. Zhukov, J.-B. Alayrac, M. Tapaswi, I. Laptev and J. Sivic, in Proc. ICCV'19, Seoul, South Korea.

- Contact: Ivan Laptev
- URL: https://www.di.ens.fr/willow/research/howto100m/

6.9. ObMan

Learning joint reconstruction of hands and manipulated objects

KEYWORDS: Computer vision - 3D reconstruction

FUNCTIONAL DESCRIPTION: Implements the method and provides dataset used in the paper Learning joint reconstruction of hands and manipulated objects (2019), Y. Hasson, G. Varol, D. Tzionas, I. Kalevatykh, M. Black, I. Laptev and C. Schmid, in Proc. CVPR'19, Long Beach, CA, USA.

- Contact: Yana Hasson
- URL: https://hassony2.github.io/obman.html

7. New Results

7.1. 3D object and scene modeling, analysis, and retrieval

7.1.1. Learning joint reconstruction of hands and manipulated objects

Participants: Yana Hasson, Gül Varol, Dimitrios Tzionas, Igor Kalevatykh, Michael Black, Ivan Laptev, Cordelia Schmid.

Estimating hand-object manipulations is essential for interpreting and imitating human actions. Previous work has made significant progress towards reconstruction of hand poses and object shapes in isolation. Yet, reconstructing hands and objects during manipulation is a more challenging task due to significant occlusions of both the hand and object. While presenting challenges, manipulations may also simplify the problem since the physics of contact restricts the space of valid hand-object configurations. For example, during manipulation, the hand and object should be in contact but not interpenetrate. In [14] we regularize the joint reconstruction of hands and objects with manipulation constraints. We present an end-to-end learnable model that exploits a novel contact loss that favors physically plausible hand-object constellations. Our approach improves grasp quality metrics over baselines, using RGB images as input. To train and evaluate the model, we also propose a new large-scale synthetic dataset, ObMan, with hand-object manipulations. We demonstrate the transferability of ObMan-trained models to real data. Figure 1 presents some example results.

7.1.2. D2-Net: A Trainable CNN for Joint Detection and Description of Local Features

Participants: Mihai Dusmanu, Ignacio Rocco, Tomas Pajdla, Marc Pollefeys, Josef Sivic, Akihiko Torii, Torsten Sattler.

In [13], we address the problem of finding reliable pixel-level correspondences under difficult imaging conditions. We propose an approach where a single convolutional neural network plays a dual role: It is simultaneously a dense feature descriptor and a feature detector, as illustrated in Figure 2. By postponing the detection to a later stage, the obtained keypoints are more stable than their traditional counterparts based on early detection of low-level structures. We show that this model can be trained using pixel correspondences extracted from readily available large-scale SfM reconstructions, without any further annotations. The proposed method obtains state-of-the-art performance on both the difficult Aachen Day-Night localization dataset and the InLoc indoor localization benchmark, as well as competitive performance on other benchmarks for image matching and 3D reconstruction.

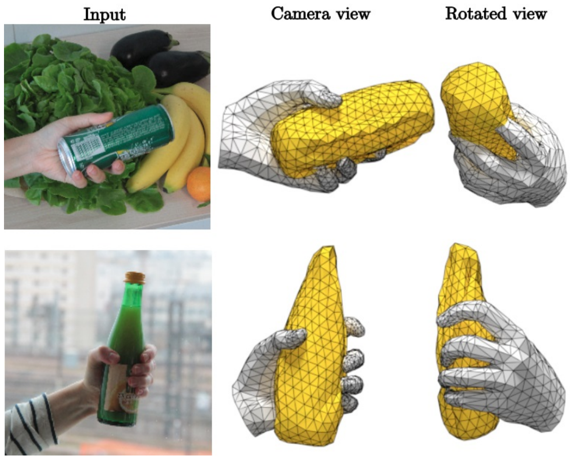

Figure 1. Our method jointly reconstructs hand and object meshes from a monocular RGB image. Note that the model generating the predictions for the above images, which we captured with an ordinary camera, was trained only on images from our synthetic dataset, ObMan.

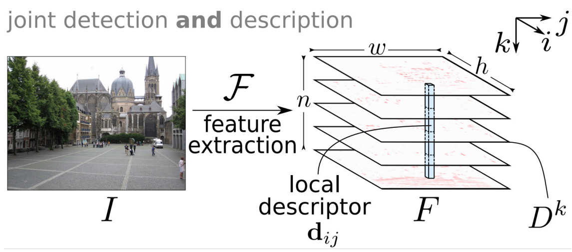

Figure 2. A feature extraction CNN $\mathcal{F}$ is used to extract feature maps that play a dual role: (i) local descriptors $\mathbf{d}_{ij}$ are simply obtained by traversing all the n feature maps D^k at a spatial position (i, j); (ii) detections are obtained by performing a non-local-maximum suppression on a feature map followed by a non-maximum suppression across each descriptor.

7.1.3. Is This The Right Place? Geometric-Semantic Pose Verification for Indoor Visual Localization

Participants: Hajime Taira, Ignacio Rocco, Jiri Sedlar, Masatoshi Okutomi, Josef Sivic, Tomas Pajdla, Torsten Sattler, Akihiko Torii.

Visual localization in large and complex indoor scenes, dominated by weakly textured rooms and repeating geometric patterns, is a challenging problem with high practical relevance for applications such as Augmented Reality and robotics. To handle the ambiguities arising in this scenario, a common strategy is, first, to generate multiple estimates for the camera pose from which a given query image was taken. The pose with the largest geometric consistency with the query image, e.g., in the form of an inlier count, is then selected in a second stage. While a significant amount of research has concentrated on the first stage, there is considerably less work on the second stage. In [21], we thus focus on pose verification. We show that combining different modalities, namely appearance, geometry, and semantics, considerably boosts pose verification and consequently pose accuracy, as illustrated in Figure 3. We develop multiple hand-crafted as well as a trainable approach to join into the geometric-semantic verification and show significant improvements over state-of-the-art on a very challenging indoor dataset.

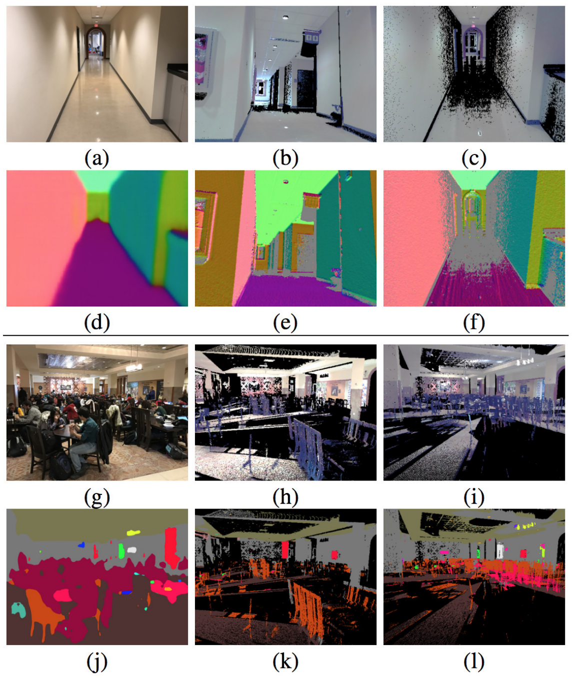

Figure 3. Given a set of camera pose estimates for a query image (a, g), we seek to identify the most accurate estimate. (b, h) Due to severe occlusion and weak textures, a state-of-the-art methodfails to identify the correct camera pose. To overcome those difficulties, we use several modalities along with visual appearance: (top) surface normals and (bottom) semantics. (c, i) Our approach verifies the estimated pose by comparing the semantics and surface normals extracted from the query (d, j) and database (f, l).

7.1.4. An Efficient Solution to the Homography-Based Relative Pose Problem With a Common Reference Direction

Participants: Yaqing Ding, Jian Yang, Jean Ponce, Hui Kong.

In [12], we propose a novel approach to two-view minimal-case relative pose problems based on homography with a common reference direction. We explore the rank-1 constraint on the difference between the Euclidean homography matrix and the corresponding rotation, and propose an efficient two-step solution for solving both the calibrated and partially calibrated (unknown focal length) problems. We derive new 3.5-point, 3.5-point, 4-point solvers for two cameras such that the two focal lengths are unknown but equal, one of them is unknown, and both are unknown and possibly different, respectively. We present detailed analyses and comparisons with existing 6-and 7-point solvers, including results with smart phone images.

7.1.5. Coordinate-Free Carlsson-Weinshall Duality and Relative Multi-ViewGeometry

Participants: Matthew Trager, Martial Hebert, Jean Ponce.

In [23], we present a coordinate-free description of Carlsson-Weinshall duality between scene points and camera pinholes and use it to derive a new characterization of primal/dual multi-view geometry. In the case of three views, a particular set of reduced trilinearities provide a novel parameterization of camera geometry that, unlike existing ones, is subject only to very simple internal constraints. These trilinearities lead to new "quasi-linear" algorithms for primal and dual structure from motion. We include some preliminary experiments with real and synthetic data.

7.1.6. Build your own hybrid thermal/EO camera for autonomous vehicle

Participants: Yigong Zhang, Yicheng Gao, Shuo Gu, Yubin Guo, Minghao Liu, Zezhou Sun, Zhixing Hou, Hang Yang, Ying Wang, Jian Yang, Jean Ponce, Hui Kong.

In [24], we propose a novel paradigm to design a hybrid thermal/EO (Electro-Optical or visible-light) camera, whose thermal and RGB frames are pixel-wisely aligned and temporally synchronized. Compared with the existing schemes, we innovate in three ways in order to make it more compact in dimension, and thus more practical and extendable for real-world applications. The first is a redesign of the structure layout of the thermal and EO cameras. The second is on obtaining a pixel-wise spatial registration of the thermal and RGB frames by a coarse mechanical adjustment and a fine alignment through a constant homography warping. The third innovation is on extending one single hybrid camera to a hybrid camera array, through which we can obtain wide-view spatially aligned thermal, RGB and disparity images simultaneously. The experimental results show that the average error of spatial-alignment of two image modalities can be less than one pixel. Some results of our method are illustrated in Figure 4.

7.2. Category-level object and scene recognition

7.2.1. Detecting unseen visual relations using analogies

Participants: Julia Peyre, Ivan Laptev, Cordelia Schmid, Josef Sivic.

In [19], we seek to detect visual relations in images of the form of triplets t = (subject, predicate, object), such as "person riding dog", where training examples of the individual entities are available but their combinations are unseen at training. This is an important set-up due to the combinatorial nature of visual relations: collecting sufficient training data for all possible triplets would be very hard. The contributions of this work are three-fold. First, we learn a representation of visual relations that combines (i) individual embeddings for subject, object and predicate together with (ii) a visual phrase embedding that represents the relation triplet. Second, we learn how to transfer visual phrase em- beddings from existing training triplets to unseen test triplets using analogies between relations that involve similar objects. Third, we demonstrate the benefits of our approach on three challenging datasets : on HICO-DET, our model achieves significant improvement over a strong baseline for both frequent and unseen triplets, and we observe similar improvement for the retrieval of unseen triplets with out-of- vocabulary predicates on the COCO-a dataset as well as the challenging unusual triplets in the UnRel dataset. Figure 5 presents an illustration of the approach.

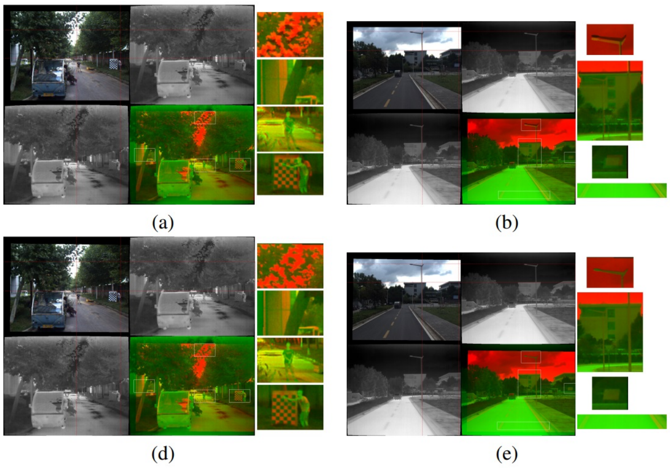

Figure 4. Results of alignment between the thermal and RGB frames of three sets of hybrid cameras before and after homography warping, respectively. (a), (b) and (c) are the alignment results before the homography warping, respectively. In each sub-figure, the layout of images is arranged as follows. Top-left: the aligned RGB image. Top-middle and bottom-left: the same aligned thermal image. Bottom-middle: the fusion image. (d), (e) and (f) are the alignment results after the homography warping, respectively. Likewise, the layout of images in each sub-figure is the same as those of (a), (b) and (c). To show the effect of homography rectification, we have overlaid red dotted lines horizontally and vertically onto the each sub-figure. In addition, the right column of each sub-figure zooms in four selected image regions to help us to view the warping result.

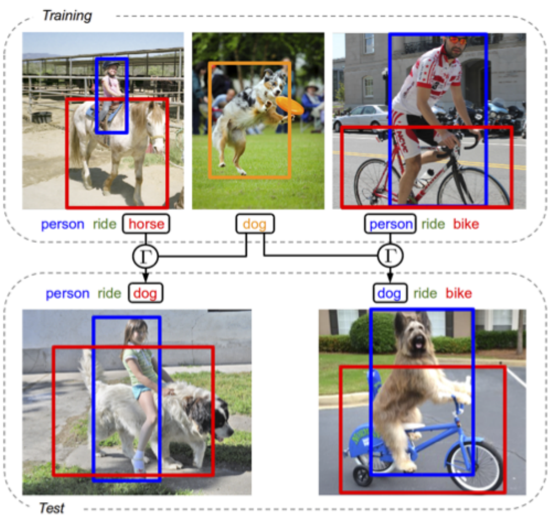

Figure 5. Illustration of transfer by analogy. We transfer visual representations of relations seen in the training set such as "person ride horse" to represent new unseen relations in the test set such as "person ride dog".

7.2.2. SFNet: Learning Object-aware Semantic Correspondence

Participants: Junghyup Lee, Dohyung Kim, Jean Ponce, Bumsub Ham.

In [15], we address the problem of semantic correspondence, that is, establishing a dense flow field between images depicting different instances of the same object or scene category. We propose to use images annotated with binary foreground masks and subjected to synthetic geometric deformations to train a convolutional neural network (CNN) for this task. Using these masks as part of the supervisory signal offers a good compromise between semantic flow methods, where the amount of training data is limited by the cost of manually selecting point correspondences, and semantic alignment ones, where the regression of a single global geometric transformation between images may be sensitive to image-specific details such as background clutter. We propose a new CNN architecture, dubbed SFNet, which implements this idea. It leverages a new and differentiable version of the argmax function for end-to-end training, with a loss that combines mask and flow consistency with smoothness terms. Experimental results demonstrate the effectiveness of our approach, which significantly outperforms the state of the art on standard benchmarks. Figure 6 presents an illustration of the approach.

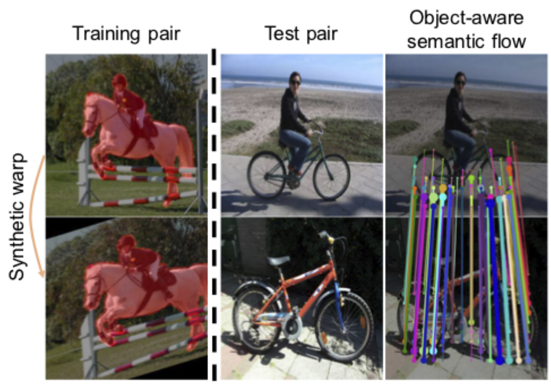

Figure 6. We use pairs of warped foreground masks obtained from a single image (left) as a supervisory signal to train our model.This allows us to establish object-aware semantic correspondences across images depicting different instances of the same object or scene category (right). No masks are required at test time.

7.2.3. Hyperpixel Flow: Semantic Correspondence with Multi-layer Neural Features

Participants: Juhong Min, Jongmin Kim, Jean Ponce, Minsu Cho.

In [17], we establish visual correspondences under large intra-class variations requires analyzing images at different levels , from features linked to semantics and context to local patterns, while being invariant to instance-specific details. To tackle these challenges, we represent images by "hyper-pixels" that leverage a small number of relevant features selected among early to late layers of a convolutional neural network. Taking advantage of the condensed features of hyperpixels, we develop an effective real-time matching algorithm based on Hough geometric voting. The proposed method, hyperpixel flow, sets a new state of the art on three standard benchmarks as well as a new dataset, SPair-71k, which contains a significantly larger number of image pairs than existing datasets, with more accurate and richer annotations for in-depth analysis. Figure 7 presents an illustration of the approach.

7.2.4. Exploring Weight Symmetry in Deep Neural Networks

Participants: Xu Shell Hu, Sergey Zagoruyko, Nikos Komodakis.

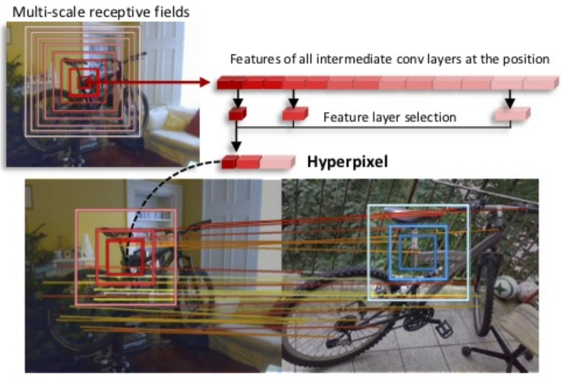

Figure 7. Hyperpixel flow. Top: The hyperpixel is a multi-layer pixel representation created with selected levels of features optimized for semantic correspondence. It provides multi-scale features, resolving local ambiguities. Bottom: The proposed method, hyperpixel flow, establishes dense correspondences in real time using hyperpixels.

In [7], we propose to impose symmetry in neural network parameters to improve parameter usage and make use of dedicated convolution and matrix multiplication routines. Due to significant reduction in the number of parameters as a result of the symmetry constraints, one would expect a dramatic drop in accuracy. Surprisingly, we show that this is not the case, and, depending on network size, symmetry can have little or no negative effect on network accuracy, especially in deep overparameterized networks. We propose several ways to impose local symmetry in recurrent and convolutional neural networks, and show that our symmetry parameterizations satisfy universal approximation property for single hidden layer networks. We extensively evaluate these parameterizations on CIFAR, ImageNet and language modeling datasets, showing significant benefits from the use of symmetry. For instance, our ResNet-101 with channel-wise symmetry has almost 25% less parameters and only 0.2% accuracy loss on ImageNet.

7.2.5. *Bilinear image translation for temporal analysis of photo collections*

Participants: Théophile Dalens, Mathieu Aubry, Josef Sivic.

In [5], we propose an approach for analyzing unpaired visual data annotated with time stamps by generating how images would have looked like if they were from different times. To isolate and transfer time dependent appearance variations, we introduce a new trainable bilinear factor separation module. We analyze its relation to classical factored representations and concatenation-based auto-encoders. We demonstrate this new module has clear advantages compared to standard concatenation when used in a bottleneck encoder-decoder convolutional neural network architecture. We also show that it can be inserted in a recent adversarial image translation architecture, enabling the image transformation to multiple different target time periods using a single network. We apply our model to a challenging collection of more than 13,000 cars manufactured between 1920 and 2000 and a dataset of high school yearbook portraits from 1930 to 2009, as illustrated in Figure 8. This allows us, for a given new input image, to generate a "history-lapse video" revealing changes over time by simply varying the target year. We show that by analyzing the generated history-lapse videos we can identify object deformations across time, extracting interesting changes in visual style over decades.

7.3. Image restoration, manipulation and enhancement

7.3.1. *Deformable Kernel Networks for Joint Image Filtering*

Participants: Beomjun Kim, Jean Ponce, Bumsub Ham.

Figure 8. Our method takes as input an image of an object (in green), such as a car, and generates what it would have looked like in another time-period (in blue). Each row shows temporal translation for a different input car image (in green). The translation model is trained on an unpaired dataset of cars with time stamps. We show that analyzing changes between the generated images reveal structural deformations in car shape and appearance over time.

Joint image filters are used to transfer structural details from a guidance picture used as a prior to a target image, in tasks such as enhancing spatial resolution and suppressing noise. Previous methods based on convolutional neural networks (CNNs) combine nonlinear activations of spatially-invariant kernels to estimate structural details and regress the filtering result. In this paper, we instead learn explicitly sparse and spatially-variant kernels. In [28], we propose a CNN architecture and its efficient implementation, called the deformable kernel network (DKN), that outputs sets of neighbors and the corresponding weights adaptively for each pixel. The filtering result is then computed as a weighted average. We also propose a fast version of DKN that runs about four times faster for an image of size 640×480. We demonstrate the effectiveness and flexibility of our models on the tasks of depth map upsampling, saliency map upsampling, cross-modality image restoration, texture removal, and semantic segmentation. In particular, we show that the weighted averaging process with sparsely sampled 3×3 kernels outperforms the state of the art by a significant margin.

7.3.2. Revisiting Non Local Sparse Models for Image Restoration

Participants: Bruno Lecouat, Jean Ponce, Julien Mairal.

In [29], we propose a differentiable algorithm for image restoration inspired by the success of sparse models and self-similarity priors for natural images. Our approach builds upon the concept of joint sparsity between groups of similar image patches, and we show how this simple idea can be implemented in a differentiable architecture, allowing end-to-end training. The algorithm has the advantage of being interpretable, performing sparse decompositions of image patches, while being more parameter efficient than recent deep learning methods. We evaluate our algorithm on grayscale and color denoising, where we achieve competitive results, and on demoisaicking, where we outperform the most recent state-of-the-art deep learning model with 47 times less parameters and a much shallower architecture. Figure 9 shows results of the proposed approach.

7.4. Human activity capture and classification

7.4.1. Video Face Clustering with Unknown Number of Clusters

Figure 9. Demosaicking result obtained by our method. Top right: Ground truth. Middle: Image demosaicked with our sparse coding baseline without non-local prior. Bottom: demosaicking with sparse coding and non-local prior. The reconstruction does not exhibit any artefact on this image which is notoriously difficult for demosaicking.

Participants: Makarand Tapaswi, Marc T. Law, Sanja Fidler.

Understanding videos such as TV series and movies requires analyzing who the characters are and what they are doing. We address the challenging problem of clustering face tracks based on their identity. Different from previous work in this area, we choose to operate in a realistic and difficult setting where: (i) the number of characters is not known a priori; and (ii) face tracks belonging to minor or background characters are not discarded.

To this end, we propose Ball Cluster Learning (BCL), a supervised approach to carve the embedding space into balls of equal size, one for each cluster (see Figure 10). The learned ball radius is easily translated to a stopping criterion for iterative merging algorithms. This gives BCL the ability to estimate the number of clusters as well as their assignment, achieving promising results on commonly used datasets. We also present a thorough discussion of how existing metric learning literature can be adapted for this task. This work has been published in [22].

7.4.2. Cross-task weakly supervised learning from instructional videos

Participants: Dimitri Zhukov, Jean-Baptiste Alayrac, Ramazan Gokberk Cinbis, David Fouhey, Ivan Laptev, Josef Sivic.

In [25], we investigate learning visual models for the steps of ordinary tasks using weak supervision via instructional narrations and an ordered list of steps instead of strong supervision via temporal annotations. At the heart of our approach is the observation that weakly supervised learning may be easier if a model shares components while learning different steps: “pour egg” should be trained jointly with other tasks involving “pour” and “egg”. We formalize this in a component model for recognizing steps and a weakly supervised learning framework that can learn this model under temporal constraints from narration and the list of steps. Past data does not permit systematic studying of sharing and so we also gather a new dataset, CrossTask, aimed at assessing cross-task sharing. Our experiments demonstrate that sharing across tasks im- proves performance, especially when done at the component level and that our component model can parse previously unseen tasks by virtue of its compositionality. Figure 11 illustrates the idea of sharing step components between different tasks.

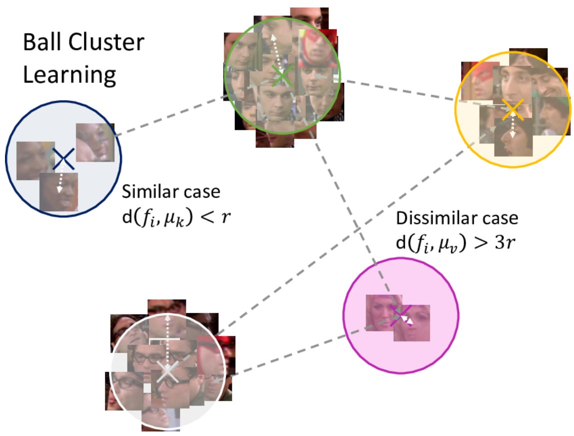

Figure 10. Ball Cluster Learning carves the feature space into balls of equal radius. The number of samples in the cluster does not affect the ball radius or minimum separation to other balls.

Figure 11. Our method begins with a collection of tasks, each con- sisting of an ordered list of steps and a set of instructional videos from YouTube. It automatically discovers both where the steps oc- cur and what they look like. To do this, it uses the order, narration and commonalities in appearance across tasks (e.g., the appear- ance of pour in both making pancakes and making meringue).

7.4.3. Leveraging the Present to Anticipate the Future in Videos

Participants: Antoine Miech, Ivan Laptev, Josef Sivic, Heng Wang, Lorenzo Torresani, Du Tran.

Anticipating actions before they are executed is crucial for a wide range of practical applications including autonomous driving and the moderation of live video streaming. While most prior work in this area requires partial observation of executed actions, in the paper we focus on anticipating actions seconds before they start (see Figure 12) Our proposed approach is the fusion of a purely anticipatory model with a complementary model constrained to reason about the present. In particular, the latter predicts present action and scene attributes, and reasons about how they evolve over time. By doing so, we aim at modeling action anticipation at a more conceptual level than directly predicting future actions. Our model outperforms previously reported methods on the EPIC-KITCHENS and Breakfast datasets. This paper was presented at the CVPR 2019 precognition workshop [34] and ranked second at the EPIC-KITCHENS action anticipation challenge.

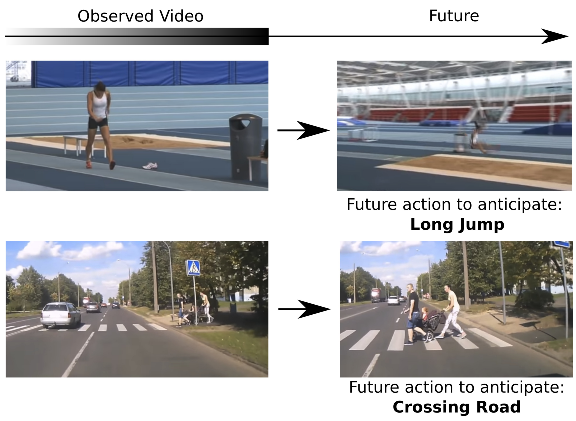

Figure 12. Examples of action anticipation in which the goal is to anticipate future actions in videos seconds before they are performed.

7.4.4. HowTo100M: Learning a Text-Video Embedding by Watching Hundred Million Narrated Video Clips

Participants: Antoine Miech, Dimitri Zhukov, Jean-Baptiste Alayrac, Makarand Tapaswi, Ivan Laptev, Josef Sivic.

Learning text-video embeddings usually requires a dataset of video clips with manually provided captions. However, such datasets are expensive and time consuming to create and therefore difficult to obtain on a large scale. In this work, we propose instead to learn such embeddings from video data with readily available natural language annotations in the form of automatically transcribed narrations (see Figure 13). The contributions of this work are three-fold. First, we introduce HowTo100M: a large-scale dataset of 136 million video clips sourced from 1.22M narrated instructional web videos depicting humans performing and describing over 23k different visual tasks. Our data collection procedure is fast, scalable and does not require any additional manual annotation. Second, we demonstrate that a text-video embedding trained on this data leads to state-of-the-art results for text-to-video retrieval and action localization on instructional video datasets such as YouCook2 or CrossTask. Finally, we show that this embedding transfers well to other domains: fine-tuning on generic Youtube videos (MSR-VTT dataset) and movies (LSMDC dataset) outperforms models trained on these datasets alone. This work was presented at ICCV 2019 [16].

Figure 13. We learn a joint text-video embedding by watching millions of narrated video clips of people performing diverse visual tasks. The learned embedding transfers well to other instructional and non-instructional text-video datasets.

7.4.5. Are Large-Scale 3D Models Really Necessary for Accurate Visual Localization?

Participants: Akihiko Torii, Hajime Taira, Josef Sivic, Marc Pollefeys, Masatoshi Okutomi, Tomas Pajdla, Torsten Sattler.

Accurate visual localization is a key technology for autonomous navigation. 3D structure-based methods, as illustrated in Figure 14, employ 3D models of the scene to estimate the full 6 degree-of-freedom (DOF) pose of a camera very accurately. However, constructing (and extending) large-scale 3D models is still a significant challenge. In contrast, 2D image retrieval-based methods only require a database of geo-tagged images, which is trivial to construct and to maintain. They are often considered inaccurate since they only approximate the positions of the cameras. Yet, the exact camera pose can theoretically be recovered when enough relevant database images are retrieved. In [8], we demonstrate experimentally that large-scale 3D models are not strictly necessary for accurate visual localization. We create reference poses for a large and challenging urban dataset. Using these poses, we show that combining image-based methods with local reconstructions results in a higher pose accuracy compared to state-of-the-art structure-based methods, albeight at higher run-time costs. We show that some of these run-time costs can be alleviated by exploiting known database image poses. Our results suggest that we might want to reconsider the need for large-scale 3D models in favor of more local models, but also that further research is necessary to accelerate the local reconstruction process.

7.4.6. End-to-End Learning of Visual Representations from Uncurated Instructional Videos

Participants: Antoine Miech, Jean-Baptiste Alayrac, Lucas Smaira, Ivan Laptev, Josef Sivic, Andrew Zisserman.

Annotating videos is cumbersome, expensive and not scalable. Yet, many strong video models still rely on manually annotated data. With the recent introduction of the HowTo100M dataset, narrated videos now offer the possibility of learning video representations without manual supervision. In this work we propose a new learning approach, MIL-NCE, capable of addressing misalignments inherent to narrated videos (see Figure 15). With this approach we are able to learn strong video representations from scratch, without the need for any manual annotation. We evaluate our representations on a wide range of four downstream tasks over eight datasets: action recognition (HMDB-51, UCF-101, Kinetics-700), text-to-video retrieval (YouCook2, MSR-VTT), action localization (YouTube-8M Segments, CrossTask) and action segmentation (COIN). Our method outperforms all published self-supervised approaches for these tasks as well as several fully supervised baselines. This preprint [32] is currently under review.

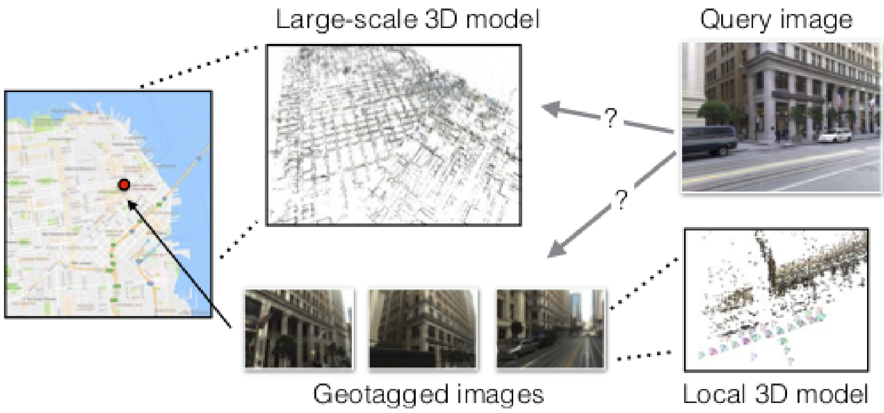

Figure 14. The state-of-the-art for large-scale visual localization. 2D image-based methods (bottom) use image retrieval and return the pose of the most relevant database image. 3D structure-based methods (top) use 2D-3D matches against a 3D model for camera pose estimation. Both approaches have been developed largely independently of each other and never compared properly before.

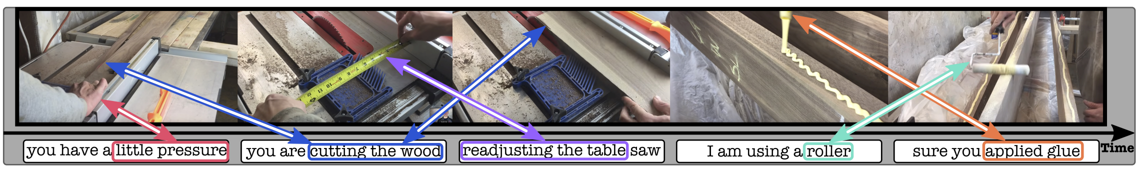

Figure 15. We describe an efficient approach to learn visual representations from highly misaligned and noisy narrations automatically extracted from instructional videos. Our video representations are learnt from scratch without relying on any manually annotated visual dataset yet outperform all self-supervised and many fully-supervised methods on several video recognition benchmarks.

7.4.7. Synthetic Humans for Action Recognition from Unseen Viewpoints

Participants: Gul Varol, Ivan Laptev, Cordelia Schmid, Andrew Zisserman.

In [35], the goal is to improve the performance of human action recognition for viewpoints unseen during training by using synthetic training data. Although synthetic data has been shown to be beneficial for tasks such as human pose estimation, its use for RGB human action recognition is relatively unexplored. We make use of the recent advances in monocular 3D human body reconstruction from real action sequences to automatically render synthetic training videos for the action labels. We make the following contributions: (i) we investigate the extent of variations and augmentations that are beneficial to improving performance at new viewpoints. We consider changes in body shape and clothing for individuals, as well as more action relevant augmentations such as non-uniform frame sampling, and interpolating between the motion of individuals performing the same action; (ii) We introduce a new dataset, SURREACT, that allows supervised training of spatio-temporal CNNs for action classification; (iii) We substantially improve the state-of-the-art action recognition performance on the NTU RGB+D and UESTC standard human action multi-view benchmarks; Finally, (iv) we extend the augmentation approach to in-the-wild videos from a subset of the Kinetics dataset to investigate the case when only one-shot training data is available, and demonstrate improvements in this case as well. Figure 16 presents an illustration of the approach.

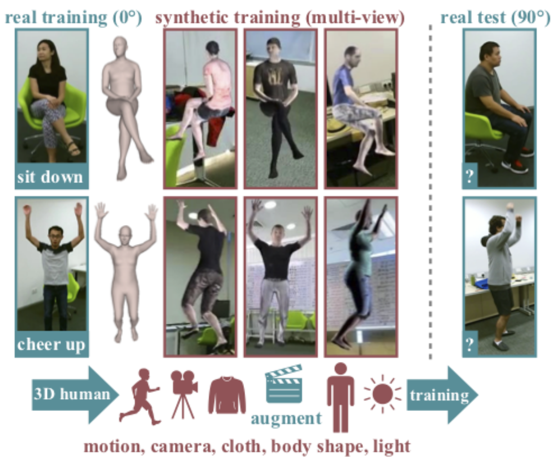

Figure 16. We estimate 3D shape from real videos and automatically render synthetic videos with action labels. We explore various augmentations for motions, viewpoints, and appearance. Training temporal CNNs with this data significantly improves the action recognition from unseen viewpoints.

7.5. Learning embodied representations and robotics

7.5.1. Roboticists and Reporters

Participants: Celine Pieters, Emmanuelle Danblon, Jean-Paul Laumond.

This paper reports on an experiment organized at the Cité des Sciences et de l'Industrie (CSI) of Paris in order to assess the importance of language in the representation and the integration of robots into the human culture. The experiment gathered specialized reporters and experts in robotics around a practical exercise of rhetoric. The objective of this work is to show that rhetoric is not a matter of communication, but a technique that allows to better understand the way roboticists understand their own discipline.

7.5.2. Robots

Participants: Jean-Paul Laumond, Denis Vidal.

What is a robot? How does it work? How is research progressing, what are the challenges and the economic and social questions posed by robotics in the twenty-first century? Today, as robots are becoming increasingly present in our professional, public and private lives, it is vital to understand their technological capabilities. We must more fully comprehend how they can help us and master their uses. Robots continue to fascinate us but our idea of them, stemming from literature and cinema, is often a purely imaginary one. This illustrated book accompanies the Robots exhibition at the Cité des sciences et de l'industrie. Figure 17 presents the front page of the exhibition.

Figure 17. Front page of the permanent exhibition at Cité des Sciences et de l'Industrie about Robotics.

7.5.3. Learning to Augment Synthetic Images for Sim2Real Policy Transfer

Participants: Alexander Pashevich, Robin Strudel, Igor Kalevatykh, Ivan Laptev, Cordelia Schmid.

Vision and learning have made significant progress that could improve robotics policies for complex tasks and environments. Learning deep neural networks for image understanding, however, requires large amounts of domain-specific visual data. While collecting such data from real robots is possible, such an approach limits the scalability as learning policies typically requires thousands of trials. In this paper [18], we attempt to learn manipulation policies in simulated environments. Simulators enable scalability and provide access to the underlying world state during training. Policies learned in simulators, however, do not transfer well to real scenes given the domain gap between real and synthetic data. We follow recent work on domain randomization and augment synthetic images with sequences of random transformations. Our main contribution is to optimize the augmentation strategy for sim2real transfer and to enable domain-independent policy learning. We design an efficient search for depth image augmentations using object localization as a proxy task. Given the resulting sequence of random transformations, we use it to augment synthetic depth images during policy learning. Our augmentation strategy is policy-independent and enables policy learning with no real images. We demonstrate our approach to significantly improve accuracy on three manipulation tasks evaluated on a real robot. Figure 18 presents an illustration of the approach.

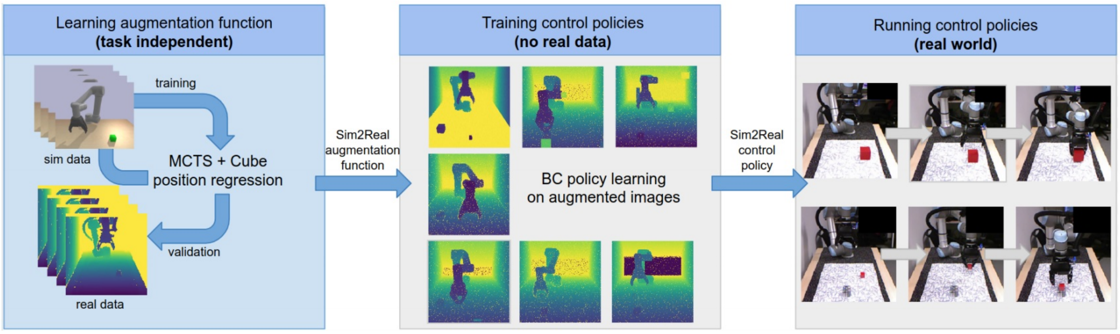

Figure 18. Overview of the method. Our contribution is the policy-independent learning of depth image augmentations (left). The resulting sequence of augmentations is applied to synthetic depth images while learning manipulation policies in a simulator (middle). The learned policies are directly applied to real robot scenes without finetuning on real images (right).

7.5.4. Learning to combine primitive skills: A step towards versatile robotic manipulation

Participants: Robin Strudel, Alexander Pashevich, Igor Kalevatykh, Ivan Laptev, Josef Sivic, Cordelia Schmid.

Manipulation tasks such as preparing a meal or assembling furniture remain highly challenging for robotics and vision. Traditional task and motion planning (TAMP) methods can solve complex tasks but require full state observability and are not adapted to dynamic scene changes. Recent learning methods can operate directly on visual inputs but typically require many demonstrations and/or task-specific reward engineering. In this paper [20], we aim to overcome previous limitations and propose a reinforcement learning (RL) approach to task planning that learns to combine primitive skills. First, compared to previous learning methods, our approach requires neither intermediate rewards nor complete task demonstrations during training. Second, we demonstrate the versatility of our vision-based task planning in challenging settings with temporary occlusions and dynamic scene changes. Third, we propose an efficient training of basic skills from few synthetic demonstrations by exploring recent CNN architectures and data augmentation. Notably, while all of our policies are learned on visual inputs in simulated environments, we demonstrate the successful transfer and high success rates when applying such policies to manipulation tasks on a real UR5 robotic arm. Figure 19 presents an illustration of the approach.

7.5.5. Monte-Carlo Tree Search for Efficient Visually Guided Rearrangement Planning

Participants: Sergey Zagoruyko, Yann Labbé, Igor Kalevatykh, Ivan Laptev, Justin Carpentier, Mathieu Aubry, Josef Sivic.

We address the problem of visually guided rearrangement planning with many movable objects, i.e., finding a sequence of actions to move a set of objects from an initial arrangement to a desired one, while relying on visual inputs coming from RGB camera. To do so, we introduce a complete pipeline relying on two key contributions. First, we introduce an efficient and scalable rearrangement planning method, based on a Monte-Carlo Tree Search exploration strategy. We demonstrate that because of its good trade-off between exploration and exploitation our method (i) scales well with the number of objects while (ii) finding solutions which require a smaller number of moves compared to the other state-of-the-art approaches. Note that on the contrary to many approaches, we do not require any buffer space to be available. Second, to precisely localize movable objects in the scene, we develop an integrated approach for robust multi-object workspace state estimation from a single uncalibrated RGB camera using a deep neural network trained only with synthetic data. We validate our multi-object visually guided manipulation pipeline with several experiments on a real

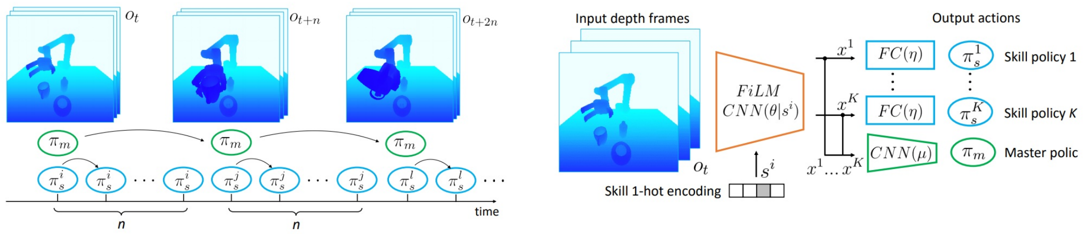

Figure 19. Illustration of our approach. (Left): Temporal hierarchy of master and skill policies. The master policy π_m is executed at a coarse interval of n time-steps to select among K skill policies $\pi_s^1 ... \pi_s^K$. Each skill policy generates control for a primitive action such as grasping or pouring. (Right): CNN architecture used for the skill and master policies.

UR-5 robotic arm by solving various rearrangement planning instances, requiring only 60 ms to compute the complete plan to rearrange 25 objects. In addition, we show that our system is insensitive to camera movements and can successfully recover from external perturbation. Figure 20 shows an example of the problems we consider. This work is under-review and an early pre-print is available [37].

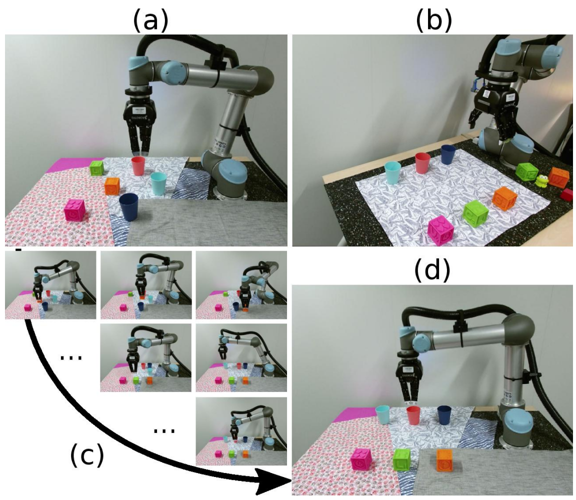

*Figure 20. **Visually guided rearrangement planning.** Given a source (a) and target (b) RGB images depicting a robot and multiple movable objects, our approach estimates the positions of objects in the scene without the need for explicit camera calibration and efficiently finds a sequence of robot actions (c) to re-arrange the scene into the target scene. Final object configuration after re-arrangement by the robot is shown in (d).*

7.5.6. Estimating the Center of Mass and the Angular Momentum Derivative for Legged Locomotion — A recursive approach

Participants: François Bailly, Justin Carpentier, Mehdi Benallegue, Bruno Watier, Philippe Soueres.

Estimating the center of mass position and the angular momentum derivative of legged systems is essential for both controlling legged robots and analyzing human motion. In this paper[4], a novel recursive approach to concurrently and accurately estimate these two quantities together is introduced. The proposed method employs kinetic and kinematic measurements from classic sensors available in robotics and biomechanics, to effectively exploits the accuracy of each measurement in the spectral domain. The soundness of the proposed approach is first validated on a simulated humanoid robot, where ground truth data is available, against an Extend Kalman Filter. The results demonstrate that the proposed method reduces the estimation error on the center of mass position with regard to kinematic estimation alone, whereas at the same time, it provides an accurate estimation of the derivative of angular momentum. Finally, the effectiveness of the proposed method is illustrated on real measurements, obtained from walking experiments with the HRP-2 humanoid robot. Figure 21 presents an illustration of the approach.

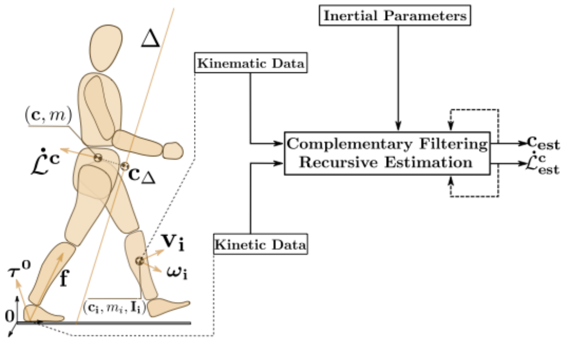

Figure 21. Illustration of the measurement apparatus. The several physical quantities involved in the estimation framework are displayed, as well as a simplified sketch of the estimation algorithm.

7.5.7. Dynamics Consensus between Centroidal and Whole-Body Models for Locomotion of Legged Robots

Participants: Rohan Budhiraja, Justin Carpentier, Nicolas Mansard.

It is nowadays well-established that locomotion can be written as a large and complex optimal control problem. Yet, current knowledge in numerical solver fails to directly solve it. A common approach is to cut the dimensionality by relying on reduced models (inverted pendulum, capture points, centroidal). However it is difficult both to account for whole-body constraints at the reduced level and also to define what is an acceptable trade-off at the whole-body level between tracking the reduced solution or searching for a new one. The main contribution of this paper [9] is to introduce a rigorous mathematical framework based on the Alternating Direction Method of Multipliers, to enforce the consensus between the centroidal state dynamics at reduced and whole-body level. We propose an exact splitting of the whole-body optimal control problem between the centroidal dynamics (under-actuation) and the manipulator dynamics (full actuation), corresponding to a rearrangement of the equations already stated in previous works. We then describe with details how alternating descent is a good solution to implement an effective locomotion solver. We validate this approach in simulation with walking experiments on the HRP-2 robot. Figure 22 presents a resulting motion of the proposed approach.

7.5.8. The Pinocchio C++ library – A fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives

Participants: Justin Carpentier, Guilhem Saurel, Gabriele Buondonno, Joseph Mirabel, Florent Lamiraux, Olivier Stasse, Nicolas Mansard.

Figure 22. Walking sequence generated for HRP-2 robot using the proposed ADMM solver.

In this paper [10], we introduce Pinocchio, an open-source software framework that implements rigid body dynamics algorithms and their analytical derivatives. Pinocchio does not only include standard algorithms employed in robotics (e.g., forward and inverse dynamics) but provides additional features essential for the control, the planning and the simulation of robots. In this paper, we describe these features and detail the programming patterns and design which make Pinocchio efficient. We evaluate the performances against RBDL, another framework with broad dissemination inside the robotics community. We also demonstrate how the source code generation embedded in Pinocchio outperforms other approaches of state of the art. Figure 23 presents the logo of Pinocchio.

Pinocchio
Efficient and versatile rigid body dynamics algorithms

Figure 23. Logo of Pinocchio.

7.5.9. Crocoddyl: An Efficient and Versatile Framework for Multi-Contact Optimal Control

Participants: Carlos Mastalli, Rohan Budhiraja, Wolfgang Merkt, Guilhem Saurel, Bilal Hammoud, Maximilien Naveau, Justin Carpentier, Sethu Vijayakumar, Nicolas Mansard.

In this paper [31], we introduce Crocoddyl (Contact RObot COntrol by Differential DYnamic Library), an open-source framework tailored for efficient multi-contact optimal control. Crocoddyl efficiently computes the state trajectory and the control policy for a given predefined sequence of contacts. Its efficiency is due to the use of sparse analytical derivatives, exploitation of the problem structure, and data sharing. It employs differential geometry to properly describe the state of any geometrical system, e.g. floating-base systems. We have unified dynamics, costs, and constraints into a single concept-action-for greater efficiency and easy prototyping. Additionally, we propose a novel multiple-shooting method called Feasibility-prone Differential Dynamic Programming (FDDP). Our novel method shows a greater globalization strategy compared to classical Differential Dynamic Programming (DDP) algorithms, and it has similar numerical behavior to state-of-the-art multiple-shooting methods. However, our method does not increase the computational complexity typically encountered by adding extra variables to describe the gaps in the dynamics. Concretely, we propose two modifications to the classical DDP algorithm. First, the backward pass accepts infeasible state-control trajectories. Second, the rollout keeps the gaps open during the early "exploratory" iterations (as expected in multiple-shooting methods). We showcase the performance of our framework using different tasks. With our method, we can compute highly-dynamic maneuvers for legged robots (e.g. jumping, front-flip) in the order of milliseconds. Figure 24 presents a resulting motion of the proposed approach.

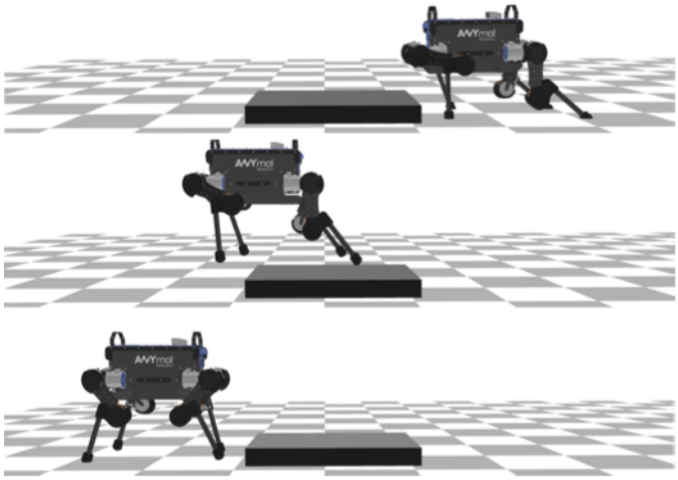

Figure 24. Crocoddyl: an efficient and versatile framework for multi-contact optimal control. Highly-dynamic maneuvers needed to traverse an obstacle with the ANYmal robot.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. MSR-Inria joint lab: Image and video mining for science and humanities (Inria)

Participants: Yana Hasson, Ivan Laptev, Jean Ponce, Josef Sivic, Dimitri Zhukov, Cordelia Schmid [Inria Thoth].

This collaborative project brings together the WILLOW and LEAR project-teams with MSR researchers in Cambridge and elsewhere. The concept builds on several ideas articulated in the 2020 Sciencea report, including the importance of data mining and machine learning in computational science. Rather than focusing only on natural sciences, however, we propose here to expand the breadth of e-science to include humanities and social sciences. The project we propose will focus on fundamental computer science research in computer vision and machine learning, and its application to archaeology, cultural heritage preservation, environmental science, and sociology, and it will be validated by collaborations with researchers and practitioners in these fields.

In October 2013 a new agreement has been signed for 2013-2017 with the research focus on automatic understanding of dynamic video content. Recent studies predict that by 2018 video will account for 80-90% of traffic on the Internet. Automatic understanding and interpretation of video content is a key enabling factor for a range of practical applications such as organizing and searching home videos or content aware video advertising. For example, interpreting videos of "making a birthday cake" or "planting a tree" could provide effective means for advertising products in local grocery stores or garden centers. The goal of this project is to perform fundamental computer science research in computer vision and machine learning in order to enhance the current capabilities to automatically understand, search and organize dynamic video content.

In 2018 a new agreement has been signed with a new focus on video understanding for personal assistants. The scientific objectives are to develop models, representations and learning algorithms for (i) automatic understanding of task-driven complex human activities from videos narrated with natural language in order to (ii) give people instructions in a new environment via an augmented reality device such as the Microsoft HoloLens. Besides the clear scientific interest of automatically understanding human activities in video streams, the main high-impact motivation of this project it to develop virtual assistants that may guide a child through simple games to improve his/her manipulation and language skills; help an elderly person to achieve everyday tasks; or facilitate the training of a new worker for highly-specialized machinery maintenance.

8.1.2. Louis Vuitton/ENS chair on artificial intelligence

Participants: Ivan Laptev, Jean Ponce, Josef Sivic.

The scientific chair Louis Vuitton - École normale supérieure in Artificial Intelligence has been created in 2017 and inaugurated on April 12, 2018 by the ENS Director Marc Mézard and the LV CEO Michael Burke. The goal of the chair is to establish a close collaboration between LV and ENS in the area of Artificial Intelligence. The chair enjoys the generous annual contribution of 200K Euros provided by LV in support of research activities in statistical learning and computer vision. In particular, the chair supports the costs of researchers, students, missions, computational resources as well as seminars and meetings, including the two days of meeting annually organized by LV and ENS. During 2018 ENS and LV have organized several joint meetings with the participation of researchers from SIERRA and WILLOW teams. The chair has also supported the hiring of one PhD student at the WILLOW team, missions to conferences and international research labs as well as data collection for research projects.

8.2. Bilateral Grants with Industry

8.2.1. Facebook AI Research Paris: Weakly-supervised interpretation of image and video data (Inria)

Participants: Jean Ponce, Minsu Cho, Ivan Laptev, Josef Sivic.

We will develop in this project (Facebook gift) new models of image and video content, as well as new recognition architectures and algorithms, to address the problem of understanding the visual content of images and videos using weak forms of supervision, such as the fact that multiple images contain instances of the same objects, or the textual information available in television or film scripts.

8.2.2. Google: Structured learning from video and natural language (Inria)

Participants: Simon Lacoste-Julien, Ivan Laptev, Josef Sivic.

People can easily learn how to change a flat tire of a car or assemble an IKEA shelve by observing other people doing the same task, for example, by watching a narrated instruction video. In addition, they can easily perform the same task in a different context, for example, at their home. This involves advanced visual intelligence abilities such as recognition of objects and their function as well as interpreting sequences of human actions that achieve a specific task. However, currently there is no artificial system with a similar cognitive visual competence. The goal of this proposal is to develop models, representations and learning algorithms for automatic understanding of complex human activities from videos narrated with natural language.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. PRAIRIE

Participants: Ivan Laptev, Jean-Paul Laumond, Jean Ponce, Josef Sivic.

The Prairie Institute (PaRis AI Research InstitutE) is one of the four French Institutes for Interdisciplinary Artificial Intelligence Research (3IA), which were created as part of the national French initiative on AI announced by President Emmanuel Macron on May 29, 2018. It brings together five academic partners (CNRS, Inria, Institut Pasteur, PSL University, and University of Paris) as well as 17 industrial partners, large corporations which are major players in AI at the French, European and international levels, as well as 45 Chair holders, including four of the members of WILLOW (Laumond, Laptev, Ponce, Sivic). Ponce is the scientific director of PRAIRIE.

9.1.2. DGA - RAPID project DRAAF

Participant: Ivan Laptev.

DGA DRAAF is a two-year collaborative effort with University of Caen (F. Jurie) and the industrial partner EVITECH (P. Bernas) focused on modelling and recognition of violent behaviour in surveillance videos. The project aims to develop image recognition models and algorithms to automatically detect weapons, gestures and actions using recent advances in computer vision and deep learning to provide an affordable real-time solution reducing effects of threats in public places.

9.2. European Initiatives

9.2.1. IMPACT: Intelligent machine perception

Participants: Josef Sivic, Jean Ponce, Ivan Laptev.

IMPACT is a 5-year collaborative project with Czech Technical University, Center for Robotics, Informatics and Cybernetics (CIIRC) (2017-2022). The IMPACT project focuses on fundamental and applied research in computer vision, machine learning and robotics to develop machines that learn to perceive, reason, navigate and interact with complex dynamic environments. For example, people easily learn how to change a flat tire of a car or perform resuscitation by observing other people doing the same task. This involves advanced visual intelligence abilities such as interpreting sequences of human actions that manipulate objects to achieve a specific task. Currently, however, there is no artificial system with a similar level of cognitive visual competence. Breakthrough progress in intelligent machine perception will have profound implications on our everyday lives as well as science and commerce, with smart assistive robots that automatically learn new skills from the Internet, safer cars that autonomously navigate in difficult changing conditions, or intelligent glasses that help people navigate never seen before environments.

9.3. International Initiatives

9.3.1. Associate team GAYA

Participants: Jean Ponce, Matthew Trager.

GAYA is a joint research team bringing together two Inria project-teams (Thoth, Grenoble and WILLOW, Paris) and Carnegie Mellon University, USA. It focuses on two research themes: (i) semantic structured interpretation of videos, and (ii) studying the geometric properties of object shapes to enhance state-of-the-art object recognition approaches.

Interpreting videos semantically in a general setting, involving various types of video content like home video clips, news broadcasts, feature films, which contain a lot of clutter, non-rigid motion, many “actors” performing actions, person-object and person-person interactions, varying viewpoints, is challenging. This task is being examined increasingly over the past decade, with the availability of large video resources, e.g., YouTube. Despite this progress, an effective video representation for recognizing actions is still missing. To address this critical challenge, we propose a joint optimization framework, wherein we learn the video representation and also develop models for action recognition. Specifically, we aim to exploit the spatio-temporal relations among pixels in a video through graphical models and novel deep learning feature representations.

The second research theme explores geometric aspects of computer vision, in particular how to model three-dimensional objects from their two-dimensional projections, and how the appearance of these objects evolves with changes in viewpoint. Beyond its theoretical interest, this work is critical for developing object recognition algorithms that take into account the three-dimensional nature of the visual world and go beyond the template-matching approaches dominant today. Duality is an important concept in this area, and we are investigating its application to the construction of visual hulls as well as the characterization of the topology of image contours using the Gauss map. Existing results are essentially limited to the Euclidean setting, and we are investigating their generalization to the general projective case.

Partners: CMU (Deva Ramanan, Martial Hebert, Abhinav Gupta, Gunnar Sigurdsson), Inria Thoth (Cordelia Schmid, Karteek Alahari, Pavel Tokmakov).

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Pierre-Yves Masse (post-doc, Czech Technical University) spent 50% of his time at Sierra (F. Bach) and Willow teams as a visiting post-doc within the framework of collaboration with the Intelligent Machine Perception project lead by J. Sivic at the Czech Technical University in Prague.
- Vladimir Petrik spent October - Janurary 2020 in Willow as a visiting post-doc within the framework of collaboration with the Intelligent Machine Perception project.
- Mircea Cimpoi spent three weeks in March 2019 in Willow as a visiting post-doc within the framework of collaboration with the Intelligent Machine Perception project.

9.4.1.1. Internships

- Anna Kukleva (Master student, University of Bonn) spent six months in the Willow team working on her Master project under supervision of M. Tapaswi and I. Laptev.

9.4.2. Visits to International Teams

9.4.2.1. Explorer programme

- J.Ponce, multiple visits to CMU's Robotics Institute within the framework of the Gaia associated team

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. Member of the Organizing Committees

- I. Laptev was co-organizedr of BMVA Symposium on Video Understanding in London, September 2019.
- P.-L. Guhur was co-organizer of Junior Conference on Data Science and Engineering at Paris-Saclay, September 2019.

10.1.2. Scientific Events: Selection

10.1.2.1. Area chairs

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019 (I. Laptev).
- Neural Information Processing Systems (NeurIPS), 2019 (J. Sivic).

10.1.2.2. Member of the Conference Program Committees / Reviewer

- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019 (A. Miech, M. Tapaswi).
- IEEE/CVF International Conference on Computer Vision (ICCV), 2019 (A. Miech, M. Tapaswi).
- Neural Information Processing Systems (NeurIPS), 2019 (I. Laptev, I. Rocco, M. Tapaswi).
- British Machine Vision Conference (BMVC), 2019 (I. Rocco).
- IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019 (I. Laptev, J. Carpentier, J. Sivic, Y. Labbe).
- IEEE International Conference on Robotics and Automation (ICRA), 2019 (J. Carpentier, Y. Labbe, J. Sivic).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- International Journal of Computer Vision (I. Laptev, J. Ponce, J. Sivic).
- IEEE Transactions on Pattern Analysis and Machine Intelligence (I. Laptev, J. Sivic).
- Foundations and Trends in Computer Graphics and Vision (J. Ponce).

10.1.3.2. Reviewer - Reviewing Activities

- International Journal of Computer Vision (M. Tapaswi).
- IEEE Transactions on Pattern Analysis and Machine Intelligence (I. Rocco).
- IEEE Transactions on Robotics (J. Carpentier).
- IEEE Robotics and Automation Letters (J. Carpentier).
- Plos One (M. Tapaswi).

10.1.4. Invited Talks

- I. Laptev, invited talk, AI Video Summit, FAIR, Los Angeles, June 2019.
- I. Laptev, invited talk, ActivityNet CVPR Workshop, Long Beach, June 2019.
- I. Laptev, keynote, AIST, Kazan, July 2019.
- I. Laptev, invited talk, ELLIS Workshop, San Sebastian, September 2019.
- I. Laptev, seminar, KTH, Stockholm, September 2019.
- I. Laptev, invited talk, BMVA Symposium on Video Understanding, London, September 2019.
- I. Laptev, invited talk, Qualcomm-UvA Deep Learning Seminars, Amsterdam, September 2019.
- I. Laptev, invited talk, Extreme Vision ICCV Workshop, Seoul, October 2019.
- I. Laptev, invited talk, CoVieW ICCV Workshop, Seoul, October 2019.
- I. Laptev, invited talk, Hands in Action ICCV Workshop, Seoul, October 2019.
- I. Laptev, keynote, DICTA, Perth, December 2019.
- I. Laptev, invited talk, SkolTech, Moscow, December 2019.
- J. Ponce, invited talk, Pavillon des Sciences CCI, Montbelliard, March 2019.
- J. Ponce, invited talk, France is AI, Paris, Oct. 2019.
- J. Ponce, invited talk, Ecole navale, Brest, Jan. 2019.
- J. Sivic, Keynote, European Conference on Mobile Robots (ECMR), Prague, September 2019.
- J. Sivic, Seminar on Weakly supervised learning for visual recognition, College de France, February 2019
- J. Sivic, Invited speaker, ML in PL conference, Warsaw, November 2019
- J. Sivic, Seminar, DeepMind, London, February 2019

- J. Sivic, Talk at invited workshop of the ELLIS computer vision program, San Sebastian, September 2019
- I. Rocco, invited talk, SMILE seminar at Télécom Paris, Paris, March 2019.
- I. Rocco, seminar, IMAGINE group at ENPC, Paris, April 2019.
- J-P. Laumond, invited talk, La robotique, Café Scientifique, Cité des sciences et de l'industrie, Paris, Janurary 2019.
- J-P. Laumond, invited talk, Une machine peut-elle apprendre ?, Les Découvrades, Toulouse, February 2019.
- J-P. Laumond, invited talk, La robotique : retour vers le réel, Théatre Tourski, Marseille, February 2019.
- J-P. Laumond, invited talk, Robots, Exhibition opening, Cité des sciences et de l'industrie, Paris, March 2019.
- J-P. Laumond, scientific talk, Robot Motion, Workshop on Robotics and Art, IEEE ICRA, Montreal, May 2019.
- J-P. Laumond, scientific talk, Wording Robotics, Workshop on Rhetoric and Robotics, IEEE ICRA, Montreal, May 2019.
- J-P. Laumond, scientific talk, Robotique et Biomécanique, Keynote, Congrès de physiologie, Montpellier, June 2019.
- J-P. Laumond, scientific talk, Robotics and AI, Istituto di Robotica e Macchine Intelligenti Grand Opening, Roma, October 2019.
- J. Carpentier, scientific talk, La robotique envisagée comme une science des données, CNRS Grenoble, January 2019.
- J. Carpentier, scientific talk, La robotique envisagée comme une science des données, Inria Grenoble, February 2019.
- J. Carpentier, scientific talk, Analytical Derivatives of Rigid Body Dynamics algorithms, RSS, June 2019.
- J. Carpentier, scientific talk, La robotique : la fabrique du mouvement artificiel, Journées DGDI Inria, Rocquencourt, November 2019.
- M. Tapaswi, scientific talk, Understanding Humans and the Stories they Tell, LIMSI Paris, February 2019.
- M. Tapaswi, scientific talk, Understanding Humans and the Stories they Tell, UPC Barcelona, August 2019.

10.1.5. Leadership within the Scientific Community

- Member of the advisory board for the IBM Watson AI Xprize (J. Ponce).
- Member of the steering committee of France AI (J. Ponce).
- Member of advisory board, Computer Vision Foundation (J. Sivic).
- Board Member Deputy, European Laboratory for Learning and Intelligent Systems – ELLIS (J. Sivic)

10.1.6. Scientific Expertise

- J. Ponce, coordinator of the AI theme for the joint French-American Committee on Science and Technology, 2018-.
- I. Laptev, head of scientific board at VisionLabs, 2019-.

10.1.7. Research Administration

- Member, Bureau du comité des projets, Inria, Paris (J. Ponce)

- Member, Scientific academic council, PSL Research University (J. Ponce)
- Member, Research representative committee, PSL Research University (J. Ponce).
- Member of Inria Commission de développement technologique (CDT), 2012-2018 (J. Sivic).
- Member of the Hiring Committee at Computer Science department, École normale supérieure (I. Laptev).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Master: M. Aubry, K. Alahari, I. Laptev and J. Sivic "Introduction to computer vision", M1, Ecole normale supérieure, 36h.
- Master: I. Laptev, J. Ponce and J. Sivic (together with C. Schmid, Inria Grenoble), "Object recognition and computer vision", M2, Ecole normale superieure, and MVA, Ecole normale superieure de Cachan, 36h.
- Master: J-P. Laumond and J. Carpentier, "Robotics", M1 MPRI, Ecole normale supérieure and Ecole normale superieure de Cachan, 48h.
- Master: J-P. Laumond, "Robotics", Ecole des Mines de Paris, 4h.
- Master: J. Sivic, three lectures (3 x 1.5h) in the 3D computer vision class of V. Hlavac at Charles University in Prague.
- License: P.L. Guhur, "Developping web applications with React Js", L3, Université Grenoble Alpes, 30h.
- J. Ponce co-organized the PRAIRIE AI Summer School in Paris, 2019.
- Bachelor: J. Ponce, "Inroduction to computer vision" MS level class, NYU Center for Data Science, Fall 2019

10.2.2. Supervision

PhD in progress : Vo Van Huy, started in Dec 2018, J. Ponce.

PhD in progress : Pierre-Louis Guhur, “Learning Visual Language Manipulation”, started in Oct 2019, I. Laptev and C. Schmid.

PhD in progress : Aamr El Kazdadi, started in Oct 2019, J. Carpentier and J. Ponce.

PhD in progress : Bruno Lecouat, started in Sept 2019, J. Ponce and J. Mairal (Inria Grenoble).

PhD in progress : Robin Strudel, “Learning and transferring complex robot skills from human demonstrations”, started in Oct 2018, I. Laptev, C. Schmid and J. Sivic.

PhD in progress : Yann Labbe, “Generalizing robotic sensorimotor skills to new tasks and environments”, started in Oct 2018, J. Sivic and I. Laptev.

PhD in progress : Minttu Alakuijala, started in Feb 2019, J. Ponce and C. Schmid.

PhD in progress : Thomas Eboli, started in Oct 2017, J. Ponce.

PhD in progress : Zongmian Li, “Learning to manipulate objects from instructional videos”, started in Oct 2017, I. Laptev, J. Sivic and N. Mansard (LAAS/CNRS, Toulouse).

PhD in progress : Yana Hasson, “Reconstruction and recognition of hand-object manipulations”, started in Nov 2017, I. Laptev and C. Schmid.

PhD in progress : Ronan Riochet, “Reconstruction and recognition of hand-object manipulations”, started in Oct 2017, E. Dupoux, I. Laptev and C. Schmid.

PhD in progress : Dmitry Zhukov, “Learning from instruction videos for personal assistants”, started in Oct 2017, I. Laptev and J. Sivic.

PhD in progress : Ignacio Rocco, “Estimating correspondence between images via convolutional neural networks”, started in Jan 2017, J. Sivic, R. Arandjelovic (Google DeepMind).

PhD in progress : Antoine Miech, “Understanding long-term temporal structure of videos”, started in Oct 2016, I. Laptev, J. Sivic.

PhD defended : Gul Varol, “Deep learning methods for video interpretation”, graduated in May 2019, I. Laptev, C. Schmid.

PhD defended : Julia Peyre, “Learning to reason about scenes from images and language”, graduated in August 2019, C. Schmid, I. Laptev, J. Sivic.

PhD defended : Theophile Dalens, “Learning to analyze and reconstruct architectural scenes”, graduated in September 2019, M. Aubry and J. Sivic.

10.2.3. Juries

- PhD thesis committee:
  - Dmitry Ulyanov, SkolTech, Russia, 2019 (I. Laptev, rapporteur).
  - Judith Bütepage, KTH, Sweden, 2019 (I. Laptev, rapporteur).
  - Alexander Richard, University of Bonn, Germany, 2019 (I. Laptev, rapporteur).
  - Rıza Alp Güler, Université Paris Saclay, France, 2019 (I. Laptev, rapporteur).
  - Tatiana Shpakova, l’École Normale Supérieure, France, 2019 (I. Laptev, examinateur).
  - Konstantin SHMELKOV, Universite Grenoble Alpes, (J. Sivic, examinateur)
  - Sanjeel PAREKH, Télécom Paris, (J. Sivic, rapporteur)
  - Filip RADENOVIC, Czech Technical University (J. Sivic, rapporteur)
  - Nathan PIASCO, Université Bourgogne Franche-Comte (J. Sivic, examinateur)
- HDR committee:
  - Karteek Alahari, HDR defense, Jan. 2019 (J. Ponce)
  - Mathieu Aubry, HDR defense, Aug. 2019 (J. Ponce)

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

J.P. Laumond is the scientific curator of the permanent exhibition at Cité des Sciences et de l’Industrie. The aim of this exhibition is to help the general audience to understand the concepts of robotics. Indeed, the notion of robotics today is packed with many preconceived notions, phobias, and utopias, all fed by literature and a rich film culture. The real challenge of the exhibition is the presentation of authentic working robots that raises awareness of our relationship to these singular machines. How do they work? What are they for? What are their performances today and what will they be tomorrow? The exhibition lays bare the actual capabilities of robots and provides insight into the current issues.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] T. DALENS.*Learnable factored image representations for visual discovery*, Ecole Normale Superieure de Paris - ENS Paris, September 2019, https://hal.inria.fr/tel-02296150

[2] J. PEYRE.*Learning to detect visual relations*, Ecole Normale Superieure de Paris - ENS Paris, August 2019, https://hal.inria.fr/tel-02332673

[3] G. VAROL.*Learning human body and human action representations from visual data*, Ecole Normale Supérieure (ENS) ; ED 386 : École doctorale de sciences mathématiques de Paris centre, UPMC, May 2019, https://hal.inria.fr/tel-02266593

Articles in International Peer-Reviewed Journal

[4] F. BAILLY, J. CARPENTIER, M. BENALLEGUE, B. WATIER, P. SOUÈRES.*Estimating the Center of Mass and the Angular Momentum Derivative for Legged Locomotion — A recursive approach*, in "IEEE Robotics and Automation Letters", October 2019, vol. 4, n^{o} 4, p. 4155-4162 [*DOI :* 10.1109/LRA.2019.2931200], https://hal.archives-ouvertes.fr/hal-02058890

[5] T. DALENS, M. AUBRY, J. SIVIC.*Bilinear image translation for temporal analysis of photo collections*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", October 2019 [*DOI :* 10.1109/TPAMI.2019.2950317], https://hal.inria.fr/hal-02452091

[6] P. FERNBACH, S. TONNEAU, O. STASSE, J. CARPENTIER, M. TAÏX.*C-CROC: Continuous and Convex Resolution of Centroidal dynamic trajectories for legged robots in multi-contact scenarios*, in "IEEE Transactions on Robotics", January 2020, p. 1-16, forthcoming [*DOI :* 10.1109/TRO.2020.2964787], https://hal.laas.fr/hal-01894869

[7] X. S. HU, S. ZAGORUYKO, N. KOMODAKIS.*Exploring Weight Symmetry in Deep Neural Networks*, in "Computer Vision and Image Understanding", August 2019, vol. 187, https://arxiv.org/abs/1812.11027 [*DOI :* 10.1016/J.CVIU.2019.07.006], https://hal.archives-ouvertes.fr/hal-01978633

[8] A. TORII, H. TAIRA, J. SIVIC, M. POLLEFEYS, M. OKUTOMI, T. PAJDLA, T. SATTLER.*Are Large-Scale 3D Models Really Necessary for Accurate Visual Localization?*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", September 2019, vol. XX, p. 1-14 [*DOI :* 10.1109/TPAMI.2019.2941876], https://hal.inria.fr/hal-02448029

International Conferences with Proceedings

[9] R. BUDHIRAJA, J. CARPENTIER, N. MANSARD.*Dynamics Consensus between Centroidal and Whole-Body Models for Locomotion of Legged Robots*, in "ICRA 2019 - IEEE International Conference on Robotics and Automation", Montreal, Canada, May 2019, https://hal.laas.fr/hal-01875031

[10] J. CARPENTIER, G. SAUREL, G. BUONDONNO, J. MIRABEL, F. LAMIRAUX, O. STASSE, N. MANSARD.*The Pinocchio C++ library – A fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives*, in "SII 2019 - International Symposium on System Integrations", Paris, France, January 2019, https://hal.laas.fr/hal-01866228

[11] *Best Paper*
H. CISNEROS, J. SIVIC, T. MIKOLOV.*Evolving Structures in Complex Systems*, in "SSCI 2019 - IEEE Symposium Series on Computational Intelligence", Xiamen, China, December 2019, https://arxiv.org/abs/1911.01086 - IEEE Symposium Series on Computational Intelligence 2019 (IEEE SSCI 2019), https://hal.inria.fr/hal-02448134.

[12] Y. DING, J. YANG, J. PONCE, H. KONG.*An Efficient Solution to the Homography-Based Relative Pose Problem With a Common Reference Direction*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, https://hal.archives-ouvertes.fr/hal-02266544

[13] M. DUSMANU, I. ROCCO, T. PAJDLA, M. POLLEFEYS, J. SIVIC, A. TORII, T. SATTLER.*D2-Net: A Trainable CNN for Joint Detection and Description of Local Features*, in "CVPR 2019 - IEEE Conference on Computer Vision and Pattern Recognition", Long Beach, United States, June 2019, https://arxiv.org/abs/1905.03561 - Accepted at CVPR 2019, https://hal.archives-ouvertes.fr/hal-02438461

[14] Y. HASSON, G. VAROL, D. TZIONAS, I. KALEVATYKH, M. J. BLACK, I. LAPTEV, C. SCHMID.*Learning joint reconstruction of hands and manipulated objects*, in "CVPR 2019 - IEEE Conference on Computer Vision and Pattern Recognition", Long Beach, United States, IEEE, June 2019, p. 1-14, https://hal.archives-ouvertes.fr/hal-02429093

[15] J. LEE, D. KIM, J. S. PONCE, B. HAM.*SFNet: Learning Object-aware Semantic Correspondence*, in "CVPR 2019 - Computer Vision and Pattern Recognition", Longbeach, United States, June 2019, https://hal.archives-ouvertes.fr/hal-02088666

[16] A. MIECH, D. ZHUKOV, J.-B. ALAYRAC, M. TAPASWI, I. LAPTEV, J. SIVIC.*HowTo100M: Learning a Text-Video Embedding by Watching Hundred Million Narrated Video Clips*, in "ICCV 2019 - International Conference on Computer Vision", Séoul, South Korea, International Conference on Computer Vision, October 2019, https://arxiv.org/abs/1906.03327 - Accepted at ICCV 2019, https://hal.archives-ouvertes.fr/hal-02433497

[17] J. MIN, J. LEE, J. PONCE, M. CHO.*Hyperpixel Flow: Semantic Correspondence with Multi-layer Neural Features*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, https://hal.archives-ouvertes.fr/hal-02267044

[18] A. PASHEVICH, R. STRUDEL, I. KALEVATYKH, I. LAPTEV, C. SCHMID.*Learning to Augment Synthetic Images for Sim2Real Policy Transfer*, in "IROS 2019 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Macao, China, November 2019, p. 1-6, https://arxiv.org/abs/1903.07740 - 7 pages, https://hal.archives-ouvertes.fr/hal-02273326

[19] J. PEYRE, I. LAPTEV, C. SCHMID, J. SIVIC.*Detecting unseen visual relations using analogies*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, https://arxiv.org/abs/1812.05736v3 , https://hal.archives-ouvertes.fr/hal-01975760

[20] R. STRUDEL, A. PASHEVICH, I. KALEVATYKH, I. LAPTEV, J. SIVIC, C. SCHMID.*Learning to combine primitive skills: A step towards versatile robotic manipulation*, in "ICRA", Paris, France, May 2020, https://arxiv.org/abs/1908.00722 - 11 pages, https://hal.archives-ouvertes.fr/hal-02274969

[21] H. TAIRA, I. ROCCO, J. SEDLAR, M. OKUTOMI, J. SIVIC, T. PAJDLA, T. SATTLER, A. TORII.*Is This The Right Place? Geometric-Semantic Pose Verification for Indoor Visual Localization*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, https://arxiv.org/abs/1908.04598 , https://hal.archives-ouvertes.fr/hal-02438468

[22] M. TAPASWI, M. T. LAW, S. FIDLER.*Video Face Clustering with Unknown Number of Clusters*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, https://arxiv.org/abs/1908.03381 - Accepted to ICCV 2019, code and data at https://github.com/makarandtapaswi/BallClustering_ICCV2019, https://hal.inria.fr/hal-02435407

[23] M. TRAGER, M. HEBERT, J. PONCE.*Coordinate-Free Carlsson-Weinshall Duality and Relative Multi-ViewGeometry*, in "CVPR 2019 - IEEE Conference on Computer Vision and Pattern Recognition", Long Beach, United States, June 2019, https://hal.inria.fr/hal-01676732

[24] Y. ZHANG, Y. GAO, S. GU, Y. GUO, M. LIU, Z. SUN, Z. HOU, H. YANG, Y. WANG, J. YANG, J. PONCE, H. KONG.*Build your own hybrid thermal/EO camera for autonomous vehicle*, in "ICRA 2019 - IEEE International Conference on Robotics and Automation", Montreal, Canada, May 2019, to appear in the Proc. of the IEEE International Conference on Robotics and Automation, 2019, https://hal.inria.fr/hal-02051880

[25] D. ZHUKOV, J.-B. ALAYRAC, R. G. CINBIS, D. F. FOUHEY, I. LAPTEV, J. SIVIC.*Cross-task weakly supervised learning from instructional videos*, in "CVPR 2019 - IEEE Conference on Computer Vision and Pattern Recognition", Long Beach, CA, United States, June 2019, https://hal.archives-ouvertes.fr/hal-02434806

Other Publications

[26] G. CHÉRON, A. OSOKIN, I. LAPTEV, C. SCHMID.*Modeling Spatio-Temporal Human Track Structure for Action Localization*, January 2019, https://arxiv.org/abs/1806.11008 - working paper or preprint, https://hal.inria.fr/hal-01979583

[27] M. HAHN, N. RUIZ, J.-B. ALAYRAC, I. LAPTEV, J. M. REHG.*Learning to Localize and Align Fine-Grained Actions to Sparse Instructions*, January 2019, https://arxiv.org/abs/1809.08381 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01979719

[28] B. KIM, J. PONCE, B. HAM.*Deformable Kernel Networks for Joint Image Filtering*, January 2020, working paper or preprint [*DOI :* 10.1007/S11263-018-1074-6], https://hal.archives-ouvertes.fr/hal-01857016

[29] B. LECOUAT, J. PONCE, J. MAIRAL.*Revisiting Non Local Sparse Models for Image Restoration*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02414291

[30] J. LEE, D. KIM, W. LEE, J. PONCE, B. HAM.*Learning Semantic Correspondence Exploiting an Object-level Prior*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02433226

[31] C. MASTALLI, R. BUDHIRAJA, W. MERKT, G. SAUREL, B. HAMMOUD, M. NAVEAU, J. CARPENTIER, S. VIJAYAKUMAR, N. MANSARD.*Crocoddyl: An Efficient and Versatile Framework for Multi-Contact Optimal Control*, September 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02294059

[32] A. MIECH, J.-B. ALAYRAC, L. SMAIRA, I. LAPTEV, J. SIVIC, A. ZISSERMAN.*End-to-End Learning of Visual Representations from Uncurated Instructional Videos*, January 2020, https://arxiv.org/abs/1912.06430 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02433504

[33] A. MIECH, I. LAPTEV, J. SIVIC.*Learning a Text-Video Embedding from Incomplete and Heterogeneous Data*, January 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01975102

[34] A. MIECH, I. LAPTEV, J. SIVIC, H. WANG, L. TORRESANI, D. TRAN.*Leveraging the Present to Anticipate the Future in Videos*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02433506

[35] G. VAROL, I. LAPTEV, C. SCHMID, A. ZISSERMAN.*Synthetic Humans for Action Recognition from Unseen Viewpoints*, January 2020, https://arxiv.org/abs/1912.04070 - working paper or preprint, https://hal.inria.fr/hal-02435731

[36] T.-H. VU, A. OSOKIN, I. LAPTEV.*Tube-CNN: Modeling temporal evolution of appearance for object detection in video*, January 2019, https://arxiv.org/abs/1812.02619 - 13 pages, 8 figures, technical report, https://hal.archives-ouvertes.fr/hal-01980339

[37] S. ZAGORUYKO, Y. LABBÉ, I. KALEVATYKH, I. LAPTEV, J. CARPENTIER, M. AUBRY, J. SIVIC.*Monte-Carlo Tree Search for Efficient Visually Guided Rearrangement Planning*, April 2019, https://arxiv.org/abs/1904.10348 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02108930