Activity Report Paris 2016

Edition: 2017-06-08

List of Inria's Research Teams

Project-Team ALPAGE

Large-scale deep linguistic processing

IN COLLABORATION WITH: Analyse Linguistique Profonde A Grande Echelle (ALPAGE)

IN PARTNERSHIP WITH:
Université Denis Diderot (Paris 7)

RESEARCH CENTER
Paris

THEME
Language, Speech and Audio

Table of contents

Project-Team ALPAGE

Creation of the Project-Team: 2008 January 01, end of the Project-Team: 2016 December 31

Keywords:

Computer Science and Digital Science:

3.1.1. - Modeling, representation
3.1.7. - Open data
3.2.2. - Knowledge extraction, cleaning
3.2.4. - Semantic Web
3.3.2. - Data mining
5.8. - Natural language processing
8.2. - Machine learning
8.4. - Natural language processing

Other Research Topics and Application Domains:

1.3.2. - Cognitive science
9.5.8. - Linguistics
9.5.10. - Digital humanities
9.7.1. - Open access
9.7.2. - Open data

1. Members

Research Scientists

Benoît Sagot [Team leader, Inria, Researcher]
Pierre Boullier [Inria, Senior Researcher (Emeritus)]
Laurent Romary [Inria, Senior Researcher, HDR]
Éric Villemonte de La Clergerie [Inria, Researcher]

Faculty Members

Lucie Barque [Univ. Paris XIII, Associate Professor]
Marie-Hélène Candito [Univ. Paris VII, Associate Professor]
Mathieu Constant [Univ. Paris Est, Associate Professor, until Aug 2016]
Benoît Crabbé [Univ. Paris VII, Associate Professor]
Laurence Danlos [Univ. Paris VII, Professor, HDR]
Djamé Seddah [Univ. Paris IV, Associate Professor]

Technical Staff

Noemie Faivre [Inria, until May 2016]
Luca Foppiano [Inria]
Pierre Magistry [Inria, until Jan 2016, granted by Caisse des Dépôts et Consignations]
Héctor Martínez Alonso [Inria, from Feb 2016]
Marie Puren [Inria]
Charles Riondet [Inria]
Stéphane Riou [Institut de Linguistique Française (CNRS), until Oct 2016]
Dorian Seillier [Inria, from May 2016]

PhD Students

Timothée Bernard [ENS Lyon]

Maximin Coavoux [Univ. Paris VII]
Daniel Dakota [Inria, Visiting PhD Student, from Oct 2016]
Marianne Djemaa [Univ. Paris VII, until Sep 2016, granted by ANR ASFALDA- project]
Mohamed Khemakhem [Inria, from Jun 2016]
Corentin Ribeyre [Univ. Paris VII, until Jan 2016]
Raphaël Salmon [Yseop and Univ. Paris VII, until Jan 2016, granted by CIFRE]

Administrative Assistant

Christelle Guiziou [Inria]

Others

Graziella Pastore [Inria, from Apr 2016 until Jul 2016]
Laura Ramirez Sanchez [Inria, Intern, from Jul 2016 until Sep 2016]
Adrien Roux [Inria, Intern, from Mar 2016 until Jul 2016]
Vincent Segonne [Inria, Intern, from Feb 2016 until Jul 2016]
De Zhao [Inria, Intern, from Jun 2016 until Jul 2016]

2. Overall Objectives

2.1. Overall Objectives

The Alpage team is specialised in **Language modelling**, **Computational linguistics** and **Natural Language Processing (NLP)**. These fields are of crucial importance for the new information society. Applications of this domain of research include the numerous technologies grouped under the term of 'language engineering'. This includes domains such as machine translation, question answering, information retrieval, information extraction, data mining, text simplification, automatic or computer-aided translation, automatic summarisation, foreign language reading and writing aid. From a more research-oriented point of view, experimental linguistics can be also viewed as an 'application' of NLP.

NLP, the domain of Alpage, is a **transdisciplinary** domain: it requires an expertise in formal and descriptive linguistics (to develop linguistic models of human languages), in computer science and algorithmics (to design and develop efficient programs that can deal with such models) and in applied mathematics (to automatically acquire linguistic or general knowledge). It is one of the specificities of Alpage to put together both researchers with a background in computer science (Inria members) and researchers with a background more oriented towards linguistics, all of them working on a single topic: simulation on computers of human understanding and production of language.

Natural language understanding systems convert samples of human language into more formal representations that are easier for computer programs to manipulate. Natural language generation systems convert information from computer databases into human language. Alpage focuses on *text* understanding and generation (by opposition to *speech* processing and generation).

One specificity of NLP is the diversity of human languages it has to deal with. Alpage focuses mostly on French. One of the main objectives of the team is to develop **generic** linguistically relevant *and* computationally efficient tools and resources for French which are freely distributed. These products are dedicated to the francophone community so as to help French to be part of the new information society. However, Alpage does not ignore other languages, through collaborations, in particular with those that are already studied by its members or by long-standing collaborators (e.g., English, Spanish, Polish, Persian and others). This is of course of high relevance, among others, for language-independant modelling and multi-lingual tools and applications.

Alpage covers all linguistics domains, although not at the same level. At the creation of the team, the morphological and syntactic levels was the most developed and led to a number of applications, especially with industrial partners. However, the importance of the semantic and discourse levels has increased during the evaluation period and the interface between syntax and semantics has been better worked on. Our goal is also to apply our knowledge, tools and resources in various contexts such as research in experimental linguistics, operational applications and prototypes as well as standardisation of linguistic resources and annotations.

Our four main objectives, as reworded and updated while writing the 2015 Inria evaluation report, are the following:

- **Objective i: Towards large scale natural language understanding at the sentence level** This objective covers all the work carried out on shallow processing, tagging, syntactic parsing, deep-syntactic parsing and shallow semantic parsing.
- **Objective ii : Language resource development, evaluation and use** This objective covers all language resource development efforts that range from morphology to semantics including syntax, but not including supra-sentential (discourse) resources.
- **Objective iii : Modelling and parsing supra-sentential phenomena** This objectives covers all efforts, including language resource development efforts, regarding discourse and other phenomena that cross sentence boundaries (e.g. anaphora).
- **Objective iv : Application domains** This objectives regroups the three main application domains for Alpage: empirical linguistics, academic downstream NLP applications and industrial applications.

3. Research Program

3.1. From programming languages to linguistic grammars

Participants: Éric Villemonte de La Clergerie, Benoît Sagot, Pierre Boullier, Djamé Seddah, Corentin Ribeyre.

Historically, several members of Alpage were originally specialists in the domain of modeling and parsing for programming languages, and have been working for more than 15 years on the generalization and extension of the techniques involved to the domain of natural language. The shift from programming language grammars to NLP grammars seriously increases complexity (e.g., grammar size [0]) and requires ways to handle the ambiguities inherent in every human language. It is well known that these ambiguities are the sources of many badly handled combinatorial explosions.

Furthermore, while most programming languages are expressed by (subclasses) of well-understood context-free grammars (CFGs), no consensual grammatical formalism has yet been accepted by the whole linguistic community for the description of human languages. On the contrary, new formalisms (or variants of older ones) appear constantly. Many of them may be classified into the three following large families:

Mildly Context-Sensitive (MCS) formalisms They manipulate possibly complex elementary structures with enough restrictions to ensure the possibility of parsing with polynomial time complexities. They include, for instance, Tree Adjoining Grammars (TAGs) and Multi-component TAGs with trees as elementary structures, Linear Indexed Grammars (LIGs). Although they are strictly more powerful than MCS formalisms, Range Concatenation Grammars (RCGs, introduced and used by Alpage members, such as Pierre Boullier and Benoît Sagot [56], [79], [84]) are also parsable in polynomial time.

Unification-based formalisms They combine a context-free backbone with logic arguments as decoration on non-terminals. Most famous representatives are Definite Clause Grammars (DCGs) where PROLOG powerful unification is used to compute and propagate these logic arguments. More recent formalisms, like Lexical Functional Grammars (LFGs) and Head-Driven Phrasal Structure Grammars (HPSGs) rely on more expressive Typed Feature Structures (TFS) or constraints.

Unification-based formalisms with an MCS backbone The two above-mentioned characteristics may be combined, for instance by adding logic arguments or constraints to non-terminals in TAGs.

[0] boullier:2010:inria-00516341:1

An efficient way to develop large-coverage hand-crafted symbolic grammars is to use adequate tools and adequate levels of representation, and in particular Meta-Grammars, one of Alpage's areas of expertise, especially with the FRMG grammar and parser for French based on the DyALog logic programming environment [92], [91]. Meta-Grammars (MGs) allows the linguist to focus on a modular description of the linguistic aspects of a grammar, rather than focusing on the specific aspects of a given grammatical formalism. Translation from MGs to grammatical formalisms such as TAG or LFG may be automatically handled. Graphical environments can be used to design MGs and their modularity provides a promising way for sharing the description of common linguistic phenomena across human languages.

3.2. Statistical Parsing

Participants: Djamé Seddah, Marie-Hélène Candito, Benoît Crabbé, Éric Villemonte de La Clergerie, Benoît Sagot, Corentin Ribeyre, Pierre Boullier, Maximin Coavoux.

Contrary to symbolic approaches to parsing, in statistical parsing, the grammar is extracted from a corpus of syntactic trees : a treebank. The main advantage of the statistical approach is to encode within the same framework the parsing and disambiguating tasks. The extracted grammar rules are associated with probabilities that allow to score and rank the output parse trees of an input sentence. This obvious advantage of probabilistic context-free grammars has long been counterbalanced by two main shortcomings that resulted in poor performance for plain PCFG parsers: (i) the generalization encoded in non terminal symbols that stand for syntagmatic phrases is too coarse (so probabilistic independence between rules is too strong an assertion) and (ii) lexical items are underused. In the last decade though, effective solutions to these shortcomings have been proposed. Symbol annotation, either manual [72] or automatic [75], [76] captures inter-dependence between CFG rules. Lexical information is integrated in frameworks such as head-driven models that allow lexical heads to percolate up the syntagmatic tree [59], or probabilistic models derived from lexicalized Tree Adjoining grammars, such as Stochastic Tree Insertion Grammars [58].

In the same period, totally different parsing architectures have been proposed, to obtain dependency-based syntactic representations. The properties of dependency structures, in which each word is related to exactly one other word, make it possible to define dependency parsing as a sequence of simple actions (such as read buffer and store word on top of a stack, attach read word as dependent of stack top word, attach read word as governor of stack top word ...) [94], [74]. Classifiers can be trained to choose the best action to perform given a partial parsing configuration. In another approach, dependency parsing is cast into the problem of finding the maximum spanning tree within the graph of all possible word-to-word dependencies, and online classification is used to weight the edges [73]. These two kinds of statistical dependency parsing allow to benefit from discriminative learning, and its ability to easily integrate various kinds of features, which is typically needed in a complex task such as parsing.

Statistical parsing is now effective, both for syntagmatic representations and dependency-based syntactic representations. Alpage has obtained state-of-the-art parsing results for French, by adapting various parser learners for French, and works on the current challenges in statistical parsing, namely (1) robustness and portability across domains and (2) the ability to incorporate exogenous data to improve parsing attachment decisions. Alpage is the first French team to have turned the French TreeBank into a resource usable for training statistical parsers, to distribute a dependency version of this treebank, and to make freely available various state-of-the art statistical POS-taggers and parsers for French. We review below the approaches that Alpage has tested and adapted, and the techniques that we plan to investigate to answer these challenges.

In order to investigate statistical parsers for French, we have first worked how to use the French Treebank [53], [52] and derive the best input for syntagmatic statistical parsing [60]. Benchmarking several PCFG-based learning frameworks [86] has led to state-of-the-art results for French, the best performance being obtained with the split-merge Berkeley parser (PCFG with latent annotations) [76].

In parallel to the work on dependency based representation, presented in the next paragraph, we also conducted a preliminary set of experiments on richer parsing models based on Stochastic Tree Insertion Grammars as used in [58] and which, besides their inferior performance compared to PCFG-LA based parser, raise promising results with respect to dependencies that can be extracted from derivation trees. One variation we explored, that uses a specific TIG grammar instance, a *vertical* grammar called *spinal* grammars, exhibits interesting properties wrt the grammar size typically extracted from treebanks (a few hundred unlexicalized trees, compared to 14 000 CFG rules). These models are currently being investigated in our team [89].

Pursuing our work on PCFG-LA based parsing, we investigated the automatic conversion of the treebank into dependency syntax representations [57], that are easier to use for various NLP applications such as question-answering or information extraction, and that are a better ground for further semantic analysis. This conversion can be applied on the treebank, before training a dependency-based parser, or on PCFG-LA parsed trees. This gives the possibility to evaluate and compare on the same gold data, both syntagmatic- and dependency-based statistical parsing. This also paved the way for studies on the influence of various types of lexical information.

3.3. Robust linguistic processing

Participants: Djamé Seddah, Benoît Sagot, Éric Villemonte de La Clergerie, Marie-Hélène Candito, Pierre Magistry.

The constitution of resources such as lexica or grammars raises the issues of the evaluation of these resources to assess their quality and coverage. For this reason, Alpage was the leader of the PASSAGE ANR project (ended in June 2010), which is the follow-up of the EASy parsing evaluation campaign held in 2004 and conducted by team LIR at LIMSI.

However, although developing parsing techniques, grammars (symbolic or probabilistic), and lexica constitute the key efforts towards deep large-scale linguistic processing, these components need to be included inside a full and robust processing chain, able to handle any text from any source, especially out-of-domain text genres. Such texts that exhibit properties (e.g., lexical and syntactic properties) that are different or differently distributed than what is found on standard data (e.g., training corpora for statistical parsers). The development of shallow processing chains, such as SxPipe , is not a trivial task [80]. Obviously, they are often used as such, and not only as pre-processing tools before parsing, since they perform the basic tasks that produce immediately usable results for many applications, such as tokenization, sentence segmentation, spelling correction (e.g., for improving the output of OCR systems), named entity detection, disambiguation and resolution, as well as morphosyntactic tagging.

Still, when used as a preliminary step before parsers, the quality of parsers' results strongly depends on the quality of such chains. This is especially the case, beyond the standard out-of-domain corpora mentioned above, for user-generated content. Indeed, until very recently out-of-domain text genres that have been prioritized have not been Web 2.0 sources, but rather biomedical texts, child language and general fiction (Brown corpus). Adaptation to user-generated content is a particularly difficult instance of the domain adaptation problem since Web 2.0 is not really a domain: it consists of utterances that are often ungrammatical. It even shares some similarities with spoken language [90]. The poor overall quality of texts found on such media lead to weak parsing and even POS-tagging results. This is because user-generated content exhibits both the same issues as other out-of-domain data, but also tremendous issues related to tokenization, typographic and spelling issues that go far beyond what statistical tools can learn from standard corpora. Even lexical specificities are often more challenging than on edited out-of-domain text, as neologisms built using productive morphological derivation, for example, are less frequent, contrarily to slang, abbreviations or technical jargon that are harder to analyse and interpret automatically.

In order to fully prepare a shift toward more robustness, we developed a first version of a richly annotated corpus of user-generated French text, the French Social Media Bank [7], which includes not only POS, constituency and functional information, but also a layer of "normalized" text. This corpus is fully available and constitutes the first data set on Facebook data to date and the first instance of user generated content for a morphologically-rich language. Thanks to the support of the Labex EFL through, we are currently the finalizing the second release of this data set, extending toward a full treebank of over 4,000 sentences.

Besides delivering a new data set, our main purpose here is to be able to compare two different approaches to user-generated content processing: either training statistical models on the original annotated text, and use them on raw new text; or developing normalization tools that help improving the consistency of the annotations, train statistical models on the normalized annotated text, and use them on normalized texts (before un-normalizing them).

However, this raises issues concerning the normalization step. A good sandbox for working on this challenging task is that of POS-tagging. For this purpose, we did leverage Alpage's work on MElt, a state-of-the art POS tagging system [68]. A first round of experiments on English have already led to promising results during the shared task on parsing user-generated content organized by Google in May 2012 [77], as Alpage was ranked second and third [88]. For achieving this result, we brought together a preliminary implementation of a normalization wrapper around the MElt POS tagger followed by a state-of-the art statistical parser improved by several domain adaptation techniques we originally developed for parsing edited out-of-domain texts. Those techniques are based on the unsupervized learning of word clusters *a la* Brown and benefit from morphological treatments (such as lemmatization or desinflection) [87].

One of our objectives is to generalize the use of the normalization wrapper approach to both POS tagging and parsing, for English and French, in order to improve the quality of the output parses. However, this raises several challenges: non-standard contractions and compounds lead to unexpected syntactic structures. A first round of experiments on the French Social Media Bank showed that parsing performance on such data are much lower than expected. This is why, we are actively working to improve on the baselines we established on that matter.

3.4. Dynamic wide coverage lexical resources

Participants: Benoît Sagot, Laurence Danlos, Éric Villemonte de La Clergerie, Marie-Hélène Candito, Lucie Barque, Marianne Djemaa.

Grammatical formalisms and associated parsing generators are useful only when used together with linguistic resources (lexicons, grammars) so as to build operational parsers, especially when considering modern lexically oriented grammatical formalisms. Hence, linguistic resources are the topic of the following section.

However, wide coverage linguistic resources are scarce and expensive, because they are difficult to build, especially when hand-crafted. This observation motivates us to investigate methods, along to manual development techniques, to automatically or semi-automatically acquire, supplement and correct linguistic resources.

Linguistic expertise remains a very important asset to benefit efficiently from such techniques, including those described below. Moreover, linguistically oriented environments with adequate collaborative interfaces are needed to facilitate the edition, comparison, validation and maintenance of large scale linguistic resources. Just to give some idea of the complexity, a syntactic lexicon, as described below, should provide rich information for several tens of thousands of lemma and several hundreds of thousands of forms.

Successful experiments have been conduced by Alpage members with different languages for the automatic acquisition of morphological knowledge from raw corpora [83]. At the syntactic level, work has been achieved on automatic acquisition of atomic syntactic information and automatic detection of errors in the lexicon [95],[6]. At the semantic level, automatic wordnet development tools have been described [78], [93], [71], [69]. All such techniques need of course to be followed by manual validation, so as to ensure high-quality results.

For French, these techniques, and others, have lead some Alpage members to develop one of the main syntactic resources for French, the Le*fff* [81], [85], developed within the Alexina framework. At the semantic level, Alpage members have developed or are developing various syntactico-semantic or semantic resources, including:

- a wordnet for French, the WOLF [82], [70], the first freely available resource of the kind;
- a French FrameNet lexicon (together with an annotated corpus) within the ASFALDA ANR project;
- and a French VerbNet.

In the last few years, Alpage members have shown how to benefit from other more linguistically-oriented resources, such as the Lexique-Grammaire and DICOVALENCE , in order to improve the coverage and quality of the Le*fff* , the WOLF, the French FrameNet lexicon and the French VerbNet. This work is a good example of how Inria and Paris 7 members of Alpage fruitful collaborate: this collaboration between NLP computer scientists and NLP linguists have resulted in significant advances which would have not been possible otherwise.

Moreover, an increasing effort has been made towards multilingual aspects. In particular, Alexina lexicons exist for German, Slovak, Polish, English, Spanish, Persian, Latin (verbs only), Kurmanji Kurdish, Maltese (verbs only, restricted to the so-called first *binyan*) and Khaling, not including freely-available lexicons adapted to the Alexina framework.

3.5. Discourse structures

Participants: Laurence Danlos, Timothée Bernard, Raphaël Salmon.

Until now, the linguistic modeling and automatic processing of sentences has been the main focus of the community. However, many applications would benefit from more large-scale approaches which go beyond the level of sentences. This is not only the case for automatic translation: information extraction/retrieval, summarizing, and other applications do need to resolve anaphora, which in turn can benefit from the availability of hierarchical discourse structures induced by discourse relations (in particular through the notion of right frontier of discourse structures). Moreover, discourse structures are required to extract sequential (chronological, logical,...) or hierarchical representations of events. It is also useful for topic extraction, which in turns can help syntactic and semantic disambiguation.

Although supra-sentential problematics received increasing attention in the last years, there is no satisfying solution to these problems. Among them, anaphora resolution and discourse structures have a far-reaching impact and are domains of expertise of Alpage members. But their formal modeling has now reached a maturity which allows to integrate them, in a near future, inside future Alpage tools, including parsing systems inherited from Atoll.

It is well known that a text is not a random sequence of sentences: sentences are linked the ones to the others by “discourse relations”, which give to the text a hierarchical structure. Traditionally, it is considered that discourse relations are lexicalized by connectors (adverbial connectors like *ensuite*, conjunctions like *parce que*), or are not lexicalized. This vision is however too simple:

- first, some connectors (in particular conjunctions of subordination) introduce pure modifiers and must not be considered as bearing discourse relations,
- second, other elements than connectors can lexicalize discourse relations, in particular verbs like *précéder / to precede* or *causer / to cause*, which have facts or fact eventualities as arguments [62].

There are three main frameworks used to model discourse structures: RST, SDRT, and, more recently, the TAG-based formalism D-LTAG. Inside Alpage, Laurence Danlos has introduced D-STAG (Discourse Synchronous TAGs, [63],[4]), which subsumes in an elegant way both SDRT and RST, to the extent that SDRT and RST structures can be obtained by two different partial projections of D-STAG structures. As done in D-LTAG, D-STAG extends a lexicalized TAG analysis so as to deal with the level of discourse. D-STAG has been fully formalized, and is hence possible to implement (thanks to Synchronous TAG, or even TAG parsers), provided one develops linguistic descriptions in this formalism.

4. Highlights of the Year

4.1. Highlights of the Year

In 2016, Alpage has obtained several new national fundings: the team is the leader of a new ANR project (Parsiti), and a partner of a new ANR project (Profiterole) and of a new ANR-NSF project (MCM-NL).

5. New Software and Platforms

5.1. Alexina

Atelier pour les LEXiques INformatiques et leur Acquisition
FUNCTIONAL DESCRIPTION

Alexina is Alpage'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.

- Participants: Benoît Sagot and Laurence Danlos
- Contact: Benoît Sagot
- URL: http://gforge.inria.fr/projects/alexina/

5.2. Bonsai

FUNCTIONAL DESCRIPTION

Alpage has developed a statistical parser for French, named Bonsai, trained on the French Treebank. This parser provides both a phrase structure and a projective dependency structure as output. This parser operates sequentially: (1) it first outputs a phrase structure analysis of sentences reusing the Berkeley implementation of a PCFG-LA trained on French by Alpage (2) it applies on the resulting phrase structure trees a process of conversion to dependency parses using a combination of heuristics and classifiers trained on the French treebank. The parser currently outputs several well known formats such as Penn treebank phrase structure trees, Xerox like triples and CONLL-like format for dependencies. The parsers also comes with basic preprocessing facilities allowing to perform elementary sentence segmentation and word tokenisation, allowing in theory to process unrestricted text. However it is believed to perform better on newspaper-like text.

- Participants: Marie-Hélène Candito, Djame Seddah and Benoît Crabbe
- Contact: Marie-Hélène Candito
- URL: http://alpage.inria.fr/statgram/frdep/fr_stat_dep_parsing.html

5.3. Crapbank

French Social Media Bank
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: Djame Seddah

5.4. DyALog

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 Villemonte De La Clergerie
- Contact: Eric Villemonte De La Clergerie
- URL: http://dyalog.gforge.inria.fr/

5.5. FDTB1

- Contact: Laurence Danlos

5.6. FQB

French QuestionBank
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: Djame Seddah

5.7. FRMG

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

5.8. Extreme UGC corpus

FUNCTIONAL DESCRIPTION

The Extreme UGC corpus is French three-domain data set focusing on user-generated content, made up of noisy question headlines from a cooking forum, live game chat logs and associated forums from two popular online games (MINECRAFT and LEAGUE OF LEGENDS). Building such an out of domain corpus, allowed us to consider the limits of our current normalization approaches. Currently annotated with part-of-speech, we plan to add other annotations layers.

- Contact: Djame Seddah

5.9. LexConn

- Contact: Laurence Danlos

5.10. LexViz

FUNCTIONAL DESCRIPTION

In the context of the industrial collaboration of ALPAGE with the company Lingua & Machina, we have extended their WEB plateform Libellex with a new component used to visualize and collaboratively validate lexical resources. In particular, this extension is used to manage terminological lists and lexical networks. The implemented graph-based representation has proved to be intuitive and quite useful for navigating in such large lexical resources (on the order to 10K to 100K entries).

- Participants: Eric Villemonte De La Clergerie and Mickaël Morardo
- Contact: Eric Villemonte De La Clergerie

5.11. 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://www.rocq.inria.fr/alpage-wiki/tiki-index.php?page=MElt

5.12. Mgwiki

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.

- Participants: Eric Villemonte De La Clergerie and Paul Bui-Quang
- Contact: Eric Villemonte De La Clergerie
- URL: http://alpage.inria.fr/frmgwiki/

5.13. OGRE

Optimized Graph Rewriting Engine
FUNCTIONAL DESCRIPTION

OGRE is a graph rewriting system specifically designed for manipulating linguistic trees and graphs, It relies on a rule specification language for expressing graph rewriting patterns. The transformation is performed in two steps:

First, the system performs simple transformations following the rewriting patterns,

Second, constraints can be applied on edges, which applies transformations depending on their environment that are propagated while all constraints are satisfied.

The system has been designed for the analysis and manipulation of attributed oriented and multi-relational graphs. Itis currently being used to convert existing universal dependencies for French to the upcoming 2.0 scheme to be used for the next “big” CoNLL parsing Shared Task of 2017.

- Participants: Corentin Ribeyre, Djame Seddah, Eric Villemonte De La Clergerie and Marie-Hélène Candito
- Contact: Corentin Ribeyre
- URL: http://www.corentinribeyre.fr/projects/view/OGRE

5.14. SYNTAX

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

5.15. Sequoia corpus

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: Djame Seddah

5.16. SxPipe

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 chain aimed to apply to raw corpora a cascade of surface processing steps. It is used as a preliminary step before Alpage's parsers (e.g., FRMG) and for surface processing (named entities recognition, text normalization, unknown word extraction and processing...).

- Participants: Pierre Boullier, Benoît Sagot, Eric Villemonte De La Clergerie and Djame Seddah
- Contact: Benoît Sagot
- URL: http://lingwb.gforge.inria.fr/

5.17. Verb∋net

- Contact: Laurence Danlos

5.18. dyalog-sr

KEYWORD: Parsing

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.

- Contact: Éric De La Clergerie

5.19. hyparse

Alpage Hybrid Parser
KEYWORDS: Parsing - NLP
FUNCTIONAL DESCRIPTION

Multilingual Phrase Structure Parser

- Contact: Benoît Crabbe
- URL: http://hyparse.gforge.inria.fr

5.20. vera

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

6. New Results

6.1. Deep syntactic parsing

Participants: Corentin Ribeyre, Marie-Hélène Candito.

Syntax plays an important role in the task of predicting the semantic structure of a sentence. But syntactic phenomena such as alternations, control and raising tend to obfuscate the relation between syntax and semantics. We have investigated how to predict the semantic structure of a sentence, encoded using the FrameNet model, taking advantage of deeper syntactic information than what is usually used. This deep syntactic representation abstracts away from purely syntactic phenomena and proposes a structural organization of the sentence that is closer to the semantic representation, by normalising the syntactic paths between a verb and its arguments. This reduces the variety of the syntactic realization of semantic roles, as shown by a decrease of the entropy of the syntactic paths of a given role.

Experiments conducted on a French corpus annotated with semantic frames showed that a FrameNet semantic parser reaches better performances with such a deep syntactic information [31]. For instance, switching from surface to deep syntactic information leads to a significant gain in FrameNet role identification, especially when this information is predicted (rather than reference information): $+5.1$ points (56.7 to 61.7) on all triggers [0] and $+6.7$ points (61.3 to 68.0) on verbal triggers only. These results clearly show the benefit of using deep syntactic features.

[0] In the sense of FrameNet, i.e. predicative lexial units, which should be assigned a frame.

6.2. Multilingual POS-tagging

Participant: Benoît Sagot.

Morphosyntactic lexicons and word vector representations have both proven useful for improving the accuracy of statistical part-of-speech taggers. We compare the performances of four systems on datasets covering 16 languages, two of these systems being feature-based (MEMMs —in the case of our own system MElt— and CRFs) and two of them being neural-based (bi-LSTMs). We show that, on average, all four approaches perform similarly and reach state-of-the-art results. Yet we obtained better performances with feature-based models on lexically richer datasets (e.g. for morphologically rich languages), whereas neural-based results are higher on datasets with less lexical variability (e.g. for English). These conclusions hold in particular for the MEMM models relying on our system MElt, which benefited from newly designed features [32], [44]. Thus we have shown that, under certain conditions, feature-based approaches enriched with morphosyntactic lexicons are competitive with respect to neural methods.

6.3. Transition-based constituency parsing with HyParse

Participants: Benoît Crabbé, Maximin Coavoux.

Transition-based parsing reduces the parsing task to predict a sequence of atomic decisions. These decisions are taken while sequentially reading words from a buffer and combining them incrementally into syntactic structures. The resulting structures are often dependency structures but can also be constituents, as is the case for our parser HyParse. Such an approach is therefore linear in the length of the input sentence, making transition-based parsing computationally efficient relative to other approaches. The challenge in transition-based parsing is modelling which action should be taken in each state it encounters as it progresses in a sentence provided as an input.

Training of a transition-based parser therefore consists in training a function that maps each of the unboundedly many states the parser might encounter to the best possible action, or transition, it should take. This function generally relies on a huge set of features, often conveniently grouped in the form of more abstract feature templates. Yet selecting the optimal subset of feature(template)s remains a challenge.

The training procedure therefore requires the help of an “oracle”, that is a function that returns the action that the parser should take in a given parser state given the gold parse. If the oracle assumes that the next action is necessarily the one given in the gold parse, it is said to be "static” and the oracle is deteminist. In order to train the parser to take relevant decisions when in an erroneous state, we can introduce some non determinism in the oracle in order to explore not only gold transition sequences but also near-gold transition sequences. This is the purpose of a dynamic oracle. Dynamic oracle training has shown substantial improvements for dependency parsing in various settings, but had not previously been explored for constituent parsing.

The two research directions we have investigated reflect the two above-mentioned challenges.

First, in collaboration with Rachel Bawden, now PhD student at LIMSI, we resumed our work on developing an efficient, language-independent model selection method for our parser HyParse [61]. It is designed for model selection when faced with a large number of possible feature templates, which is typically the case for morphologically rich languages, for which we want to exploit morphological information. The method we proposed uses multi-class boosting for iterative selection in constant time, using virtually no *a priori* constraints on the search space. We did however use a pre-ranking step before selection in order to guide the selection process. Our experiments have illustrated the feasibility of the method for our working language, French and resulted in high-performing, compact models much more efficiently than naive methods [22].

Second, we developed a dynamic oracle for HyParse. First, we replaced the traditional feature-based approach used in the above-described experiments by a neural approach. This is a way to overcome the feature selection issue addressed in the above-described work. The neural network weighting function we developed uses a non-linear hidden layer to automatically capture interactions between variables, and embeds morphological information in a vector space, as is usual for words and other symbols. Then, we developed our dynamic oracle based on this neural function and conducted experiments on the 9 languages of the SPMRL dataset in order to assess the impact of this oracle [25]. The experiments have shown that a neural greedy parser with morphological features, trained with a dynamic oracle, leads to accuracies comparable with the best currently available non-reranking and non-ensemble parsers.

6.4. French FrameNet

Participants: Marie-Hélène Candito, Marianne Djemaa.

In 2016 we have continued the development of a French FrameNet, within the ASFALDA project. While the first phase of the project focused on the development of a French set of frames and corresponding lexicon (Candito et al., 2014), we have focused this year on the subsequent corpus annotation phase, which targeted four notional domains (commercial transactions, cognitive stances, causality and verbal communication). Given full coverage is not reachable for a relatively "new" FrameNet project such as ours, focusing on specific notional domains allowed us to obtain full lexical coverage for the frames of these domains, while partially reflecting word sense ambiguities. Furthermore, as frames and roles were annotated on two main French Treebanks (the French Treebank and the Sequoia Treebank), we were able to extract a syntactico-semantic lexicon from the annotated frames. In the resource's current status [28], there are 98 frames, 662 frame-evoking words or "triggers", 872 senses, and about 13,000 annotated frames, with their semantic roles assigned to portions of text [0]

During this year's resource development efforts, we have put a specific emphasis on the causality domain (about 4000 instances of causal lexical items with their corresponding semantic frames are included in our resource). In the process of building the French lexicon and preparing the annotation of the corpus, we had to remodel some of the frames proposed in FrameNet based on English data, with hopefully more precise frame definitions to facilitate human annotation. This includes semantic clarifications of frames and frame elements, redundancy elimination, and added coverage. The result is arguably a significant improvement of the treatment of causality in FrameNet itself [34].

6.5. Verb∋net

Participants: Lucie Barque, Laurence Danlos.

VerbNet is a lexical resource for English verbs in which verbs are grouped together based on their ability to appear in similar sets of syntactic frames that correspond as well to alternations exhibited by verbs as to alternative syntactic realizations (Kipper et al. 2004). A French Verbnet, named Verb∋net, was first automatically derived from English VerbNet (Pradet et al., 2014) and is still under development. [13] details how Verb∋net was developed from the English VerbNet while using as far as possible the available lexical resources for French and how the various French alternations are coded, focusing on differences with English (e.g. existence of pronominal forms). One difficulty encountered in the development of Verb∋net springs from the fact that the list of (potentially numerous) frames has no internal organization in VerbNet. [26] proposes a type system for frames that shows whether two frames are variants of a given alternation. Frame typing facilitates coherence checking of the resource in a "virtuous circle".

6.6. French FrameNet

Participant: Benoît Crabbé.

Elaborating on our previous work on Medieval French in collaboration with Sasha Simonenko (McGill) and Sophie Prévost (LATTICE), we have conducted the first large-scale quantitative investigation of the syncretisation of verbal subject agreement in this language and test a classic analysis which relates non-syncretic agreement and null subjects as parts of the same grammar. We have shown that agreement syncretisation and the emergence of overt pronominal subjects proceeded at the same rate. Under the Constant Rate Hypothesis of Kroch (1989), which states that a grammatical change has the same rate in different contexts, these results are compatible with the traditional analysis [40], [39], [33]. However, we show that this analysis also

[0] The French FrameNet is freely available at http://asfalda.linguist.univ-paris-diderot.fr/frameIndex.xml.

generates a number of predictions which are not borne out by the quantitative data. We conclude that a more complex model of interaction of subject and inflection parameters is needed. Such a model may for instance be one where the type of an ending (non-syncretic vs. syncretic), presumably dependent on some unrelated phonological mechanism, presents a parsing difficulty for a null subject-licensing grammar and thus lowers its probability to be chosen by the speaker, which eventually drives it to extinction, similarly to the grammar competition model proposed in Yang (2010).

We have also investigated the effects of the text form (prose vs. verse) on diachronic grammatical changes in Medieval French using parsed treebanks and (1 million words with PTB-like annotations). Despite the common intuition that the prose is somehow more "advanced" than the verse contemporary to it with respect to grammatical changes, the magnitude of the difference has remained unknown in the absence of quantificational evaluations. At the same time, the prevalence of verse in the earliest periods of documented French (i.e. X–XII c.) results in a strong and unavoidable correlation between time and form, which potentially undermines the results of the studies attempting to formally model Medieval French evolution. We have compared two historical changes across text forms (namely the loss of pro-drop and that of OV_{finite} order), and shown that verse and prose behave differently, at least regarding the OV_{finite} order, thus contradicting Kroch's (1989) Constant Rate Hypothesis [38].

6.7. Modelling discourse-level information

Participants: Laurence Danlos, Timothée Bernard.

We have continued our work on the formalisation of discourse-level information. First, we have proposed in [24] a new model in STAG syntax and semantics for subordinate conjunctions (SubConjs) and attributing phrases —attitude/reporting verbs (AVs; *believe*, *say*) and attributing prepositional phrase (APPs; *according to*). This discourse-oriented model is based on the observation that SubConjs and AVs are not homogeneous categories. Indeed, previous work has shown that SubConjs can be divided into two classes according to their syntactic and semantic properties. Similarly, AVs have two different uses in discourse: evidential and intentional. While evidential AVs and APPs have strong semantic similarities, they do not appear in the same contexts when SubConjs are at play. Our proposition aims at representing these distinctions and capturing these various discourse-related interactions.

We have also investigated how sentential and discourse TAG-based grammars can be interfaced, in collaboration with Aleksandre Maskharashvili and Sylvain Pogodalla (LORIA). Tree-Adjoining Grammars (TAG) have been used both for syntactic parsing, with sentential grammars, and for discourse parsing, with discourse grammars (see for example our D-STAG model or the D-LTAG model). Yet the modelling of discourse connectives (coordinate conjunctions, subordinate conjunctions, adverbs...) in TAG-based formalisms for discourse differ from their modelling in sentential grammars. Because of this mismatch, an intermediate processing step is required between the sentential and the discourse processes, both in parsing and in generation [27]. We have developed a method to smoothly interface sentential and discourse TAG grammars, without using such an intermediate processing step. This method, based on Abstract Categorial Grammars (ACG), allows for building D-STAG discourse structures that are direct acyclic graphs (DAG) and not only trees.

6.8. Detecting omissions in journalistic texts

Participants: Héctor Martínez Alonso, Benoît Sagot.

In the journalistic genre that is characteristic of online news, editors make frequent use of citations as prominent information; yet these citations are not always in full. The reasons for leaving information out are often motivated by the political leaning of the news platform.

Existing approaches to the detection of political bias rely on bag-of-words models that examine the words present in the writings. In the context of the VerDI project (see below), we have initiated work aimed at going beyond such approaches, which focus on what is said, by instead focusing on what is *ommited*. Thus, this method requires a pair of statements; an original one, and a shortened version with some deleted words or spans. The task is then to determine whether the information left out in the second statement conveys *substantial* additional information. If so, we consider that a certain statement pair presents an omission. To tackle this question, we used a supervised classification framework, for which we require a dataset of sentence pairs, each pair manually annotated for omission.

We have developed a small reference corpus for evaluation purposes, using and comparing both crowd and expert annotation. This corpus has allowed us to examine which features help automaticallly identify cases of omission. In addition to straightforward measures of word overlap (the Dice coefficient), we also determined that there is a good deal of lexical information that determines whether there is an omission. This work is, to the best of our knowledge, the first empirical study on omission identification in statement pairs. We shall make all data and annotations freely available upon publication.

6.9. Models for interoperable lexical data

Participants: Mohamed Khemakhem, Laurent Romary.

Lexical data play an essential role in computational linguistic in two complementary ways:

- They serve as basic resources with which computational linguistic process can be parameterized. Such lexical resources are usually automatically or semi-automatically produced, are highly structured and may cover various levels of linguistic description from basic morpho-syntactic content to semantic representations;
- When created manually either for the purpose of describing a language (mono- or multilingual dictionary) or as a by product other language based activities (e.g. technical writing, translation), they may serve as a primary source of observation to analyse the way the lexicon of a language is organized, is used in domain oriented content, or how languages vary across time, space and usage.

The Alpage team has a specific expertise in the domain of lexical data, having been involved in the recent years in the creation of reference resources for the French language in particular, but also as driving force in the definition of international standards for the modelling and representation of both semasiological (word to sense) and onomasiological (concept to term) lexical information:

- ISO 16642 (TMF, Terminological Markup framework) and ISO 30042 (TBX, TermBase eXchange) as reference standards for the interchange of terminological data, for instance between translators' workbenches, but also for the modelling of dialectal information in linguistics;
- ISO 24613 (LMF, Lexical Markup Framework), a modular modelling framework for the representation of both machine and human semasiological resources;
- The Text Encoding Initiative (TEI), which since its inception has provided an XML based format for human readable dictionaries, widely used in most last scale dictionary projects worldwide.

One of the difficulties in lexical modelling is to identify the proper modelling framework for a given lexical resource but also to ensure maximal interoperability across heterogeneous lexical content. In the recent period, we have been working on the following aspects:

- Participation in the on going revision of ISO 30046, and planning of a possible integration of a TBX dialect in the TEI guidelines;
- Setting up the revision of ISO 24613 as a multi-part standard. Alpage is now involved in the provision of a reference TEI based serialisation of LMF and the part dedicated to etymological/diachronical information;
- Proposing an extension to the TEI guidelines for the representation of etymological information in dictionaries thus offering a formal basis for the study of diachronical phenomena across dictionaries [46];

- Organising a workshop in the context of the COST action eNEL that brought together the most relevant experts in the field in order to provide a set of constraints to apply the TEI guidelines in a more interoperable way across dictionary projects;
- Starting working on a machine learning based process to extract lexical content and structure automatically from digitized legacy dictionaries, This activity, base don the architecture of the Grobid library, is the basis of the PhD work by Mohamed Khemakhem.

6.10. Open data in the arts and humanities

Participants: Luca Foppiano, Marie Puren, Charles Riondet, Laurent Romary, Dorian Seillier.

The issue of open data has become increasingly important in various scholarly domains for it impacts on the visibility of the corresponding works, the capacity to provide evidence for reported facts and results, but also let other scholars build up new research on existing data sets. This is particularly acute in the humanities where primary sources play an essential role in providing the core material of scholarly results and for which the digital turn has offered a unique perspective of building up a wealth of structure information about human traces at large.

Based upon the experience gained in the definition of the open access policy at Inria [42], [50], [43], we have pursued various activities leading to a better understanding of the technical, editorial and political factors that may improve the wide dissemination of scholarly data sets in the humanities:

- Carry out a large scale questionnaire on data re-use within the partnership of the Iperion projects, which showed the lack of a coherent data management policy across cultural heritage laboratories in Europe from the points of view of documentation, archiving, licencing and re-use [49];
- Design a concept [16], [41] to improve the general fluidity of research results in the humanities based on data quality assessment, data journals and above all the setting of of a data re-use charter between scholars and cultural research institutions in the humanities. This action, carried out in the context of the Parthenos project has started with the organisation of two high level workshops in Berlin and Paris with representatives of major cultural research institutions;
- Coordinate as leader of WP 4 (Standards) in the Parthenos project a major overview of the needs and possible deployment of standards in the humanities based of an in depth survey of possible research scenario and associated practices in the domain of standards (Deliverable 4.1 published in October 2016). This has been accompanied by specific technical developments such as the proposition of an extension to the TEI guidelines for the representation of embedded stand-off annotations [45], [51];
- Develop specific modules for mining digital sources in the humanities, in particular in the domain of named entity recognition as an improvement of the NERD software initially developed in the European Cendari project.

7. Bilateral Contracts and Grants with Industry

7.1. Contracts with Industry

Alpage has developed several collaborations with industrial partners. Apart from grants described in the next section and informal discussions, specific collaboration agreements have been set up with the following companies:

- Verbatim Analysis (license agreement, transfer agreement, “CIFRE” PhD (contract ended in Dec 2014), see section 5.20),
- Yseop (“CIFRE” PhD of Raphael Salmon started in 2012 about automatic text generation)
- Agence France-Presse (on-going discussions aimed at a renewal of a long-lasting collaborations, involving several joint projects and a CIFRE PhD)

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. LabEx EFL (Empirical Foundations of Linguistics) (2011 – 2021)

Participants: Laurence Danlos, Benoît Sagot, Marie-Hélène Candito, Benoît Crabbé, Pierre Magistry, Djamé Seddah, Maximin Coavoux, Éric Villemonte de La Clergerie.

Linguistics and related disciplines addressing language have achieved much progress in the last two decades but improved interdisciplinary communication and interaction can significantly boost this positive trend. The LabEx (excellency cluster) EFL (Empirical Foundations of Linguistics), launched in 2011 and headed by Jacqueline Vaissière, opens new perspectives by adopting an integrative approach. It groups together some of the French leading research teams in theoretical and applied linguistics, in computational linguistics, and in psycholinguistics. Through collaborations with prestigious multidisciplinary institutions (CSLI, MIT, Max Planck Institute, SOAS...) the project aims at contributing to the creation of a Paris School of Linguistics, a novel and innovative interdisciplinary site where dialog among the language sciences can be fostered, with a special focus on empirical foundations and experimental methods and a valuable expertise on technology transfer and applications.

Alpage is a very active member of the LabEx EFL together with other linguistic teams we have been increasingly collaborating with: LLF (University Paris 7 & CNRS) for formal linguistics, LIPN (University Paris 13 & CNRS) for NLP, LPNCog (University Paris 5 & CNRS) LSCP (ENS, EHESS & CNRS) for psycholinguistics, MII (University Paris 4 & CNRS) for Iranian and Indian studies. Alpage resources and tools have already proven relevant for research at the junction of all these areas of linguistics, both before the start of the LabEx EFL and within several EFL "scientific operations". Moreover, the LabEx provides Alpage with opportunities for collaborating with new teams, e.g., on language resource development and empirical studies in collaboration with descriptive linguists.

The LabEx EFL's scientific activities are spread accross 7 autonomous scientific "strands". In 2016,Benoît Sagot, Marie Candito and Benoît Crabbé were respectively deputy-head of strand 6 on "Language Resources", strand 5 on "Computational semantic analysis" and strand 2 on "Experimental grammar from a cross-linguistic perspective". Several project members are in charge of research operations within these 3 strands.

8.1.2. ANR

8.1.2.1. ANR project Profiterole (2017 - 2020)

Participants: Benoît Crabbé, Éric Villemonte de La Clergerie, Benoît Sagot.

PROFITEROLE is a 4-year ANR research project led by Sophie Prévost (LATTICE) that involves computational linguists and specialists of Medieval French from LATTICE (Univ. Paris 3, CNRS, ENS), ALPAGE and ICAR (Univ. Lyon, ENS).

PROFITEROLE has three closely correlated main goals that fall within the fields of linguistics and Natural Language Processing (NLP): (1) formal and computational modeling phonological, morphological and syntactic aspects of the diachronic evolution of French; (2) targeting the development of a methodology to explore and annotate heterogeneous linguistic data while providing automatic analysers for various stages of the French language; (3) expanding linguistic resources for French, by building a large annotated corpus (1 million words) of Medieval French (9th-15th centuries) and morphological lexicons (plus NLP tools) covering several stages of French. Alpage members will essentially be involved on the computational and formal modeling aspects of the project and on the design of automated processing tools for lexicon and syntax.

8.1.2.2. ANR project PARSITI (2016 - 2020)

Participants: Marie-Hélène Candito, Djamé Seddah [principal investigator], Benoît Crabbé, Éric Villemonte de La Clergerie, Benoît Sagot.

Exploiting multilingual user-generated content (UGC), for applications such as information extraction, text mining or summarization, and facilitate their access to a wider audience implies a qualitative step-ahead in Natural Language Understanding. This is because UGC differs from better-studied edited data in many ways, including by non-canonical syntax, highly contextualised nature and rich lexical variability. The ParSiTi ANR project focuses on three critical aspects: (1) Robust Parsing Technologies, (2) Accurate Machine Translation Engines and (3) Context-aware Methods, all backed by State-of-the-Art Morphological Analysers and Normalization tools. To showcase the different models and algorithms designed during the project, a Machine Translation System will be developed that will be able to translate UGC between French, Arabic and English.

8.1.2.3. ANR project PARSEME-FR (2016 - 2019)

Participants: Marie-Hélène Candito, Mathieu Constant [principal investigator], Benoît Crabbé, Laurence Danlos, Éric Villemonte de La Clergerie, Djamé Seddah.

PARSEME-FR is a 4-year ANR research project headed by Mathieu Constant (LIGM, Université Paris-Est Marne-la-Vallée, currently in “délégation” at Alpage). PARSEME-FR partners are LIGM, Alpage, LI (Université de Tours), LIF (Aix-Marseille Université) and LIFO (Université d’Orléans). This project aims at improving linguistic representativeness, precision and computational efficiency of Natural Language Processing (NLP) applications, notably parsing. The project focuses on the major bottleneck of these applications: Multi-Word Expressions (MWEs), i.e. groups of words with a certain degree of idiomaticity such as “hot dog”, “to kick the bucket”, “San Francisco 49ers” or "to take a haircut". In particular, it aims at investigating the syntactic and semantic representation of MWEs in language resources, the integration of MWE analysis in (deep) syntactic parsing and its links to semantic processing. Expected deliverables include enhanced language resources (lexicons, grammars and annotated corpora) for French, MWE-aware (deep) parsers and tools linking predicted MWEs to knowledge bases. This proposal is a spin-off of the European IC1207 COST action PARSEME on the same topic.

Alpage is participating mainly to two tasks: (i) the production of an evaluation corpus annotated with MWE and (ii) the production of MWE-aware statistical parsers, both for surface syntax and deep syntax. MWE recognition can be viewed as part of a more ambitious task of recovering the semantic units of a sentence. Combining it to deep syntactic parsing will provide a further step towards semantic parsing.

8.1.2.4. ANR project SoSweet (2015 - 2019)

Participants: Djamé Seddah, Marie-Hélène Candito, Benoît Sagot, Éric Villemonte de La Clergerie, Benoît Crabbé.

Led by Jean-Phillipe Magué (ENS Lyon), the SoSweet project focuses on the synchronic variation and the diachronic evolution of the variety of French language used on Twitter. Its goal is to provide a state-of-the-art socio-linguistic description of half a billion tweets collected over 5 years.

Alpage, specialized in natural language processing, takes care of the linguistics enrichment part, which provides the other partners with normalized and structurally enriched forms of text. Alpage is also responsible of providing distributional analysis of our corpus, by the means of various forms of word clustering in order to define sociolinguistic variants in the tweets.

8.1.2.5. ANR project ASFALDA (2012 – 2016)

Participants: Marie-Hélène Candito [principal investigator], Marianne Djemaa, Benoît Sagot, Éric Villemonte de La Clergerie, Laurence Danlos.

Alpage is principal investigator team for the ANR project ASFALDA, lead by Marie Candito. The other partners are the Laboratoire d’Informatique Fondamentale de Marseille (LIF), the CEA-List, the MELODI team (IRIT, Toulouse), the Laboratoire de Linguistique Formelle (LLF, Paris Diderot) and the Ant’inno society.

The project aims to provide both a French corpus with semantic annotations and automatic tools for shallow semantic analysis, using machine learning techniques to train analyzers on this corpus. The target semantic annotations are structured following the FrameNet framework [54] and can be characterized roughly as an explicitation of "who does what when and where", that abstracts away from word order / syntactic variation, and to some of the lexical variation found in natural language.

The project relies on an existing standard for semantic annotation of predicates and roles (FrameNet), and on existing previous effort of linguistic annotation for French (the French Treebank). The original FrameNet project provides a structured set of prototypical situations, called frames, along with a semantic characterization of the participants of these situations (called *roles*). We propose to take advantage of this semantic database, which has proved largely portable across languages, to build a French FrameNet, meaning both a lexicon listing which French lexemes can express which frames, and an annotated corpus in which occurrences of frames and roles played by participants are made explicit. The addition of semantic annotations to the French Treebank, which already contains morphological and syntactic annotations, will boost its usefulness both for linguistic studies and for machine-learning-based Natural Language Processing applications for French, such as content semantic annotation, text mining or information extraction.

To cope with the intrinsic coverage difficulty of such a project, we adopt a hybrid strategy to obtain both exhaustive annotation for some specific selected concepts (commercial transaction, communication, causality, sentiment and emotion, time), and exhaustive annotation for some highly frequent verbs. Pre-annotation of roles will be tested, using linking information between deep grammatical functions and semantic roles.

The project is structured as follows:

- Task 1 concerns the delimitation of the focused FrameNet substructure, and its coherence verification, in order to make the resulting structure more easily usable for inference and for automatic enrichment (with compatibility with the original model);
- Task 2 concerns all the lexical aspects: which lexemes can express the selected frames, how they map to external resources, and how their semantic argument can be syntactically expressed, an information usable for automatic pre-annotation on the corpus;
- Task 3 is devoted to the manual annotation of corpus occurrences (we target 20000 annotated occurrences);
- In Task 4 we will design a semantic analyzer, able to automatically make explicit the semantic annotation (frames and roles) on new sentences, using machine learning on the annotated corpus;
- Task 5 consists in testing the integration of the semantic analysis in an industrial search engine, and to measure its usefulness in terms of user satisfaction.

The scientific key aspects of the project are:

- an emphasis on the diversity of ways to express the same frame, including expression (such as discourse connectors) that cross sentence boundaries;
- an emphasis on semi-supervised techniques for semantic analysis, to generalize over the available annotated data.

8.1.2.6. ANR project Polymnie (2012-2016)

Participants: Laurence Danlos, Éric Villemonte de La Clergerie, Timothée Bernard.

Polymnie is an ANR research project headed by Sylvain Podogolla (Sémagramme, Inria Lorraine) with Melodi (INRIT, CNRS), Signes (LABRI, CNRS) and Alpage as partners. This project relies on the grammatical framework of Abstract Categorial Grammars (ACG). A feature of this formalism is to provide the same mathematical perspective both on the surface forms and on the more abstract forms the latter correspond to. ACG allows for the encoding of a large variety of grammatical formalisms, in particular Tree Adjoining grammars (TAG).

The role of Alpage in this project is to develop sentential or discursive grammars written in TAG and to participate in their conversion in ACG. Results were first achieved in 2014 concerning text generation: GTAG formalism created by Laurence Danlos in the 90's has been rewritten in ACG [64], [65], [66]. As regards discursive analysis, D-STAG formalism created by Laurence Danlos in the 00's has also been rewritten in ACG in 2015 [67] (see also [27]).

8.1.3. Other national initiatives

8.1.3.1. "RAPID" project VerDI (2016 – 2019)

Participants: Benoît Sagot, Héctor Martínez Alonso.

The ANR "RAPID" project VerDI focuses on the automatic identification of information dissimulation on the Internet and on social networks. Such dissimulations can be produced by omiting crucial pieces of information within documents or during written online discussions, by hiding them within a massive information flow, or using other techniques. VerDI aims at extending an existing journalistic fact-checking tool developed by Trooclick, the company that leads the project.

8.1.3.2. FUI project COMBI (2014-2016)

Participant: Laurence Danlos.

COMBI is is an "FUI 16" project. It started in February 2014 for a two year duration. It groups 5 industrial partners (Temis, Isthma, Kwaga, Yseop and Qunb) and Alpage. Temis and Istma work on data mining from texts and big data. Kwaga works on the interpretation and inferences that can be drawn from the data retrieved in the analysis module. Alpage and Qunb work, under the supervision of Yseop, on the production of respectively texts and graphics describing the results of the interpretation module. Currently, COMBI aims at creating the full chain for a user case concerning the weekly activity of an on-line service.

Alpage works on text generation, with the adaptation of TextElaborator, a generation system developed in the 10's by WatchAssistance and based on G-TAG. Alpage also works on the opportunity to describe pieces of information by texts, graphics or both.

8.1.3.3. Institut de Linguistique Française and Consortium CORLI within the TGIR Huma-Num

Participants: Benoît Sagot, Stéphane Riou, Djamé Seddah.

Huma-Num is a TGIR (Very Large Research Infrastructure) dedicated to digital humanities. Among Huma-Num initiatives are a dozen of consortia, which bring together most members of various research communities. Among them is the CORLI consortium (following, among other, the *Corpus Écrits* consortium in which previously participating), which is dedicated, among other topics, to all aspects related to written corpora, from NLP to corpus development, corpus specification, standardization, and others. All types of written corpora are covered (French, other languages, contemprorary language, medieval language, specialized text, non-standard text, etc.). The consortium CORLI is managed by the Institut de Linguistique Française, a CNRS federation of which Alpage is a member since June 2013, under the supervision of Franck Neveu.

Alpage is involved in various projects within this consortium, and especially in the development of corpora for CMC texts (blogs, forum posts, SMSs, textchat...) and shallow corpus annotation, especially with MElt, and in the development of a preliminary version of the future Corpus de Référence du Français (French Reference Corpus).

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. H2020 PARTHENOS

Participants: Laurent Romary, Luca Foppiano, Mohamed Khemakhem, Marie Puren, Charles Riondet, Dorian Seillier.

This EU project Parthenos of the H2020 INFRADEV program aims et 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. Within this project started in May 2015, Alpage has the leadership over the work package dedicated to the promotion and development of standards in the humanities.

In 2015, Laurent Romary and Charles Riondet have identified digital humanities use cases where standards play a central role and specified an architecture for organising standards related information (specification, software, bibliography, reference material, experts) at the service of scholars in the humanities.

8.2.1.2. H2020 EHRI

Participants: Laurent Romary, Luca Foppiano, Charles Riondet.

The EHRI 2 (European Holocaust Research Infrastructure), also in the INFRADEV program of H2020, seeks to transform archival research on the Holocaust, by providing methods and tools to integrate and provide access to a wide variety of archival content. The project has started in June 2015 and will led us to work on both standards for the representation of archival content and develop data mining components for archival textual data.

In 2015, we have focused on the identification of available data sources resulting from the first phase of the project in the previous years and compile specifications for the description of authorities according to the EAC (Encoded Archival Context) standard.

8.2.1.3. H2020 Iperion

Participants: Laurent Romary, Luca Foppiano, Marie Puren.

The H2020 Iperion project aims at coordinating infrastructural activities in the cultural heritage domain. Our team has a small participation in relation to the definition of data management and representation issues. This will directly contribute to increase our experience in curating the kind of heterogeneous linguistic data that we gathered over the years.

In 2015, we have designed a questionnaire for all data producers in the project in order to gather feedback on their existing practices (data flows, licences, formats) concerning the creation, management and dissemination of cultural heritage data. On this basis, we have produced a first version of the data management plan for the project.

8.2.2. Collaborations in European Programs, Except FP7 & H2020

Program: IC1207 COST

Project acronym: PARSEME

Project title: PARSing and Multi-word Expressions

Duration: March 2013 - March 2017

Coordinator: Agata Savary

Other partners: interdisciplinary experts (linguists, computational linguists, computer scientists, psycholinguists, and industrials) from 30 countries

Abstract: The aim of this project is to improve linguistic representativeness, precision and computational efficiency of Natural Language Processing (NLP) applications, focusing on the major bottleneck of these applications: Multi-Word Expressions (MWEs), i.e., sequences of words with unpredictable properties such as "to count somebody in" or "to take a haircut". A breakthrough in their modelling and processing is targeted, as the result of a coordinated effort of multidisciplinary experts working on fourteen different languages.

Program: ISCH COST Action IS1312

Project acronym: TextLink

Project title: Structuring Discourse in Multilingual Europe

Duration: April 2014 - April 2018

Coordinator: Liesbeth Degand

Other partners: experts in computational linguistics and discourse from 24 countries

France MC members: Laurence Danlos and Philippe Muller (IRIT)

Abstract: This action will facilitate European multilingualism by (1) identifying and creating a portal into discourse-level resources within Europe - including annotation tools, search tools, and discourse-annotated corpora; (2) delineating the dimensions and properties of discourse annotation across corpora; (3) organising these properties into a sharable taxonomy; (4) encouraging the use of this taxonomy in subsequent discourse annotation and in cross-lingual search and studies of devices that relate and structure discourse; and (5) promoting use of the portal, its resources and sharable taxonomy.

Program: **ISCH COST Action IS1305**

Project acronym: ENeL

Project title: European Network of e-Lexicography

Duration: October 2013 - October 2017

Coordinator: Prof Martin EVERAERT (NL)

Other partners: interdisciplinary experts (linguists, computational linguists, computer scientists, lexicographers, and industrials) from 31 countries

Abstract: The proposed Action aims to establish a European network of lexicographers in order to deal with the following issues: give easier access to scholarly dictionaries, establish a systematic exchange of expertise on common standards and solutions, develop a common approach to e-lexicography that forms the basis for a new type of lexicography that fully embraces the pan-European nature of much of the vocabularies of the languages spoken in Europe

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Selection

9.1.1.1. Member of the Conference Program Committees or Reviewer

- Marie Candito was a reviewer or member of the program committee for the following events: NAACL 2016, CONLL 2016, RECITAL 2016, COLING 2016
- Benoît Crabbé was a reviewer or member of the program committee for the following events: NAACL 2016, ACL 2016, COLING 2016, EMNLP 2016, CONLL 2016.
- Maximin Coavoux was a sub-reviewer or member of the program committee for the following events: ACL 2016, EMNLP 2016.
- Héctor Martínez Alonso was a reviewer or member of the program committee for the following events: ACL 2016, COLING 2016, CONLL 2016, EMNLP 2016, EACL 2017, NAACL 2016 , MWE 2016 and WNUT 2016.
- Laurent Romary was a member of the program committee for the following events: CMLC-4, Digital Humanities and Iconography, The case of alpine mural painting, Datech 2016, CILA 2016, TOTh (Terminology & Ontology: Theories and applications), LDL-2016, LG-LP 2016, ISA-12
- Benoît Sagot was a member of the program committe for the following events: NAACL 2016, EACL 2017, EMNLP 2016, ACL 2016
- Djamé Seddah was a reviewer or member of the program committee for the following events: ACL 2016, COLING 2016, EMNLP 2016, CONLL 2016, TLT 2016, LAW 2016, TALN 2016 and others workshops.

- Éric Villemonte de La Clergerie was a reviewer or member of the program committee for the following events: COLING'16 (tracks Parsing, and "Resources, Software and Tools"), LREC'16, NAACL'16, DiscoNLP'16, Toth'16

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- Laurent Romary is co-editor of the overlay Journal of Data Mining and Digital Humanities and for the DH Commons journal
- Laurent Romary is member of the advisory board of the Journal of the Text Encoding Initiative (jTEI)

9.1.2.2. Reviewer - Reviewing Activities

- Héctor Martínez Alonso was a reviewer for the Artificial Intelligence journal
- Laurent Romary was a reviewer for the jTEI journal
- Benoît Sagot was a reviewer for the following journals: Journal of Language Modelling, Language Resources and Evaluation, Northern European Journal of Language
- Djamé Seddah was a reviewer for the following journals: Asian Languages and Information Processing, Language Ressources and Evaluation

9.1.3. Invited Talks

- Marie Candito gave a talk on 24/03/2016 at the "3ème journée TAL et IA ", Paris
- Marie Candito gave a talk on 27/10/2016 at the DIGICOSME Labex, LIMSI, Orsay
- Héctor Martínez Alonso gave an invited talk on 05/04/2016 at PARSEME/ENeL workshop on MWE e-lexicons, Skopje (Republic of Macedonia)
- Héctor Martínez Alonso gave an invited talk on 01/06/2016 at Darmstadt Technical University (Germany).
- Marie Puren gave an invited talk on 26/10/2016 for theDARIAH's Humanities At Scale Winter School at Charles University in Prague.
- Marie Puren gave an invited talk on 22/11/2016 during the colloqium "DHNord" at the "Maison Européenne des Sciences de l'Homme et de la Société" in Lille.
- Charles Riondet gave an invited talk on 26/10/2016 at the DARIAH's Humanities at Scale Winter School, Charles University, in Prag (Czech Republic)
- Laurent Romary gave a speech at Ready to Reach Out, Conference on Digitization of Cultural Heritage (NL EU presidency) "Data and Dissertations", Amsterdam, The Netherlands on 29/06/2016
- Laurent Romary gave a keynote at ETD 2016 "Data and Dissertations", Lille, France on 11/07/2016 [50]
- Laurent Romary gave a keynote at Language Technologies and & Digital Humanities 2016 on 30/09/2016 in Ljubljana, Slovenia [51]
- Benoît Sagot was an invited panelist during the AnaMorphoSys conference on 20-22/06/2016 in Lyon
- Benoît Sagot gave an invited talk at the EPHE workshop on Digital Humanities on 12/10/2016 in Paris
- Djamé Seddah gave a keynote talk at the workshop "Data Driven Approach to Networks and Language" (Lyon)
- Djamé Seddah gave a keynote talk at the workshop "Challenges for Data-driven Natural Language Analysis beyond Standard Data (Lyon)
- Djamé Seddah gave a invited talk in Dusseldorf (Formal Linguistic Departement) on deep syntax based graph-parsing

- Djamé Seddah gave a invited talk in Lyon (ENS seminar series) on morpho-syntactic analysis in challenging environments
- Djamé Seddah gave a invited talk at the 1st NLP Paris Meetup on graph parsing for French
- Éric Villemonte de La Clergerie gave an invited talk at the meeting on “Information, Me´dias et Informatique” (IRISA Rennes, 15/03/2016)
- Éric Villemonte de La Clergerie gave an invited tutorial at the COST Parseme meeting (Dubrovnik, 27/09/2016)
- Éric Villemonte de La Clergerie gave an invited talk at the AIM-WEST & PARSEME-FR Workshop (IMAG Grenoble, 4/10/2016)
- Éric Villemonte de La Clergerie gave an invited talk at the NLP meetup (Paris, 23/11/2016)

9.1.4. Leadership within the Scientific Community

9.1.4.1. Involvement in international initiatives

- Alpage is involved in the ISO subcommittee TC 37/SC 4 on “Language Resource Management”. Éric Villemonte de La Clergerie has participated in various ISO meetings as an expert, in particular on morpho-syntactic annotations (MAF), feature structures (FSR & new FSD), and syntactic annotations (SynAF) [55]. Within the same subcommittee, Laurent Romary is the convenor of the working group on lexical resources (WG4).
- Laurent Romary is chairman of ISO committee TC 37 “Terminology and other language and content resources”
- Laurent Romary chairs the Board of Directors of the European Research Infrastructure Consortium DARIAH established by the European Commission to coordinate Digital Humanities infrastructure activities in Europe.
- Laurent Romary is member of the TEI Archiving, Publishing, and Access Service (TAPAS) project advisory board
- Laurent Romary is member of the International Advisory board of the Belgrade Center for Digital Humanities

9.1.4.2. Involvement in national initiatives

- Alpage has many responsabilities wihtin the LabEx EFL. Until February 2015, Benoît Sagot is deputy head of this research strand; Marie Candito is deputy head of the research strand on computational semantics; Benoît Crabbé is deputy head of the research strand on experimental grammar; all three are therefore deputy members of the Scientific and Governing Boards of the LabEx; Laurence Danlos is a member of the Scientific Board of the LabEx EFL, representing Alpage;
- Laurent Romary is the leader of the scientific committe of the EquipEx Ortolang, of which Benoît Sagot is also a member.
- Laurent Romary is chairman of the scientific council of ABES (Agence Bibliographique de l’Enseignement Supérieur)
- Laurent Romary is also member of several scientific committee or advisory board: Labex ‘Les passes dans le présent’ (PasP), OpenAIRE 2020, OpenEdition (UMS Cleo)
- Laurent Romary is the Inria scientific advisor for Scientific and Technical Information, in charge in particular of the Open Access strategy.

9.1.4.3. Other activities for the scientific community

- Laurence Danlos is member of the Permanent Committee of the TALN conference (CPerm) organised by ATALA.

9.1.5. Scientific Expertise

- Marie Candito was member of the committee for the "prix de thèse ATALA"
- Benoît Crabbé was reviewer for a ERC starting grant
- Laurent Romary was project reviewer for Fonds National Suisse, Switzerland
- Laurent Romary was project reviewer for the Canada Foundation for Innovation, Canada
- Laurent Romary was project reviewer for LabEx PATRIMA, Switzerland
- Laurent Romary has advised the European Patent Office for the specification of their model for representing the non patent literature, based on the TEI (Text Encoding Initiative) guidelines.
- Laurent Romary has been a referee for the selection of a professor at the University of Cologne, Germany
- Laurent Romary has part of the selection committee for a professorship at the Université de Lorraine, France
- Benoît Sagot was project reviewer for the ANR
- Benoît Sagot was an elected board member of the French NLP society (ATALA) until July 2016
- Benoît Sagot was a member of the HCERES evaluation committee for the UMR ATILF (december 2016)
- Djamé Seddah was a reviewer for the Luxemrbourg research funding agency.
- Djamé Seddah was an external reviewer for a Machine Learning PhD thesis (University of Barcelona, sup: Xavier Carreras)
- Éric Villemonte de La Clergerie has participated to several AFNOR meetings in relation with ISO TC37SC4 “Language Resource Management”. He was also a member of the French delegation at the annual ISO TC37 meeting (Copenhagen, June 2016)
- Éric Villemonte de La Clergerie was one of the animators of the working group of the GFII on Knowledge Technologies, and was also involved in the discussions of the GFII about the "regulation of algorithms"

9.1.6. Research Administration

9.1.6.1. University duties

- Lucie Barque was until june 2016 deputy director of the Linguistic department at Université Paris Nord
- Marie Candito is deputy director of the UFR of Linguistics.
- Maximin Coavoux is a student member of the Administrative board of the UFR of Linguistics of University Paris Diderot.
- Benoît Crabbé is a member of the Administrative board of the UFR of Linguistics of University Paris Diderot.
- Laurence Danlos is the chair of the Scientific Committee of the Linguistics UFR of University Paris Diderot.
- Laurence Danlos is the deputy chair of the Doctoral School for Linguistic Sciences (École Doctorale de Sciences du Langage).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence: Lucie Barque, Phonétique, 22,5 heures en équivalent TD, niveau L2, Université Paris 13, France

Licence: Lucie Barque, Dictionnaires électroniques, 22,5 heures en équivalent TD, niveau L2, Université Paris 13, France

Licence: Lucie Barque, Corpus électroniques, 22,5 heures en équivalent TD, niveau L2, Université Paris 13, France

Licence: Lucie Barque, Syntaxe et sémantique, 22,5 heures en équivalent TD, niveau L3, Université Paris 13, France

Licence: Lucie Barque, Pragmatique et Énonciation, 11 heures en équivalent TD, niveau L3, Université Paris 13, France

Licence: Timothée Bernard, TD d'Algorithmique, 24 heures en équivalent TD, niveau L3, Université Paris 7 Diderot, France

Licence: Marie Candito, Linguistique de corpus, 28 heures en équivalent TD, niveau L3, Université Paris Diderot, France

Licence: Marie Candito, Probabilités et statistiques pour le TAL, 28 heures en équivalent TD, niveau L3, Université Paris Diderot, France

Licence: Maximin Coavoux, Programmation 2, 28 heures en équivalent TD, niveau L3, Université Paris Diderot, France

Licence: Benoît Crabbé, Introduction à la programmation, 24 heures en équivalent TD, niveau L3, Université Paris Diderot, France.

Licence: Laurence Danlos, Introduction au TAL, 32 heures en équivalent TD, niveau L3, Université Paris-Diderot, France

Master: Lucie Barque, Ressources lexicales pour le TAL, 24 heures en équivalent TD, niveau M2, Université Paris 13, France

Master: Lucie Barque, La langue et son enseignement, 18 heures en équivalent TD, niveau M1, Université Paris 13, France

Master: Lucie Barque, Problématiques de la néologie, 36 heures en équivalent TD, niveau M2, Université Paris 13, France

Master: Timothée Bernard, Phonétique (TD), 12 heures en équivalent TD, niveau M1, Université Paris Diderot, France

Master: Timothée Bernard, Langages formels (TD), 24 heures en équivalent TD, niveau M1, Université Paris Diderot, France

Master: Marie Candito, Analyse sémantique automatique du langage naturel, 14 heures en équivalent TD, niveau M2, Université Paris Diderot, France

Master: Marie Candito, Traduction automatique, 51 heures en équivalent TD, niveau M1, Université Paris Diderot, France

Master: Marie Candito, Apprentissage automatique pour le TAL, 60 heures en équivalent TD, niveau M1, Université Paris Diderot, France

Master: Maximin Coavoux, Approches probabilistes du TAL (TD), 24 heures en équivalent TD, niveau M1, Université Paris Diderot, France

Master: Benoît Crabbé, Linguistique empirique et expérimentale, 24 heures en équivalent TD, niveau M2, Université Paris Diderot, France.

Master: Laurence Danlos, Discours: Analyse et génération de textes, 32 heures en équivalent TD, niveau M2, Université Paris-Diderot, France

Master: Laurent Romary, Basic encoding and annotation of textual sources in TEI, 24 hours, Master of Library and Information Science, Fach Hochschule Potsdam, Germany

Continuous training: Laurent Romary, Codage de document scientifique en XML TEI, 16 hours, INIST-CNRS, France

9.2.2. Supervision

PhD in progress: Timothée Bernard, “Analyse discursive et factualité”, started in September 2015, supervised by Laurence Danlos (superviser) and Philippe de Groote (co-superviser)

PhD in progress: Raphael Salmon, “Implémentation d’un système de génération à base de contraintes”, Université Paris-Diderot, started in October 2013, supervised by Laurence Danlos (superviser) and Alain Kaeser (tutor in the company Yseop)

PhD in progress: Marianne Djemaa, "Création semi-automatique d’un FrameNet du français", started in October 2012, supervised by Marie Candito

PhD: Corentin Ribeyre, “Vers la syntaxe profonde pour l’interface syntaxe-sémantique ”, started in November 2012, supervised by Laurence Danlos (superviser), Djamé Seddah (co-superviser) and Éric Villemonte de La Clergerie (co-superviser), defended on January 27, 2016

PhD in progress: Maximin Coavoux, “Représentations continues pour l’analyse syntaxique et sémantique automatique”, started in September 2015, supervised by Benoît Crabbé.

PhD in progress: Loïc Grobol, “Reconnaissance automatique de chaînes de coréférences en français par combinaison d’apprentissage automatique et de connaissances linguistiques”, started in October 2016, supervised by Isabelle Tellier (supervisor), Marco Dinarelli (co-supervisor), and Éric Villemonte de La Clergerie (co-supervisor)

PhD in progress: Axel Herold, “Automatic identification and modeling of etymons in retro-digitized dictionaries”, started in November 2016, supervised by Laurent Romary

PhD in progress: Jack T. 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 November 2016, supervised by Laurent Romary

PhD in progress: Mohamed Khemakhem, “Structuration automatique de dictionnaires à partir de modèles lexicaux standardisés”, started in September 2016, supervised by Laurent Romary

9.2.3. Juries

Laurence Danlos was an examinator in the PhD defense committee of Aleksandre Maskharashvili. Title: Discourse Modeling with Abstract Categorial Grammars. University: Université de Lorraine. PhD supervisor: P. de Groote, Sylvain Pogodalla. Defense date: 1st December 2016.

Laurent Romary was a reviewer (*rapporteur*) in the PhD defense committee of Sirine Boukedi Troudi. Title: Outils d’analyse des structures non-têtes. University: Université de Sfax. PhD supervisor: Kais Haddar. Defense date: 11 May 2016.

Laurent Romary was a reviewer (*rapporteur*) in the PhD defense committee of Daouda Sawadogo. Title: Architectures logicielles et mécanismes pour la gestion adaptative et consolidée de ressources numériques dans une application interactive scénarisée. University: Université de La Rochelle. PhD supervisor: Pascal Estraillier and Ronan Champagnat. Defense date: 28 June 2016.

Djamé Seddah and Éric Villemonte de La Clergerie was a member in the PhD defense committee of Corentin Ribeyre as co-supervisor, together with Laurence Danlos (as PhD director). Title: Méthodes d’Analyse Supervisée pour l’Interface Syntaxe-Sémantique. University: Université Paris-Diderot.Defense date: 27 January 2016.

Éric Villemonte de La Clergerie was a member in the PhD defense committee of Julie Belião. Title: How prosody and syntax are mapping: A study of synchronization and congruence. University: Université Paris Ouest. Defense date: 5 December 2016

9.3. Popularization

- Interview of Laurent Romary on new scientific publication models for APS news, American Physical Society (http://www.aps.org/publications/apsnews/201602/arxiv.cfm).
- Laurent Romary and Marie Puren have given an interview on 24/11/2016 on the European project Parthenos for the Inria website (https://www.inria.fr/centre/paris/actualites/epi-alpage-historiens-et-informaticiens-reunis-dans-le-projet-europeen-parthenos).

- Timothée Bernard has welcomed two high school students at Alpage during 4 days, in the framework of the “Science Académie” of the “Paris-Montagne” association.

10. Bibliography

Major publications by the team in recent years

[1] A. BITTAR, P. AMSILI, P. DENIS, L. DANLOS.*French TimeBank: an ISO-TimeML Annotated Reference Corpus*, in "ACL 2011 - 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies", Portland, OR, United States, Association for Computational Linguistics, June 2011, http://hal.inria.fr/inria-00606631/en.

[2] M. CANDITO, M. CONSTANT.*Strategies for Contiguous Multiword Expression Analysis and Dependency Parsing*, in "ACL 14 - The 52nd Annual Meeting of the Association for Computational Linguistics", Baltimore, United States, ACL, June 2014, https://hal.inria.fr/hal-01022415.

[3] B. CRABBÉ.*An LR-inspired generalized lexicalized phrase structure parser*, in "COLING", Dublin, Ireland, 2014, https://hal.inria.fr/hal-01105142.

[4] L. DANLOS.*D-STAG : un formalisme d’analyse automatique de discours fondé sur les TAG synchrones*, in "Traitement Automatique des Langues", 2009, vol. 50, n^{o} 1.

[5] B. SAGOT.*Construction de ressources lexicales pour le traitement automatique des langues*, in "Ressources Lexicales – Contenu, construction, utilisation, évaluation", N. GALA, M. ZOCK (editors), Lingvisticæ Investigationes Supplementa, John Benjamins, 2013, vol. 30, p. 217-254, https://hal.inria.fr/hal-00927281.

[6] B. SAGOT, É. VILLEMONTE DE LA CLERGERIE.*Error Mining in Parsing Results*, in "Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics", Sydney, Australia, Association for Computational Linguistics, July 2006, p. 329–336.

[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] J. THUILIER, G. FOX, B. CRABBÉ.*Prédire la position de l’adjectif épithète en français : approche quantitative*, in "Lingvisticae Investigationes", June 2012, vol. 35, n^{o} 1, https://hal.inria.fr/hal-00698896.

[9] 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.

[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] Z. AGIC, A. JOHANNSEN, B. PLANK, H. MARTÍNEZ ALONSO, N. SCHLUTER, A. SØGAARD.*Multilingual Projection for Parsing Truly Low-Resource Languageš*, in " Transactions of the Association for Computational Linguistics", August 2016, https://hal.inria.fr/hal-01426754.

[12] M. COAVOUX, B. CRABBÉ.*Prédiction structurée pour l'analyse syntaxique en constituants par transitions : modèles denses et modèles creux*, in "Traitement Automatique des Langues", 2016, vol. 57, n^{o} 1, https://hal.inria.fr/hal-01365252.

[13] L. DANLOS, Q. PRADET, L. BARQUE, T. NAKAMURA, M. CONSTANT.*Un Verbenet du français*, in "Traitement Automatique des Langues", September 2016, vol. 57, n^{o} 1, 25, https://hal.inria.fr/hal-01392817.

[14] B. GAUME, K. DUVIGNAU, E. NAVARRO, Y. DESALLE, H. CHEUNG, S. HSIEH, P. MAGISTRY, L. PREVOT.*Skillex: a graph-based lexical score for measuring the semantic efficiency of used verbs by human subjects describing actions*, in "Revue TAL", 2016, vol. 55, n^{o} 3, https://hal.archives-ouvertes.fr/hal-01320416.

[15] H. MARTINEZ ALONSO, D. ZEMAN.*Universal Dependencies for the AnCora treebanks* , in "Procesamiento del Lenguaje Natural", September 2016, n^{o} 57, https://hal.inria.fr/hal-01426751.

[16] L. ROMARY, M. MERTENS, A. BAILLOT.*Datenfluss an der Schnittstelle von Forschung, Infrastruktur und Kulturerbeinstitutionen: der DARIAH-Fahrplan 2016*, in "BIBLIOTHEK Forschung und Praxis", 2016, vol. 39, n^{o} 3, p. 350–357, https://hal.inria.fr/hal-01285917.

Articles in National Peer-Reviewed Journal

[17] C. RIONDET.*Journaux Intimes de Clandestinité : Le cas de Léo Hamon (19040-1944)*, in "Vingtième siècle", October 2016, https://hal.inria.fr/hal-01416988.

Invited Conferences

[18] M. PUREN.*A l'épreuve de l'hétérogénéité : données de recherche et interdisciplinarité : L'exemple du projet européen IPERION-CH*, in "DHnord 2016 - Humanités numériques: théories, débats, approches critiques", Lille, France, Maison Européenne des Sciences de l'Homme et de la Société , November 2016, https://hal.archives-ouvertes.fr/hal-01408951.

[19] M. PUREN, C. RIONDET.*Research data management, a chance for Open Science. Methods and tutorials to create a Data Management Plan (DMP)*, in "DARIAH's Humanities at Scale Winter School", Prague, Czech Republic, Dariah and Humanities at Scale, October 2016, https://hal.inria.fr/hal-01416978.

[20] C. RIONDET.*De Gaulle et l'organisation de la résistance à Paris*, in "De Gaulle et Paris", Paris, France, Comité d'Histoire de la ville de Paris and Fondation Charles de Gaulle, April 2016, https://halshs.archives-ouvertes.fr/halshs-01301278.

International Conferences with Proceedings

[21] L. BARQUE.*A survey on semantic productivity*, in "Workshop Expanding the lexicon", Trier, Unknown or Invalid Region, D. GRAS (editor), 2016, https://halshs.archives-ouvertes.fr/halshs-01428252.

[22] R. BAWDEN, B. CRABBÉ.*Boosting for Efficient Model Selection for Syntactic Parsing*, in "COLING 2016 - 26th International Conference on Computational Linguistics", Osaka, Japan, December 2016, p. 1-11, https://hal.inria.fr/hal-01391743.

[23] T. BERNARD.*Modelling Subordinate Conjunctions in STAG: A Discourse Perspective*, in "28th European Summer School in Logic, Language & Information", Bozen-Bolzano, Italy, Proceedings of the ESSLLI 2016 Student Session, August 2016, 13, https://hal.inria.fr/hal-01363201.

[24] T. BERNARD, L. DANLOS.*Modelling Discourse in STAG: Subordinate Conjunctions and Attributing Phrases*, in "12th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+12)", Dûsseldorf, Germany, Proceedings of the 12th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+12), June 2016, p. 38-47, https://hal.archives-ouvertes.fr/hal-01329539.

[25] M. COAVOUX, B. CRABBÉ.*Neural Greedy Constituent Parsing with Dynamic Oracles*, in "Association for Computational Linguistics (ACL)", Berlin, Germany, 2016, https://hal.inria.fr/hal-01353734.

[26] L. DANLOS, M. CONSTANT, L. BARQUE.*Improvement of VerbNet-like resources by frame typing*, in "Workshop on Grammar and Lexicon: interactions and interfaces (GramLex)", Osaka, Japan, Proceedings of the Workshop on Grammar and Lexicon: interactions and interfaces (GramLex), The COLING 2016 Organizing Committee, December 2016, https://hal.inria.fr/hal-01392822.

[27] L. DANLOS, A. MASKHARASHVILI, S. POGODALLA.*Interfacing Sentential and Discourse TAG-based Grammars*, in "The 12th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+12)", Düsseldorf, Germany, Proceedings of the 12th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+12), June 2016, https://hal.inria.fr/hal-01328697.

[28] M. DJEMAA, M. CANDITO, P. MULLER, L. VIEU.*Corpus annotation within the French FrameNet: a domain-by-domain methodology*, in "Tenth International Conference on Language Resources and Evaluation (LREC 2016)", Portorož, Slovenia, May 2016, https://hal.archives-ouvertes.fr/hal-01391526.

[29] M. LHIOUI, K. HADDAR, L. ROMARY.*A new method for interoperability between lexical resources using MDA approach*, in "AISI 2016 The 2nd International Conference on Advanced Intelligent Systems and Informatics", Cairo, Egypt, October 2016, https://hal.inria.fr/hal-01350524.

[30] H. MARTINEZ ALONSO, A. JOHANNSEN, B. PLANK.*Supersense tagging with inter-annotator disagreement*, in "Linguistic Annotation Workshop 2016", Berlin, Germany, August 2016, p. 43 - 48, https://hal.inria.fr/hal-01426747.

[31] O. MICHALON, C. RIBEYRE, M. CANDITO, A. NASR.*Deeper syntax for better semantic parsing*, in "Coling 2016 - 26th International Conference on Computational Linguistics", Osaka, Japan, December 2016, https://hal.archives-ouvertes.fr/hal-01391678.

[32] B. SAGOT.*Multilingual part-of-speech tagging with MElt*, in "23ème Conférence sur le Traitement Automatique des Langues Naturelles", Paris, France, July 2016, https://hal.inria.fr/hal-01352243.

[33] A. SIMONENKO, B. CRABBÉ, S. PRÉVOST.*Taraldsen's Generalization in Diachrony: Evidence from a Diachronic Corpus*, in "West Coast Conference on Formal Linguistics", Salt Lake City, United States, 2016, https://hal.inria.fr/hal-01353741.

[34] L. VIEU, P. MULLER, M. CANDITO, M. DJEMAA.*A general framework for the annotation of causality based on FrameNet*, in "Tenth International Conference on Language Resources and Evaluation (LREC 2016)", Portorož, Slovenia, May 2016, https://hal.archives-ouvertes.fr/hal-01391542.

[35] S. M. YIMAM, H. MARTÍNEZ ALONSO, M. RIEDL, C. BIEMANN.*Learning Paraphrasing for Multi-word Expressions*, in "MWE 2016 - Multiword Expression Workshop 2016", Berlin, Germany, August 2016, https://hal.inria.fr/hal-01426749.

National Conferences with Proceeding

[36] T. BERNARD.*Conjonctions de subordination, verbes de dire et d'attitude propositionnelle : une modélisation STAG pour le discours*, in "18ème Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues", Paris, France, Actes de la conférence conjointe JEP-TALN-RECITAL 2016, July 2016, vol. volume 3 : RECITAL, p. 27-39, https://hal.archives-ouvertes.fr/hal-01357125.

Conferences without Proceedings

[37] A. BAILLOT.*A certification model for digital scholarly editions: Towards peer review-based data journals in the humanities*, in "Digital Scholarly Editing: Theory, Practice, Methods", Anvers, Belgium, Université d'Anvers, October 2016, https://halshs.archives-ouvertes.fr/halshs-01392880.

[38] A. SIMONENKO, B. CRABBÉ, S. PREVOST.*Effects of literary form on grammatical changes: A treebank study*, in "49th Annual Meeting of the Societas Linguistica Europaea (SLE 2016)", Naples, Italy, 2016, https://hal.inria.fr/hal-01365263.

[39] A. SIMONENKO, B. CRABBÉ, S. PRÉVOST.*Quantificational dimension of Taraldsen's Generalisation*, in "New Ways of Analyzing Syntactic Variation 2 (NWASV 2)", Ghent, Belgium, 2016, https://hal.inria.fr/hal-01353738.

[40] A. SIMONENKO, B. CRABBÉ, S. PRÉVOST.*Taraldsen's Generalisation in Medieval French*, in "Diachronic Generative Syntax conference (DIGS)", Ghent, Belgium, 2016, https://hal.inria.fr/hal-01353736.

Scientific Books (or Scientific Book chapters)

[41] T. BLANKE, C. KRISTEL, L. ROMARY.*Crowds for Clouds: Recent Trends in Humanities Research Infrastructures*, in "Cultural Heritage Digital Tools and Infrastructures", A. BENARDOU, E. CHAMPION, C. DALLAS, L. HUGHES (editors), Taylor & Francis Group, 2016, https://hal.inria.fr/hal-01248562.

[42] L. ROMARY.*Eléments d'une communication scientifique ouverte et publique*, in "Publier, éditer, éditorialiser. Nouveaux enjeux de la production numérique", L. CALDERAN, P. LAURENT, H. LOWINGER, J. MILLET (editors), Information & Stratégie, De Boek, 2016, https://hal.inria.fr/hal-01328192.

Research Reports

[43] J. BAETEN, P. ESTRAILLIER, C. KIRCHNER, A. MOATTI, L. ROMARY.*Open Access in Japan – a multi-institutional perspective*, Ambassade de France au Japon, March 2016, https://hal.archives-ouvertes.fr/hal-01290936.

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

Other Publications

[45] P. BANSKI, B. GAIFFE, P. LOPEZ, S. MEONI, L. ROMARY, T. SCHMIDT, P. STADLER, A. WITT.*Wake up, standOff!*, September 2016, TEI Conference 2016, https://hal.inria.fr/hal-01374102.

[46] J. BOWERS, L. ROMARY.*Deep encoding of etymological information in TEI*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01296498.

[47] L. DANLOS, B. CRABBÉ.*Natural Language Processing, 60 years after the Chomsky-Schützenberger hierarchy*, March 2016, Marie Paule Schützenberger 20 ans après, https://hal.inria.fr/hal-01392829.

[48] E. NIVAULT, A. MONTEIL, L. FARHI, L. ROMARY.*Implementation of the IFIP Digital Library in the HAL open publication repository*, June 2016, Libraries Opening Paths to Knowledge: LIBER Annual Conference 2016, Poster, https://hal.inria.fr/hal-01327170.

[49] L. ROMARY, M. PUREN.*Datasets of IPERION CH*, March 2016, Atelier interdisciplinaire « Matériaux du patrimoine et patrimoine matériel », https://hal.inria.fr/hal-01289058.

[50] L. ROMARY.*Elements of a scientific communication policy*, July 2016, ETD 2016 "Data and Dissertations", https://hal.inria.fr/hal-01345623.

[51] L. ROMARY.*The Text Encoding Initiative: 30 years of accumulated wisdom and its potential for a bright future*, September 2016, Language Technologies & Digital Humanities 2016, https://hal.inria.fr/hal-01374597.

References in notes

[52] A. ABEILLÉ, N. BARRIER.*Enriching a French Treebank*, in "Proceedings of LREC'04", Lisbon, Portugal, 2004.

[53] A. ABEILLÉ, L. CLÉMENT, F. TOUSSENEL.*Building a treebank for French*, in "Treebanks: building and using parsed corpora", A. ABEILLÉ (editor), Kluwer academic publishers, 2003, p. 165-188.

[54] C. F. BAKER, C. J. FILLMORE, J. B. LOWE.*The Berkeley FrameNet project*, in "Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 1", Montreal, Canada, 1998, p. 86-90.

[55] S. BOSCH, S. K. CHOI, É. VILLEMONTE DE LA CLERGERIE, A. CHENGYU FANG, G. FAASS, K. LEE, A. PAREJA-LORA, L. ROMARY, A. WITT, A. ZELDES, F. ZIPSER.*[tiger2] As a standardized serialisation for ISO 24615 - SynAF*, in "TLT11 - 11th international workshop on Treebanks and Linguistic Theories - 2012", Lisbon, Portugal, I. HENDRICKX, S. KÜBLER, K. SIMOV (editors), Ediçoes Colibri, November 2012, p. 37-60, https://hal.inria.fr/hal-00765413.

[56] P. BOULLIER.*Range Concatenation Grammars*, in "New Developments in Parsing Technology", H. BUNT, J. CARROLL, G. SATTA (editors), Text, Speech and Language Technology, Kluwer Academic Publishers, 2004, vol. 23, p. 269–289.

[57] M. CANDITO, B. CRABBÉ, P. DENIS, F. GUÉRIN.*Analyse syntaxique du français : des constituants aux dépendances*, in "Proceedings of TALN'09", Senlis, France, 2009.

[58] D. CHIANG.*Statistical parsing with an automatically-extracted Tree Adjoining Grammar*, in "Proceedings of the 38th Annual Meeting on Association for Computational Linguistics", 2000, p. 456–463.

[59] M. COLLINS.*Head Driven Statistical Models for Natural Language Parsing*, University of Pennsylvania, Philadelphia, 1999.

[60] B. CRABBÉ, M. CANDITO.*Expériences D'Analyse Syntaxique Statistique Du Français*, in "Actes de la 15ème Conférence sur le Traitement Automatique des Langues Naturelles (TALN'08)", Avignon, France, 2008, p. 45–54.

[61] B. CRABBÉ.*Multilingual discriminative lexicalized parsing*, in "Empirical Methods in Natural Language Processing", Lisbon, Portugal, 2015, https://hal.inria.fr/hal-01186018.

[62] L. DANLOS.*Discourse Verbs and Discourse Periphrastic Links*, in "Second International Workshop on Constraints in Discourse", Maynooth, Ireland, 2006.

[63] L. DANLOS.*D-STAG : un formalisme pour le discours basé sur les TAG synchrones*, in "Proceedings of TALN 2007", Toulouse, France, 2007.

[64] L. DANLOS, A. MASKHARASHVILI, S. POGODALLA.*An ACG Analysis of the G-TAG Generation Process*, in "INLG 2014 - 8th International Natural Language Generation Conference", Philadelphia, PA, United States, M. MITCHELL, K. MCCOY, D. MCDONALD, A. CAHILL (editors), Proceedings of the 8th International Natural Language Generation Conference (INLG), Association for Computational Linguistics, June 2014, p. 35-44, https://hal.inria.fr/hal-00999595.

[65] L. DANLOS, A. MASKHARASHVILI, S. POGODALLA.*An ACG View on G-TAG and Its g-Derivation*, in "LACL 2014 - Eight International Conference on Logical Aspects of Computational Linguistics", Toulouse, France, N. ASHER, S. SOLOVIEV (editors), Springer, June 2014, vol. 8535, p. 70-82 [*DOI :* 10.1007/978-3-662-43742-1_6], https://hal.inria.fr/hal-00999633.

[66] L. DANLOS, A. MASKHARASHVILI, S. POGODALLA.*Génération de textes : G-TAG revisité avec les Grammaires Catégorielles Abstraites*, in "TALN 2014 - 21ème conférence sur le Traitement Automatique des Langues Naturelles", Marseille, France, Actes de TALN 2014, Association pour le Traitement Automatique des Langues, July 2014, vol. 1, p. 161-172, https://hal.inria.fr/hal-00999589.

[67] L. DANLOS, A. MASKHARASHVILI, S. POGODALLA.*Grammaires phrastiques et discursives fondées sur les TAG : une approche de D-STAG avec les ACG*, in "TALN 2015 - 22e conférence sur le Traitement Automatique des Langues Naturelles", Caen, France, Actes de TALN 2015, Association pour le Traitement Automatique des Langues, June 2015, p. 158-169, https://hal.inria.fr/hal-01145994.

[68] P. DENIS, B. SAGOT.*Coupling an annotated corpus and a lexicon for state-of-the-art POS tagging*, in "Language Resources and Evaluation", 2012, vol. 46, n^{o} 4, p. 721-736 [*DOI :* 10.1007/s10579-012-9193-0], https://hal.inria.fr/inria-00614819.

[69] D. FIŠER.*Leveraging Parallel Corpora and Existing Wordnets for Automatic Construction of the Slovene Wordnet*, in "Proceedings of L&TC'07", Poznań, Poland, 2007.

[70] D. FIŠER, B. SAGOT.*Constructing a poor man's wordnet in a resource-rich world*, in "Language Resources and Evaluation", 2015, 35 [*DOI :* 10.1007/s10579-015-9295-6], https://hal.inria.fr/hal-01174492.

[71] N. IDE, T. ERJAVEC, D. TUFIS.*Sense Discrimination with Parallel Corpora*, in "Proc. of ACL'02 Workshop on Word Sense Disambiguation", 2002.

[72] D. KLEIN, C. D. MANNING.*Accurate Unlexicalized Parsing*, in "Proceedings of the 41st Meeting of the Association for Computational Linguistics", 2003.

[73] R. T. MCDONALD, F. C. N. PEREIRA.*Online Learning of Approximate Dependency Parsing Algorithms*, in "Proc. of EACL'06", 2006.

[74] J. NIVRE, M. SCHOLZ.*Deterministic Dependency Parsing of English Text*, in "Proceedings of Coling 2004", Geneva, Switzerland, COLING, Aug 23–Aug 27 2004, p. 64–70.

[75] S. PETROV, L. BARRETT, R. THIBAUX, D. KLEIN.*Learning Accurate, Compact, and Interpretable Tree Annotation*, in "Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics", Sydney, Australia, Association for Computational Linguistics, July 2006.

[76] S. PETROV, D. KLEIN.*Improved Inference for Unlexicalized Parsing*, in "Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference", Rochester, New York, Association for Computational Linguistics, April 2007, p. 404–411, http://aclweb.org/anthology/N07-1051.

[77] S. PETROV, R. T. MCDONALD.*Overview of the 2012 Shared Task on Parsing the Web*, in "Proceedings of the First Workshop on Syntactic Analysis of Non-Canonical Language (SANCL), a NAACL-HLT 2012 workshop", Montréal, Canada, 2012.

[78] P. RESNIK, D. YAROWSKY.*A perspective on word sense disambiguation methods and their evaluation*, in "ACL SIGLEX Workshop Tagging Text with Lexical Semantics: Why, What, and How?", Washington, D.C., USA, 1997.

[79] B. SAGOT, P. BOULLIER.*Les RCG comme formalisme grammatical pour la linguistique*, in "Actes de TALN'04", Fès, Maroc, 2004, p. 403-412.

[80] B. SAGOT, P. BOULLIER.*SxPipe 2: architecture pour le traitement présyntaxique de corpus bruts*, in "Traitement Automatique des Langues (T.A.L.)", 2009, vol. 50, n^{o} 1.

[81] B. SAGOT, L. CLÉMENT, É. VILLEMONTE DE LA CLERGERIE, P. BOULLIER.*The Lefff 2 syntactic lexicon for French: architecture, acquisition, use*, in "Proc. of LREC'06", 2006, http://hal.archives-ouvertes.fr/docs/00/41/30/71/PDF/LREC06b.pdf.

[82] B. SAGOT, D. FIŠER.*Building a free French wordnet from multilingual resources*, in "OntoLex", Marrakech, Morocco, May 2008, https://hal.inria.fr/inria-00614708.

[83] B. SAGOT.*Automatic acquisition of a Slovak lexicon from a raw corpus*, in "Lecture Notes in Artificial Intelligence 3658 (© Springer-Verlag), Proceedings of TSD'05", Karlovy Vary, Czech Republic, September 2005, p. 156–163.

[84] B. SAGOT.*Linguistic facts as predicates over ranges of the sentence*, in "Lecture Notes in Computer Science 3492 (© Springer-Verlag), Proceedings of LACL'05", Bordeaux, France, April 2005, p. 271–286.

[85] 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)", Malte Valletta, 2010.

[86] D. SEDDAH, M. CANDITO, B. CRABBÉ.*Cross Parser Evaluation and Tagset Variation: a French Treebank Study*, in "Proceedings of the 11th Internation Conference on Parsing Technologies (IWPT'09)", Paris, France, October 2009, p. 150-161.

[87] D. SEDDAH, G. CHRUPAŁA, Ö. ÇETINOGLU, J. VAN GENABITH, M. CANDITO.*Lemmatization and Statistical Lexicalized Parsing of Morphologically-Rich Languages*, in "Proceedings of the NAACL/HLT Workshop on Statistical Parsing of Morphologically Rich Languages - SPMRL 2010", États-Unis Los Angeles, CA, 2010.

[88] 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.

[89] D. SEDDAH.*Exploring the Spinal-Stig Model for Parsing French*, in "Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010)", Malte Malta, 2010.

[90] S. TAGLIAMONTE, D. DENIS.*Linguistic ruin? LOL! Instant messaging and teen language*, in "American Speech", 2008, vol. 83, n^{o} 1, 3.

[91] F. THOMASSET, É. VILLEMONTE DE LA CLERGERIE.*Comment obtenir plus des Méta-Grammaires*, in "Proceedings of TALN'05", Dourdan, France, ATALA, June 2005.

[92] É. VILLEMONTE DE LA CLERGERIE.*From Metagrammars to Factorized TAG/TIG Parsers*, in "Proceedings of IWPT'05", Vancouver, Canada, October 2005, p. 190–191.

[93] VOSSEN, P..*EuroWordNet: a multilingual database with lexical semantic networks for European Languages*, Kluwer, Dordrecht, 1999.

[94] H. YAMADA, Y. MATSUMOTO.*Statistical Dependency Analysis with Support Vector Machines*, in "The 8th International Workshop of Parsing Technologies (IWPT2003)", 2003.

[95] G. VAN NOORD.*Error Mining for Wide-Coverage Grammar Engineering*, in "Proc. of ACL 2004", Barcelona, Spain, 2004.

Project-Team ALPINES

Algorithms and parallel tools for integrated numerical simulations

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

IN PARTNERSHIP WITH:
CNRS
Université Pierre et Marie Curie (Paris 6)

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:

6.1.1. - Continuous Modeling (PDE, ODE)
6.1.4. - Multiscale modeling
6.1.5. - Multiphysics modeling
6.2.1. - Numerical analysis of PDE and ODE
6.2.5. - Numerical Linear Algebra
6.2.7. - High performance computing
6.3. - Computation-data interaction
6.3.1. - Inverse problems
7.1. - Parallel and distributed algorithms

Other Research Topics and Application Domains:

3.3.1. - Earth and subsoil
9.4.2. - Mathematics
9.4.3. - Physics

1. Members

Research Scientists

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

Faculty Members

Xavier Claeys [Univ. Paris VI, Associate Professor, HDR]
Frédéric Hecht [Univ. Paris VI, Professor, HDR]

Technical Staff

Simplice Donfack [Inria]
Ange Toulougoussou [Inria]

PhD Students

Hussam Al Daas [Inria]
Alan Ayala Obregon [Inria, granted by FP7 H NLAFET project]
Sebastien Cayrols [Inria]
Mohamed Ryadh Haferssas [Univ. Paris VI, until Sep 2016]
Zakariae Jorti [IFP Energies nouvelles, from Feb 2016]
Pierre Marchand [Inria]
Olivier Tissot [Inria, granted by FP7 H NLAFET project]

Post-Doctoral Fellows

Cihui Pan [Inria, until Sep 2016]
Pierre-Henri Tournier [Inria, until Nov 2016, granted by ANR MEDIMAX- project]

Visiting Scientists

Maria Barreda Vaya [Universitat Jaume, Castelló de la Plana (Spain), from Sep 2016 until Oct 2016]
Amanda Bienz [UIUC, USA, from Apr 2016 until Jul 2016]

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.

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
  - Computational kernels for numerical linear algebra
- Validation on numerical simulations

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 http://www.freefem.org/ff++), 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, 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 datat 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++.

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 direction preserving solvers in the context of multilevel domain decomposition methods with adaptive coarse spaces and multilevel incomplete decompositions. A judicious choice for the directions to be preserved through filtering or low rank approximations allows us to alleviate the effect of low frequency modes on the convergence.

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, which includes devising adapted domain decomposition approaches. The final objective is to bring the state of the art on boundary integral equations closer to contemporary industrial needs.

3.4. Computational kernels for numerical linear 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 aproach 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.

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

The research of F. Nataf on inverse problems is rather new since this activity was started from scratch in 2007. Since then, several papers were published in international journals and conference proceedings. All our numerical simulations were performed in FreeFem++.
We focus on methods related to time reversal techniques. 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 main idea is to take advantage of the reversibility of wave propagation phenomena such as it occurs in acoustics, elasticity or electromagnetism in a non-dissipative unknown medium to back-propagate signals to the sources that emitted them. Number of industrial applications have already been developped: touchscreen, medical imaging, non-destructive testing and underwater communications. 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 difraction 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. Since then, numerous applications of this physical principle have been designed, see [Fink, Renversement du temps, ondes et innovation. Ed. Fayard, 2009] or for numerical experiments [Larmat et al., Time-reversal imaging of seismic sources and application to the great sumatra earthquake. Geophys. Res. Lett., 33, 2006] and references therein.

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 natural 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 caracterizing 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 half petabyte of data, and will require 100 Exaflops per analysis as early as in 2020. 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.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

5.1.1.1. SIAM Siag on Supercomputing Best Paper Prize 2016

for the most outstanding paper published in 2012-2015 in a journal in the field of high performance computing. Co-authors are J. Demmel, L. Grigori, M. Hoemmen, and J. Langou, for the paper Communication-Optimal Parallel and Sequential QR and LU Factorizations, published in SIAM Journal on Scientific Computing 2012. Citation of the jury: *This is a cornerstone paper in Numerical Linear Algebra and Parallel Processing that lays down both theoretical and practical algorithmic frameworks for communication-avoiding algorithms. The paper provides powerful insights and renews attention on communication reduction both of which will have long-lasting and practical impact in parallel and distributed computing.*

5.1.1.2. Bull-Joseph Fourier 1st Prize 2015 (15 000 euros)

for our work *Imaging of cerebrovascular accident through High Performance Computing* by V. Dolean, F. Hecht, P. Jolivet, F. Nataf and P-H. Tournier. This was the sixth edition of this competition which corresponds to the French "Gordon Bell Prize".

6. New Software and Platforms

6.1. BFD

Block Filtering Decomposition preconditioner
KEYWORDS: Preconditioner - Linear system
FUNCTIONAL DESCRIPTION

Iterative methods are used in many industrial and academic applications to solve large sparse linear systems of equations, and preconditioning these methods is often necessary to accelerate their convergence. Several highly used preconditioners as incomplete LU factorizations are known to have scalability problems, often due to the presence of several low frequency modes that hinder the convergence of the iterative method. To address this problem, we work on filtering preconditioners. A judicious choice of the filtering vector allows to alleviate the effect of low frequency modes, and can accelerate significantly the convergence of the iterative method.

- Participants: Laura Grigori, Remi Lacroix and Frédéric Nataf
- Partners: CNRS - UPMC
- Contact: Laura Grigori
- URL: https://who.rocq.inria.fr/Laura.Grigori/

6.2. CALU : communication optimal algorithms for linear algebra

KEYWORDS: Communication avoiding - Linear algebra
FUNCTIONAL DESCRIPTION

CALU solves linear systems of equations Ax=b using Communication Avoiding LU.

- Contact: Laura Grigori
- URL: https://who.rocq.inria.fr/Laura.Grigori/

6.3. DPREPack

KEYWORD: Large scale
FUNCTIONAL DESCRIPTION

This library solves linear systems on parallel computers from PCs based on multicore processors to large scale computers. It implements recent parallel algorithms issued from domain decomposition methods and parallel approximate factorizations.

- Partners: CNRS - UPMC
- Contact: Laura Grigori
- URL: https://team.inria.fr/alpines/

6.4. 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 2^d-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: Frédéric Hecht
- URL: http://www.freefem.org/ff++/

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

6.6. LORASC

LORASC preconditioner

KEYWORD: Preconditioner

- Participants: Laura Grigori and Remi Lacroix
- Contact: Laura Grigori
- URL: not available

6.7. NFF

NFF Nested Filtering Factorization

KEYWORDS: Preconditioner - Interactive method - Linear system

- Participants: Laura Grigori, Frédéric Nataf and Long Qu
- Partners: Université Paris-Sud - UPMC
- Contact: Laura Grigori
- URL: not available

6.8. SparseToolbox

KEYWORDS: Preconditioner - Interactive method - Linear system

- Participants: Laura Grigori and Remi Lacroix
- Contact: Laura Grigori
- URL: not available

7. New Results

7.1. Communication avoiding algorithms

Our group continues to work on algorithms for dense and sparse linear algebra operations that minimize communication. During this year we focused on communication avoiding iterative methods and designing algorithms for computing rank revealing and low rank approximations of dense and sparse matrices.

In [9], we discuss sparse matrix-matrix multiplication (or SpGEMM), which is an important operation for many algorithms in scientific computing. In our previous work we have identified lower bounds on communication for this operation, which is the limiting factor of SpGEMM. Even though 3D (or 2.5D) algorithms have been proposed and theoretically analyzed in the flat MPI model on Erdos–Renyi matrices, those algorithms had not been implemented in practice and their complexities had not been analyzed for the general case. In this work, we present the first implementation of the 3D SpGEMM formulation that exploits multiple (intranode and internode) levels of parallelism, achieving significant speedups over the state-of-the-art publicly available codes at all levels of concurrencies. We extensively evaluate our implementation and identify bottlenecks that should be subject to further research.

In [10] we discuss algorithms that not only aim at minimizing communication, but they also aim at reducing the number of writes to secondary storage. Most of the prior work does not distinguish between loads and stores, i.e., between reads and writes to a particular memory unit. But in fact there are some current and emerging nonvolatile memory technologies (NVM) where writes can be much more expensive (in time and energy) than reads. NVM technologies are being considered for scientific applications on extreme scale computers and for cluster computing platforms, in addition to commodity computers.

This motivates us to first refine prior work on communication lower bounds of algorithms which did not distinguish between loads and stores to derive new lower bounds on writes to different levels of a memory hierarchy. When these new lower bounds on writes are asymptotically smaller than the previous bounds on the total number of loads and stores, we ask whether there are algorithms that attain them. We call such algorithms, that both minimize the total number of loads and stores (i.e., are CA), and also do asymptotically fewer writes than reads, *write-avoiding (WA)*. In this paper, we identify several classes of problems where either sequential or parallel WA algorithms exist, or provably cannot.

In [7] we introduce a new approach for reducing communication in Krylov subspace methods that consists of enlarging the Krylov subspace by a maximum of t vectors per iteration, based on the domain decomposition of the graph of A.We show in this paper that the enlarged Krylov projection subspace methods lead to faster convergence in terms of iterations and parallelizable algorithms with less communication, with respect to Krylov methods.

In this paper we focus on Conjugate Gradient (CG), a Krylov projection method for symmetric (Hermitian) positive definite matrices. We discuss two new versions of Conjugate Gradient. The first method, multiple search direction with orthogonalization CG (MSDO-CG), is an adapted version of MSD-CG with the A-orthonormalization of the search directions to obtain a projection method that guarentees convergence at least as fast as CG. The second projection method that we propose here, long recurrence enlarged CG (LRE-CG), is similar to GMRES in that we build an orthonormal basis for the enlarged Krylov subspace rather than finding search directions. Then, we use the whole basis to update the solution and the residual. We compare the convergence behavior of both methods using different A-orthonormalization and orthonormalization methods and then we compare the most stable versions with CG and other related methods. Both methods converge faster than CG in terms of iterations, but LRE-CG converges faster than MSDO-CG since it uses the whole basis to update the solution rather than only t search directions. And the more subdomains are introduced or the larger t is, the faster is the convergence of both methods with respect to CG in terms of iterations. For example, for $t = 64$ the MSDO-CG and LRE-CG methods converge in 75% up to 98 less iteration with respect to CG for the different test matrices.

In [12] we present an algorithm for computing a low rank approximation of a sparse matrix based on a truncated LU factorization with column and row permutations. We present various approaches for determining the column and row permutations that show a trade-off between speed versus deterministic/probabilistic accuracy. We show that if the permutations are chosen by using tournament pivoting based on QR factorization, then the obtained truncated LU factorization with column/row tournament pivoting, LU_CRTP, satisfies bounds on the singular values which have similarities with the ones obtained by a communication avoiding rank revealing QR factorization. Experiments on challenging matrices show that LU_CRTP provides a good low rank approximation of the input matrix and it is less expensive than the rank revealing QR factorization in terms of computational and memory usage costs, while also minimizing the communication cost. We also compare the computational complexity of our algorithm with randomized algorithms and show that for sparse matrices and high enough but still modest accuracies, our approach is faster.

7.2. Integral equation based domain decomposition

We kept on studying the convergence of classical domain decomposition strategies applied to multi-trace formulations (MTF). In the contribution [18], we present a gentle introduction to multi-trace formalism aimed at the domain decomposition community as well as analytical calculations in simple geometrical configuration where a full analysis of block-Jacoobi applied to MTF is possible. We only consider transmission problems in 1D with one or two interfaces. In [5], we generalize this analysis to arbitrary 2D or 3D transmission problems with arbitrary subdomain partitionning, only assuming that there is no junction point. The analysis holds mainly for completely homogeneous media with no material constrast, and in such a case we determine the spectrum of the multi-trace operator, as well as the spectrum of the Jacobi operator. We show that this spectrum only consists in a finite number of point values. In the more general case where the propagation medium is piecewise constant, this analysis still yields the location of the essential spectrum of the MTF and the Jacobi operator.

This analysis also led to an explicit expression for the inverse of the MTF operators for transmission problems in the case of perfectly homogeneous media. This was studied during the intership of Alan Ayala, and was described and tested numerically in 3D in the proceedings.

The analysis presented in [5] also shows that, in the case of purely homogeneous media, a block Jacobi strategy converges in a number of steps that exactly corresponds to the depth of the adjacency graph of the subdomain partition under consideration, which suggests a close relationship with Optimized Schwarz Methods (OSM), following the ideas of [20]. We investigated this point during the internship of Pierre Marchand, and we exhibited fully explicitely the exact relationship between block-Jacobi-MTF and OSM. Besides, we also generalized the analysis presented in [5] to the case of a competely heterogeneous problem, which involves abstract boundary integral operators that are not easily computable.

7.3. Multi-subdomain integral equations

In the context of boundary integral equations adapted to wave scattering in piecewise constant media in harmonic regime, we also made significant progress in the study of the single trace boundary integral formulation (STF) of the second kind originally introduced in [17]. This work was achieved in collaboration with Ralf Hiptmair and Elke Spindler (ETH Zürich). First of all, we proposed a version of this formulation for the solution to Maxwell's equations whereas, so far, it had been studied only in the context of scalar wave scattering (Helmholtz equation). In this direction, we conducted numerical experiments which confirmed the attractive properties of the matrices obtained when discretising such formulations (good accuracy, and good conditionning independent of discretisation parameters). For Maxwell's equations, we also established elementary theoretical results of STF 2nd kind such as Fredholmness of the corresponding integral operator.

So far, second kind STF had been stuudied for wave scatering problems where material contrasts only enter in the compact part of the partial differential operator, which is harmless regarding the Fredholmness of the corresponding boundary integral operator. Thus, in [19], we investigated the case where material contrasts come into play in the principal part of the operator, considering a pure diffusion-transmission problem. In this case, we have been able to establish well-posedness (hence Fredholmness). A rather naive approach leads to choose Sobolev spaces of fractionnal order (half-integer) as main functional setting for this formulation. We showed that this formulation can be extended so as to make sense in the space of square integrable trace functions. This is much more handy a functional setting that allows in particular discontinuous Galerkin discretisations of the corresponding boundary integral equations.

7.4. Asymptotics for a semi-linear convex problem with small inclusion

In [16], in collaboration with Lucas Chesnel (Inria Defi) and Sergei Nazarov (Saint-Petersbourg University), we recently investigated the asymptotics of the solution to a semi-linear problem in 2D with Dirichlet boundary condition. The partial differential operator under consideration was $-\Delta u + (u)^{2p+1}$ where p is a positive integer. The computational domain is assumed to contain a small Dirichlet obstacle of size $\delta > 0$. Using the method of matched asymptotic expansions, we compute an asymptotic expansion of the solution as δ tends to zero. Its relevance was justified by proving a rigorous error estimate. Then we construct an approximate model, based on an equation set in the limit domain without the small obstacle, which provides a good approximation of the far field of the solution of the original problem. The interest of this approximate model lies in the fact that it leads to a variational formulation which is very simple to discretize. We obtained numerical experiments to illustrate the analysis.

7.5. Time-dependent wave splitting and source separation

Starting from classical absorbing boundary conditions, we (M. Grote, M. Kray, F. Nataf and F. Assous) propose a method for the separation of time-dependent scattered wave fields due to multiple sources or obstacles. More precisely, we propose a method to determine the separate outgoing components of the incident and scattered wave fields for time-dependent scattering problems. In the case of two superposed wave fields, our method applies to the following three typical configurations: two distinct localized sources with unknown time history each, a single (unknown) localized source with a nearby scatterer, or two separate scatterers illuminated by a known incident wave field. In all three cases, our method permits to recover the individual outgoing components from measurements of the total scattered field at a distance. In doing so, the particular nature of the scatterer, be it an im- penetrable well-defined obstacle or a penetrable localized inhomogeneity, is immaterial; only the purely outgoing character of the individual wave fields matters. In contrast to previous work, our approach is local in space and time, deterministic, and also avoids any a priori assumptions on the frequency spectrum of the signal. Numerical simulations in FreeFem++ in two space dimensions illustrate the usefulness of wave splitting for time-dependent scattering problems. This work was presented to several international conferences and was published in J. Comput. Phys. (2016).

7.6. SORAS GenEO-2

Optimized Schwarz methods (OSM) are very popular methods which were introduced by P.L. Lions (1989) for elliptic problems and by B. Després (1990) for propagative wave phenomena. We (R. Haferssas, P. Jolivet and F. Nataf) give here a theory for Lions' algorithm that is the genuine counterpart of the theory developed over the years for the Schwarz algorithm. The first step is to introduce a new symmetric variant of the ORAS (Optimized Restricted Additive Schwarz) algorithm that is suitable for the analysis of a two-level method. Then we build a coarse space for which the convergence rate of the two-level method is guaranteed regardless of the regularity of the coefficients. We show scalability results for thousands of cores for nearly incompressible elasticity and the Stokes systems with a continuous discretization of the pressure.

7.7. Numerical modeling and high speed parallel computing: new perspectives for tomographic microwave imaging for brain stroke detection and monitoring

These works deals with microwave tomography for brain stroke imaging using state-of-the-art numerical modeling and massively parallel computing. Iterative microwave tomographic imaging requires the solution of an inverse problem based on a minimization algorithm (e.g. gradient based) with successive solutions of a direct problem such as the accurate modeling of a whole-microwave measurement system. Moreover, a sufficiently high number of unknowns is required to accurately represent the solution. As the system will be used for detecting the brain stroke (ischemic or hemorrhagic) as well as for monitoring during the treatment, running times for the reconstructions should be reasonable. The method used is based on high-order finite elements, parallel preconditioners from the Domain Decomposition method and Domain Specific Language with open source FreeFem++ solver. This work, for which we got the Joseph Fourier-Bull prize, is supported by ANR grant MEDIMAX (ANR-13-MONU-0012) and was granted access to the HPC resources of TGCC@CEA under the allocations 2016-067519 and 2016- 067730 made by GENCI.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Contract with Total, February 2015 - February 2018, that funds the PhD thesis of Hussam Al Daas on enlarged Krylov subspace methods for oil reservoir and seismic imaging applications. Supervisor L. Grigori.
- Contract with IFPen, February 2016 - February 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.
- Contract with MentorGraphics, March 2016, that funds the internship of T. Freiman on circuit simulations. Supervisor F. Nataf.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

9.1.1.1. Medimax

ANR-MN (Modèles Numériques) October 2013 - September 2017

The main goal is the methodological and numerical development of a new robust inversion tool, associated with the numerical solution of the electromagnetic forward problem, including the benchmarking of different other existing approaches (Time Reverse Absorbing Condition, Method of Small-Volume Expansions, Level Set Method). This project involves the development of a general parallel open source simulation code, based on the high-level integrated development environment of FreeFem++, for modeling an electromagnetic direct problem, the scattering of arbitrary electromagnetic waves in highly heterogeneous media, over a wide frequency range in the microwave domain. The first applications considered here will be medical applications: microwave tomographic images of brain stroke, brain injuries, from both synthetic and experimental data in collaboration with EMTensor GmbH, Vienna (Austria), an Electromagnetic Medical Imaging company.

9.1.1.2. 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.

9.1.1.3. 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.

9.1.1.4. Soilμ-3D

ANR appel à projet générique October 2015 - September 2020

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. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. NLAFET

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

Programm: H2020

Duration: November 2015 - November 2018

Coordinator: UMEÅUniversitet

Partners:

Science and Technology Facilities Council (United Kingdom)

Computer Science Department, UmeåUniversitet (Sweden)

Mathematics Department, The University of Manchester (United Kingdom)

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.

9.2.1.2. EXA2CT

Title: EXascale Algorithms and Advanced Computational Techniques

Programm: FP7

Duration: September 2013 - August 2016

Coordinator: IMEC

Partners:

Fraunhofer-Gesellschaft Zur Foerderung Der Angewandten Forschung E.V (Germany)
Interuniversitair Micro-Electronica Centrum Vzw (Belgium)
Intel Corporations (France)
Numerical Algorithms Group Ltd (United Kingdom)
T-Systems Solutions for Research (Germany)
Universiteit Antwerpen (Belgium)
Universita della Svizzera italiana (Switzerland)
Université de Versailles Saint-Quentin-En-Yvelines. (France)
Vysoka Skola Banska - Technicka Univerzita Ostrava (Czech Republic)

Inria contact: Luc Giraud

Numerical simulation is a crucial part of science and industry in Europe. The advancement of simulation as a discipline relies on increasingly compute intensive models that require more computational resources to run. This is the driver for the evolution to exascale. Due to limits in the increase in single processor performance, exascale machines will rely on massive parallelism on and off chip, with a complex hierarchy of resources. The large number of components and the machine complexity introduce severe problems for reliability and programmability. The former of these will require novel fault-aware algorithms and support software. In addition, the scale of the numerical models exacerbates the difficulties by making the use of more complex simulation algorithms necessary, for numerical stability reasons. A key example of this is increased reliance on solvers. Such solvers require global communication, which impacts scalability, and are often used with preconditioners, increasing complexity again. Unless there is a major rethink of the design of solver algorithms, their components and software structure, a large class of important numerical simulations will not scale beyond petascale. This in turn will hold back the development of European science and industry which will fail to reap the benefits from exascale. The EXA2CT project brings together experts at the cutting edge of the development of solvers, related algorithmic techniques, and HPC software architects for programming models and communication. It will take a revolutionary approach to exascale solvers and programming models, rather than the incremental approach of other projects. We will produce modular open source proto-applications that demonstrate the algorithms and programming techniques developed in the project, to help boot-strap the creation of genuine exascale codes.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

- J. Demmel, UC Berkeley, USA
- R. Hipmair, ETH Zurich
- M. Grote (Université de Bâle, Suisse)
- F. Assous (Israel)

9.4. International Research Visitors

9.4.1. Visits to International Teams

9.4.1.1. Research Stays Abroad

- Laura Grigori has spent 5 months at UC Berkeley, from January 2016 to May 2016, as a visiting Professor/Researcher.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- SIAM Conference on Parallel Processing and Scientific Computing 2016 http://www.siam.org/meetings/pp16/, Paris. Co-chairs L. Grigori and R. Vuduc. The team Alpines as well as other members of Inria Paris and UPMC, were in charge of the local organization of this conference. There were a record number of 500 participants, 88 minisymposia sessions and 44 contributed talks.
- Frederic Hecht: Organizing the 8th FreeFem++ days (December 2016, Paris)

10.1.1.2. Member of the Organizing Committees

- Laura Grigori: Member of Program Committee of IEEE International Parallel and Distributed Processing Symposium, IPDPS 2016.
- Laura Grigori: Member of Program Committee of HiPC 2016, IEEE Int'l Conference on High Performance Computing.
- Laura Grigori: Member of Program Committee of IEEE Cluster 2016.
- Laura Grigori: Member of Organizing Committee of 5th SIAM Workshop on Combinatorial Scientific Computing (CSC) 2016.

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

10.1.2.1.1. 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/.
- June 2013 – current. Area editor for Parallel Computing Journal, Elsevier
- January 2016 – current. Associate Editor, SIAM Journal on Scientific Computing.
- January 2016 – current. Editorial board, Numerical linear algebra with applications Journal, Wiley.

10.1.2.1.2. Frédéric Nataf

- 2014 – current. Associate Editor, Journal of Numerical Mathematics, de Gruyter.

10.1.3. Invited Talks

- Laura Grigori: Invited plenary speaker, Fifth IMA Conference on Numerical Linear Algebra and Optimization, Birmingham, (http://www.ima.org.uk/conferences/conferences_calendar/5th_ima_conference_on_numerical_linear_algebra_and_optimisation.cfm.html, September 7-9, 2016.
- Frederic Hecht: tutorial with FreeFem++, CEMRACS 2016 CIRM, Luminy, Marseille, july 18-22, 2016.
- Frederic Hecht: Graduate Course: An introduction to scientific computing using free software FreeFem++ , The Fields Institute for Research in Mathematical Sciences , Toronto, Canada , 7-17 March . 2016.
- Frederic Nataf: tutorial on Domain Decomposition Methods, CEMRACS 2016 CIRM, Luminy, Marseille, july 18-22, 2016.
- Frederic Nataf: Invited plenary speaker, Franco-Scottish Seminar 2016: Linear Algebra and Parallel Computing at the Heart of Scientific Computing. 21 September 2016.

10.1.4. Leadership within the Scientific Community

- Laura Grigori: Chair of the SIAM SIAG on Supercomputing (SIAM special interest group on supercomputing), January 2016 - December 2017. Nominated by a Committee and elected by the members of this SIAG.
- Laura Grigori: Member of the PRACE (Partnership for Advanced Computing in Europe, http://www.prace-ri.eu/) Scientific Steering Committee, September 2016 - current.
- Laura Grigori: Steering committee member, Challenge 7: Information and communication society, ANR (Comité de Pilotage , Défi 7), since November 2016.

10.1.5. 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.
- Xavier Claeys: Member of the hiring committee for the position of maître de conférence 26 MCF 4346 at Université Claude Bernard Lyon 1, Spring 2016

10.1.6. Research Administration

10.1.6.1. Xavier Claeys

- Correspondent of the Association pour les Mathématiques en Interactions avec les Entreprises et la Société (AMIES) for Laboratoire Jacques-Louis Lions

10.1.6.2. Laura Grigori

- Member of the Director Committee (Comité Directeur) of GIS Geosciences franciliennes, since November 2015.
- Member (nominated by Inria) of the Scientific Committee of the Mathematics Department of UPMC

10.1.6.3. Frédéric Hecht

- Assistant Director of Fédération de Recherche en Mathématiques de Paris Centre, Fédération de Recherche en Mathématiques de Paris Centre, Since 2010

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

10.2.1.1. Xavier Claeys

- Academic year 2015-2016, total number of course hours: 192 hrs
  - Licence 1: Calculus, 12hrs, UPMC
  - Licence 2: Orientation and Professional Insertion, 40 hrs, UPMC
  - Master 1: Basic computing, 72 hrs of practical works in C++, UPMC
  - Master 1: Scientific computing, 44 hrs of lectures, UPMC
  - Master 2: Solving PDEs by using finite element methods, 18 hrs of practical works in C++, UPMC

10.2.1.2. Laura Grigori

- Spring 2016, UC Berkeley: Co-teaching a class with J. Demmel on Communication Avoiding Algorithms https://who.rocq.inria.fr/Laura.Grigori/TeachingDocs/CS-294_Spr2016/CS-294_CA_Spr2016.html. The class is available for mathematics students as a *Hot Topics Course in Mathematics*, and for Computer Science students as a *Special Topics* class. It was mainly designed for graduate students.
- 4-hour lecture at the XVII Jacques-Louis Lions Spanish-French School on Numerical Simulation in Physics and Engineering, June 2016, Gijon, Spain.

10.2.1.3. Frédéric Hecht

- Academic year 2015-2016, total number of course hours: 192 hrs
  - Master 1: Initiation au C++, 24hrs, M1, Université Pierre-et-Marie Curie Paris 6, France
  - Master 2: Des EDP à leur résolution par la méthode des éléments finis (MEF), 24hrs, M2, Université Pierre-et-Marie Curie Paris 6, France
  - Master 2: Numerical methods for fluid mechanics, 10hrs, M2, Université Pierre-et-Marie Curie Paris 6, France
  - Master 2: Calcul scientifique 3 / projet industriel FreeFem++, 28hrs, M2, Université Pierre-et-Marie Curie Paris 6, France
  - Master 2: Ingénierie 1 / Logiciel pour la simulation (FreeFem++), 21hrs, M2, Université Pierre-et-Marie Curie Paris 6, France
  - Master 2: Ingénierie 2 / Projet collaboratif, 21hrs, M2, Université Pierre-et-Marie Curie Paris 6, France
- CEMRACS 2016 : Calcul scientifique à haute performance, 3h course on Tutorial with FreeFem++. The slides are available on F. Hecht website, and videos can be found here https://www.youtube.com/watch?v=NQy2kZQGBbg.

10.2.1.4. Frédéric Nataf

- CEMRACS 2016 : Calcul scientifique à haute performance, 3h course on Domain decomposition methods The slides are available on F. Nataf's website, videos can be found here https://www.youtube.com/watch?time_continue=10672&v=t-rgIFcN6w4.
- Spring 2016: Course on *Domain Decomposition Methods*, Master 2nd year Mathematics & Applications, University Pierre and Marie Curie
- Winter 2016: Course on *Domain Decomposition Methods*, Master 2nd year, Mathematics & Applications, ENSTA and UVSQ

10.2.2. Supervision

- PhD in progress: Alan Ayala, since October 2015 (funded by NLAFET H2020 project), co-advisors Xavier Claeys and Laura Grigori.
- PhD in progress : Sebastien Cayrols, since October 2013 (funded by Maison de la simulation), adivsor Laura Grigori.

- PhD in progress: Hussam Al Daas, since February 2015 (funded by contract with Total), advisor Laura Grigori.
- PhD in progress: Olivier Tissot, since October 2015 (funded by NLAFET H2020 project), advisor Laura Grigori.
- PhD in progress: Pierre Marchand, since October 2016 (funded by ANR NonLocalDD project), advisors Xavier Claeys et Frédéric Nataf.
- PhD in progress: Zakariae Jorti, since February 2016 (funded by IFPen), advisor Laura Grigori.
- PhD : Ryadh Haferssas, defended december 2016 (funded by Ecole Doctorale, UPMC), advisor F. Nataf
- PhD : Mireille El-HAddad, defended december 2016(t UPMC), advisors F. Hecht and T. Sayah.
- HDR: X. CLAEYS, Boundary integral equations of time harmonic wave scattering at complex structures, Accreditation to supervise research, (UPMC Paris 6), February 2016.

10.3. Popularization

- 2 Conferences on being a mathematicien: **Applied Mathematics, Research...**, in the high school Viollet le Duc - Villiers-Saint-Frédéric, France for 200 high school students in 11th and 12th grade, april 2016.
- Part of the teaching duty of Xavier Claeys (around 70 hours total) has been dedicated to "orientation and insertion professionnelle". This has consisted in familiarizing young students (2nd year bachelor) with potential curricula, professionnal activities and networks related to mathematics.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] X. CLAEYS.*Boundary integral equations of time harmonic wave scattering at complex structures*, Université Pierre et Marie Curie (UPMC Paris 6), February 2016, Habilitation à diriger des recherches, http://hal.upmc.fr/tel-01357729.

Articles in International Peer-Reviewed Journal

[2] I. BAJC, F. HECHT, S. ŽUMER.*A mesh adaptivity scheme on the Landau–de Gennes functional minimization case in 3D, and its driving efficiency*, in "Journal of Computational Physics", September 2016, vol. 321, p. 981–996 [*DOI :* 10.1016/J.JCP.2016.02.072], http://hal.upmc.fr/hal-01360423.

[3] L. CHESNEL, X. CLAEYS.*A numerical approach for the Poisson equation in a planar domain with a small inclusion*, in "BIT Numerical Mathematics", March 2016, https://hal.inria.fr/hal-01109552.

[4] X. CLAEYS.*Asymptotics of the eigenvalues of the Dirichlet-Laplace problem in a domain with thin tube excluded*, in "Quarterly of Applied Mathematics", December 2016, vol. 74, n^{o} 4, p. 595-605, https://hal.archives-ouvertes.fr/hal-01120422.

[5] X. CLAEYS.*Essential spectrum of local multi-trace boundary integral operators*, in "IMA Journal of Applied Mathematics", May 2016, https://hal.inria.fr/hal-01251212.

[6] X. CLAEYS, R. HIPTMAIR.*Integral Equations for Electromagnetic Scattering at Multi-Screens*, in "Integral Equations and Operator Theory", January 2016, vol. 84, n^{o} 1, 36 [*DOI :* 10.1007/s00020-015-2242-5], https://hal.inria.fr/hal-01251236.

[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 [*DOI :* 10.1137/140989492], https://hal.inria.fr/hal-01357899.

[8] 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, https://hal.archives-ouvertes.fr/hal-01216117.

[9] A. AZAD, G. BALLARD, A. BULUC, J. W. DEMMEL, L. GRIGORI, O. SCHWARTZ, S. TOLEDO.*Exploiting Multiple Levels of Parallelism in Sparse Matrix-Matrix Multiplication*, in "SIAM Journal on Matrix Analysis and Applications", November 2016, vol. 38, n^{o} 6, 27, https://hal.inria.fr/hal-01426294.

International Conferences with Proceedings

[10] C. ERIN, J. W. DEMMEL, L. GRIGORI, K. NICK, K. PENPORN, O. SCHWARTZ, V. HARSHA.*Write-Avoiding Algorithms*, in "Proceedings of IEEE International Parallel & Distributed Processing Symposium, IPDPS 2016", Chicago, United States, 2016, Short version of the technical report available at http://www.eecs.berkeley.edu/Pubs/TechRpts/2015/EECS-2015-163.pdf as Technical Report No. UCB/EECS-2015-163, https://hal.inria.fr/hal-01248678.

[11] P. JOLIVET, P.-H. TOURNIER.*Block Iterative Methods and Recycling for Improved Scalability of Linear Solvers*, in "SC16 - International Conference for High Performance Computing, Networking, Storage and Analysis", Salt Lake City, Utah, United States, Proceedings of SC16: International Conference for High Performance Computing, Networking, Storage and Analysis, November 2016, https://hal.archives-ouvertes.fr/hal-01357998.

Research Reports

[12] L. GRIGORI, S. CAYROLS, J. W. DEMMEL.*Low rank approximation of a sparse matrix based on LU factorization with column and row tournament pivoting*, inria, March 2016, n^{o} RR-8910, 35, https://hal.inria.fr/hal-01313856.

Other Publications

[13] S. DIB, T. SAYAH, V. GIRAULT, F. MURAT, C. BERNARDI, F. HECHT.*Finite element method for Darcy's problem coupled with the heat equation*, December 2016, working paper or preprint, http://hal.upmc.fr/hal-01416505.

[14] M. EL-HADDAD, F. HECHT, T. SAYAH.*Interface transport scheme of a two-phase flow by the method of characteristics*, April 2016, working paper or preprint, http://hal.upmc.fr/hal-01217940.

[15] G. VERGEZ, I. DANAILA, S. AULIAC, F. HECHT.*A finite-element toolbox for the stationary Gross-Pitaevskii equation with rotation*, 2016, working paper or preprint, http://hal.upmc.fr/hal-01277660.

References in notes

[16] L. CHESNEL, X. CLAEYS.*Small obstacle asymptotics for a 2D semi-linear convex problem*, accepted in Applicable Analysis, Oct. 2016..

[17] X. CLAEYS.*A single trace integral formulation of the second kind for acoustic scattering*, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2011, n$^{\mathrm{o}}$ 2011-14, https://www.sam.math.ethz.ch/sam_reports/reports_final/reports2011/2011-14.pdf.

[18] X. CLAEYS, V. DOLEAN, M. J. GANDER.*An introduction to Multitrace Formulations and Associated Domain Decomposition Solvers*, October 2015, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01207193.

[19] X. CLAEYS, R. HIPTMAIR, E. SPINDLER.*Second Kind Boundary Integral Equation for Multi-Subdomain Diffusion Problems*, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2016, n$^{\mathrm{o}}$ 2016-44, https://www.sam.math.ethz.ch/sam_reports/reports_final/reports2016/2016-44.pdf.

[20] F. NATAF, F. ROGIER, E. DE STURLER.*Optimal interface conditions for domain decomposition methods*, CMAP, Ecole Polytechnique, 1994.

Project-Team ANGE

Numerical Analysis, Geophysics and Ecology

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

Université Pierre et Marie Curie (Paris 6)

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:

6. - Modeling, simulation and control
6.1. - Mathematical Modeling
6.1.1. - Continuous Modeling (PDE, ODE)
6.1.4. - Multiscale modeling
6.1.5. - Multiphysics modeling
6.2. - Scientific Computing, Numerical Analysis & Optimization
6.2.1. - Numerical analysis of PDE and ODE
6.2.6. - Optimization
6.3. - Computation-data interaction
6.3.2. - Data assimilation

Other Research Topics and Application Domains:

3. - Environment and planet
3.3. - Geosciences
3.3.2. - Water: sea & ocean, lake & river
3.3.3. - Littoral
3.4. - Risks
3.4.1. - Natural risks
3.4.3. - Pollution
4. - Energy
4.3. - Renewable energy production
4.3.1. - Biofuels
4.3.2. - Hydro-energy

1. Members

Research Scientists

Jacques Sainte-Marie [Team leader, CEREMA, Senior Researcher, HDR]
Martin Parisot [Inria, Researcher]
Yohan Penel [CEREMA, Researcher]

Faculty Members

Bernard Di Martino [Inria secondment, on leave from Univ. Corse, Associate Professor, until Aug. 16, HDR]
Edwige Godlewski [Univ. Pierre et Marie Curie, Professor, HDR]
Cindy Guichard [Univ. Pierre et Marie Curie, Associate Professor]
Boris Haspot [CNRS secondment, on leave from Univ. Paris Dauphine, Associate Professor, from Sept. 16, HDR]
Anne Mangeney [CEREMA secondment, on leave from Univ. Paris Diderot, Professor, HDR]
Julien Salomon [Inria secondment, on leave from Univ. Paris Dauphine, Associate Professor, from Sept. 16, HDR]
Nicolas Seguin [Univ. Pierre et Marie Curie, Associate Professor, until Feb. 16, HDR]

Technical Staff

David Froger [Inria – SED 70%, until May 16]
Fabien Souillé [Inria, from Nov. 16]
Jérémy Ledoux [Univ. Paris Dauphine, from Sept. 16]

PhD Students

Nora Aïssiouene [Inria]
Léa Boittin [Inria]
Do Minh Hieu [Univ. Paris 13]
Dena Kazerani [Univ. Pierre et Marie Curie, until Oct. 2016]
Hugo Martin [IPGP]
Ethem Nayir [Univ. Pierre et Marie Curie]
Pierrick Quémar [Univ. Paris 13]
Sebastian Reyes-Riffo [Univ. Paris Dauphine]
Amandine Sergeant [IPGP, from Aug. 16]
Fabien Wahl [Univ. Pierre et Marie Curie]

Post-Doctoral Fellow

Pierre-Olivier Lamare [Inria]

Visiting Scientist

Marie-Odile Bristeau [Senior Researcher, retired]

Administrative Assistant

Maryse Desnous [Inria]

Others

Albin Grataloup [ENS Lyon, intern, from May 16 until July 16]
Emmanuel Audusse [Univ. Paris 13, Associate Professor]
Nicole Goutal [EDF, Senior Researcher, HDR]

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. Last but not least, efficient numerical resolutions of reduced-size models requires 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.

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 sense of the models (uniqueness, conservativity, entropy dissipation, ...) and propose effective numerical methods to estimate their solution in complex configurations (multi-dimensional, unstructured meshes, well-balanced, ...).

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 discretization 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, so-called shallow water assumptions, namely

- the horizontal 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 must rely on a special analysis.

3.2.1. Multilayer approach

As for the first shallow water assumption, *multi-layer* systems were proposed describing the flow as a superposition of Saint-Venant type systems [31], [33], [34]. Even if this approach has provided interesting results, layers are considered separate and non-miscible fluids, which imply strong limitation. That is why we proposed a slightly different approach [1], [2] based on Galerkin type decomposition along the vertical axis of all variables and leading, both for the model and its discretization, to more accurate results.

A kinetic representation of our multilayer model allows to derive robust numerical schemes endowed with 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 enable to guess the optimal number of layers.
- The introduction of several source terms due for instance to Coriolis forces 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 proved and conversely the possible loss of hyperbolicity cannot be characterized. 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.

When 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 integrating the 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-physic 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 zone 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 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 system Euler/Poisson, 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 such as threshold laws.
- Interaction between flows and floating structures is the challenge at the scale of the shallow water equations. This study needs 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 scheme. 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.

3.3. 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 has to 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

The hydrodynamic models comprise advection and sources terms. The conservation of the balance between the source terms, typically viscosity and friction, has a significant impact since the overall flow is generally a perturbation around one 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 potential waves and the physical velocity makes classical numerical schemes not efficient: 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 potential 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. Optimization

Numerical simulations are a very useful tool for the design of new processes, for instance in renewable energy or water decontamination. The optimization 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 optimization 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, biogeochemistry of oceans, resilience of society w.r.t. hazardous flows, ...

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 quasisteady 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 costly and generally not efficient in long term. The objective is to determine a configuration almost stable with 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 the 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.

5. Highlights of the Year

5.1. Highlights of the Year

While the theory and the numerics related to the nonlinear shallow water equations are extensively studied, the understanding of more complex models including dispersive ones is not achieved. Two PhD theses about these issues were defended in 2016 within the team (N. Aïssiouene and D. Kazerani). To go further, a collaboration with spanish collaborators from the university of Sevilla was launched with multiple trips in Spain and France resulting in a preprint [25]. The collaboration is expected to be made more formal in 2017.

Moreover, the team has been reinforced by two young engineers: J. Ledoux in the framework of the ANR project Hyflo-Eflu and F. Souillé. The latter recruitment has been allowed by the Inria ADT grant F2O ("Freshkiss to Others") and is aimed at easing the transfer of the Freshkiss code in cooperation with SciWorks Technologies.

6. New Software and Platforms

6.1. Freshkiss3D (FREe Surface Hydrodynamics using KInetic SchemeS)

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

- Participants: Nora Aïssiouene, Marie-Odile Bristeau, David Froger, Jacques Sainte-Marie, Fabien Souillé
- Formerly: Emmanuel Audusse, Anne-Céline Boulanger, Alain Dervieux, Raouf Hamouda, Bijan Mohammadi
- Partners: CEREMA – UPMC
- Contact: Jacques Sainte-Marie

A review of last developments is given in § 7.5.1. See [20] for recent numerical results obtained thanks to the Freshkiss3D software.

6.2. TSUNAMATHS

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, Raouf Hamouda, Jacques Sainte-Marie
- Contact: Jacques Sainte-Marie
- URL: http://ange.raoufhamouda.com/tsunami/en_motivation.htm

7. New Results

7.1. Modelling of complex flows

7.1.1. The Shallow Water model with Roof: derivation and simulation

Participants: Edwige Godlewski, Cindy Guichard, Martin Parisot, Jacques Sainte-Marie, Fabien Wahl.

In view of taking into account interactions with floating structures, a shallow water type model is derived. In a first step a constraint corresponding to a static roof is considered and a relaxation approach is proposed in order to solve the model numerically. A particular attention is paid to the energy law as an application to marine energy devices is planned. The CPR scheme proposed in [17] is adapted to our case and implemented in one space dimension. Finally the numerical results are tested on analytical solutions, as well stationary as non-stationary ones [26].

7.1.2. Modelling of Sediment Transport

Participants: Emmanuel Audusse, Léa Boittin, Martin Parisot, Jacques Sainte-Marie.

A new model for sediment transport in river context is proposed. The model is derived from the Navier-Stokes equations by performing simultaneously the thin layer approximation and the diffusive limit. The well-posedness of the model is studied in a simplified case.

7.1.3. Layer-averaged Euler and Navier-Stokes equation

Participants: Marie-Odile Bristeau, Bernard Di Martino, Cindy Guichard, Jacques Sainte-Marie.

In [3], we propose a strategy to approximate incompressible hydrostatic free surface Euler and Navier-Stokes models. The proposed strategy extends previous works approximating the Euler and Navier-Stokes systems using a multilayer description. Here, the required closure relations are obtained using an energy-based optimality criterion instead of an asymptotic expansion. Moreover, the layer-averaged description is successfully applied to the Navier-Stokes system with a general form of the Cauchy stress tensor.

7.1.4. Layerwise Discretization for Non-Hydrostatic flows

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

In collaboration with Enrique Fernández-Nieto (Sevilla).

The work presented in [25] aims at deriving a new semi-discretisation with respect to the vertical variable of the Euler equations. It results in a hierarchy of multilayer model involving both hydrostatic and non-hydrostatic parts of the pressure field. All models are proven to satisfy an energy inequality. Moreover, the linear dispersion relation is given for each one with an explicit formula which converges to the exact Airy formula when the number of layers goes to infinity.

7.1.5. Two-phase (grains/fluid) model for geophysical debris flows

Participant: Anne Mangeney.

We developped a thin-layer depth-averaged model describing the two-phase flow made of granular material saturated by a fluid and include compression/dilatation effects. We solved numerically these equations and were able to accurately reproduce laboratory experiments.

7.1.6. Multi-layer model for viscoplastic granular flows

Participant: Anne Mangeney.

In collaboration with Enrique Fernández-Nieto and Gladys Narbona-Reina (Sevilla).

A multi-layer model was developped to simulate granular flow dynamics and deposit based on viscoplastic behaviour ($\mu(I)$-rheology). The numerical model made it possible to reproduce for the first time the increase of runout distance of granular material when flowing on erodible beds.

7.2. Assessments of models by means of experimental data

7.2.1. Hydrodynamics and biology coupling in the context of algae growth

Participants: Marie-Odile Bristeau, Jacques Sainte-Marie.

In collaboration with BIOCORE (especially O. Bernard) in the framework of the IPL Algae in Silico.

Hydrodynamics in a high rate production reactor for microalgae cultivation affects light history perceived by the cells. The interplay between cell movement and medium turbidity leads to a light pattern forcing photosynthesis dynamics. The purpose of this multidisciplinary downscaling study is to reconstruct single cell trajectories in an open raceway and experimentally reproduce such high frequency light pattern to observe its effect on growth. We show that the frequency of such a realistic signal plays a determinant role in the dynamics of photosynthesis. This study highlights the need for experiments with more realistic light stimuli in order to better understand microalgal growth at high cell density.

7.2.2. 2D Drucker-Prager and $\mu(I)$ granular flow model

Participant: Anne Mangeney.

In collaboration with François Bouchut, Ioan Ionescu, Alexandre Ern, Christelle Lusso and Nathan Martin.

We developped 2D (horizontal/vertical) models of granular flows solving the yield behaviour of Drucker-Prager type laws using either a duality method or a regularization method. We included the effect of the lateral wall friction and get very good agreement with laboratory experiments of granular collapses over horizontal and inclined planes.

7.2.3. Analytical and numerical description of the static/flowing interface deduced from 2D Drucker-Prager model

Participant: Anne Mangeney.

In collaboration with François Bouchut, Alexandre Ern and Christelle Lusso.

We proposed analytical and numerical solution of the static/flowing interface and compared it with laboratory experiments of granular flows. Our study show how the static/flowing interface dynamics depends on the slope, friction angle, viscosity and normal velocity profiles.

7.2.4. Seismic inversion and numerical modelling of the force generated by landslides on the topography or by iceberg calving

Participant: Anne Mangeney.

By inverting the long period seismic signal to recover the force generating seismic waves and simulating this force with mechanical models of granular flows, we can provide a unique constraint on the dynamics of the phenomenon at stake and on its characteristics.

7.2.5. Data assimilation

Participants: Sebastian Reyes-Riffo, Julien Salomon.

In collaboration with Felix Kwok.

Taking advantage of a PROCORE-FRANCE/HONG KONG grant obtained in the latter spring, we work on a time-parallelization strategy for an assimilation algorithm. The target application also deals with wave energy: we aim at forecasting in real-time the characteristics of the wave coming on an extracting device, in order to adapt it in a continuous way.

7.3. Analysis of models in Fluid Mechanics

7.3.1. Weak solutions of multilayer Hydrostatic Flows

Participants: Bernard Di Martino, Boris Haspot, Yohan Penel.

We investigate the existence of global weak solutions for the multilayer model introduced by Audusse et al. [2] which is related to incompressible free surface flows. More precisely, in [22] we prove the global stability of weak solutions over the torus. We observe that this model admits the so called BD-entropy and a gain of integrability on the velocity in the spirit of the work of Mellet and Vasseur. The main difficulty comes from the terms describing the transfer of flux between the layers which are not taken into account in the immiscible case.

7.3.2. Hyperbolicity of the Layerwise Discretized Hydrostatic Euler equation: the bilayer case

Participants: Emmanuel Audusse, Edwige Godlewski, Martin Parisot.

In collaboration with Nina Aguillon (UPMC).

Several model of free surface flows described in the literature are based on a layerwise discretization of the Euler equations. The question addressed in the current work is about the hyperbolicity of the layerwise discretized model. More precisely, we focus on the 2-layer case and we prove the well-posedness of the Riemann problem in two dimensional framework. Due to the mass exchange, the 2D Riemann problem is not a simple extension of the 1D Riemann problem.

7.3.3. Normal mode perturbation for the shallow water equations

Participants: Emmanuel Audusse, Albin Grataloup, Yohan Penel.

This work focuses on the shallow water equations for a fluid flow in subcritical regime with an arbitrary topography. A normal mode perturbation was performed around a 1D steady state in the 2D model. The resulting system of ODE was studied in terms of eigenvalues of the corresponding matrix and the derivation of a dispersion relation.

7.3.4. Global well-posdness of the Euler-Korteweg system for small irrotational data

Participant: Boris Haspot.

In collaboration with C. Audiard (UPMC).

The Euler-Korteweg equations are a modification of the Euler equations that takes into account capillary effects. In the general case they form a quasi-linear system that can be recast as a degenerate Schrödinger type equation. We prove here that under a natural stability condition on the pressure, global well-posedness holds in dimension $d \geq 3$ for small irrotational initial data. The proof is based on a modified energy estimate, standard dispersive properties if $d \geq 5$, and a careful study of the nonlinear structure of the quadratic terms in dimensions 3 and 4 involving the theory of space time resonance.

7.4. Numerical methods for free-surface flows

7.4.1. A two-dimensional method for a dispersive shallow water model

Participants: Nora Aïssiouene, Marie-Odile Bristeau, Edwige Godlewski, Jacques Sainte-Marie.

We propose a numerical method for a two-dimensional dispersive shallow water system with topography. This model is a depth averaged Euler system and takes into account a non-hydrostatic pressure which implies to solve an incompressible system. From the variational formulation of the mixed problem proposed in [6], we apply a finite element method with compatible spaces to the two-dimensional problem on unstructured grids.

7.4.2. Numerical Discretization for Coriolis Effects

Participants: Emmanuel Audusse, Do Minh Hieu, Yohan Penel.

Efficient computations near the geostrophic equilibrium need to carefully design numerical schemes. This question is investigated in the context of colocated finite volume approach and extends previous works by Bouchut et al. [32], Dellacherie [35], Buet and Despres [36].

7.4.3. Optimization of topography

Participants: Sebastian Reyes-Riffo, Julien Salomon.

We work on a method to compute optimal topographies for wave-energy production. The first part of the work was devoted to the numerical analysis of the scheme used to simulate waves. In this way, we have obtained stability conditions that enable to couple it with an optimization loop.

7.4.4. An adaptive numerical scheme for solving incompressible two-phase and free-surface flows

Participant: Dena Kazerani.

We present a numerical scheme for solving two-phase or free surface flows. The interface/free surface is modelled using the level-set formulation. Besides, the mesh is anisotriopic and adapted at each iteration. The incompressible Navier–Stokes equations are temporally discretized using the method of characteristics and are solved at each time iteration by a first order Lagrange–Galerkin method. The level-set function representing the interface/free surface satisfies an advection equation which is also solved using the method of characteristics.

7.4.5. Propeler design

Participants: Jérémy Ledoux, Julien Salomon.

We work on a usual algorithm in propeler design: based on the so-called “Blade Element Momentum Theory”, this approach reduces the simulation to a 2D system by coupling the latter with a outer loop of low computational cost. So far, this method has not been analyzed mathematically, hence our interest.

7.5. Software developments and assessments

7.5.1. Improvements in the FRESHKISS3D code

Participants: Marie-Odile Bristeau, David Froger, Jacques Sainte-Marie, Fabien Souillé.

Several tasks have been achieved in the FRESHKISS3D software:

- Cython branch finalization (integration of second order in time and space numerical schemes)
- Project exportation on Gitlab.inria collaborative development platform
- New development tools set-up (Gitlab-ci, Git-lfs)
- Definition of new development rules and practices with gitlab
  - Use of the issue board
  - Review system and merge request rework
- Chlorides propagation in Vilaine river (Saur project)
  - Case definition in freshkiss3d
  - TracerSource class definition for floodgate modeling
  - VerticalDebit class definition for special boundary condition (siphon)
  - Simplified scenarios set-up (1day, 2days simulated)
  - First simulations and post processing
- New examples structure with introduction of two new cases to illustrate VerticalDebit and TracerSource class
- Various documentation updates

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

A contract has been made (120.000 euros) with SAUR, IAV (Institut d'Aménagement de la Vilaine) and Agence de l'eau Loire-Bretagne in collaboration with SciWorks Technologies. It deals with the modelling and the simulation of chlorides entry in the Vilaine reservoir.

The ANR project Hyflo-Eflu relies on a collaboration with the company "HydroTube Energie". It comprises the recruitment of a young engineer and regular meetings with industrial (Bordeaux) and academic partners (Nantes).

8.2. Bilateral Grants with Industry

P. Quémar's PhD thesis is funded by EDF ("thèse CIFRE"). His PhD is entitled "3D numerical simulations of environmental hydrolics: application to Telemac".

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Plasticity of geophysical flows and seismic emissions (2013-2016)

Participant: Anne Mangeney.

This project is funded by Sorbonne Paris Cité (80.000 euros) and is a collaboration between IPGP and Univ. Paris 13.

9.1.2. LRC Manon (2014-2018)

Participants: Edwige Godlewski, Yohan Penel, Nicolas Seguin.

CEA and Laboratory Jacques-Louis Lions launched a collaboration in order to carry out studies about complex fluids (modelling, numerical simulations and optimisation), in particular about compressible two-phase flows. This includes the derivation of strategies for model coupling, for instance in the case of an asymptotic hierarchy of models.

9.2. National Initiatives

9.2.1. 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).

9.2.2. ANR Hyflo-Eflu (2016-2020)

Participants: Jérérmy 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 optimization, 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 horizontal 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.

9.2.3. CNRS CORSURF (2016)

Participants: Bernard Di Martino, Cindy Guichard, Anne Mangeney, Jacques Sainte-Marie.

CNRS project call: INSU-INSMI

Project acronym: CORSURF

Project title: COmplex Rheology SURface Flows

Coordinator: Cindy Guichard

In collaboration with E. Fernández-Nieto (Sevilla, Spain).

Geophysical flows like avalanches (mud, snow) or landslides involve surface flows with non-Newtonian fluids. The goal is to develop numerical models, both accurate with respect to the material behavior and industrially efficient.

9.2.4. CNRS MOCHA (2016)

Participant: Martin Parisot.

CNRS project call: PEPS

Project acronym: MOCHA

Project title: Multi-dimensiOnal Coupling in Hydraulics and data Assimilation

Coordinator: Martin Parisot

Multi-dimensionnal coupling in river hydrodynamics offers a convenient solution to properly model complex flow while limiting the computational cost and making the most of pre-existing models. The project aims to adapt the lateral interface coupling proposed in [37] to the implicit version and test it on real data for the Garonne River.

9.2.5. CNRS Moset (2016-2017)

Participants: Emmanuel Audusse, Martin Parisot.

CNRS project call: INSU Tellus

Project acronym: Moset

Project title: Modélisation des suspensions concentrées naturelles

Coordinator: Emmanuel Audusse

In collaboration with G. Antoine (EDF), S. Boyaval (LHSV), C. Le Bouteiller (Irstea), M. Jodeau (EDF).

Gathering mathematicians (numerical analysis) and geophysicists, this project focuses on the quantitative prediction of solid transport. This issue raises several questions about rheology when the sediment concentration is high enough. It is crucial for modelling the dynamics of suspension. The collaboration aims at assessing models by means of experimental data and at providing preliminary numerical results to evaluate the order of magnitude of constraints.

9.2.6. Inria Project Lab "Algae in Silico" (2015-2018)

Participants: Nora Aïssiouene, Marie-Odile Bristeau, David Froger, Yohan Penel, Jacques Sainte-Marie, Fabien Souille.

In the aftermath of the ADT In@lgae (2013–2015), we developed a simulation tool for microalgae culture. An Inria Project Lab "Algae in Silico" has started in collaboration with Inria teams BIOCORE and DYLISS. It concerns microalgae culture for biofuel production and the aim is to provide an integrated platform for numerical simulation "from genes to industrial processes".

9.2.7. ANR MIMOSA (2014–2017)

Participants: Nora Aïssiouene, Marie-Odile Bristeau, Anne Mangeney, Bernard Di Martino, Jacques Sainte-Marie.

Program: ANR Défi 1 "Gestion sobre des ressources et adaptation au changement climatique"

Project acronym: MIMOSA

Project title: MIcroseism MOdeling and Seismic Applications

Coordinator: Eleonore Stutzmann (IPGP)

Seismic noise is recorded by broadband seismometers in the absence of earthquakes. It is generated by the atmosphere-ocean system with different mechanisms in the different frequency bands. Even though some mechanisms have been known for decades, an integrated understanding of the noise in the broadband period band 1-300sec is still missing. Using novel theoretical, numerical and signal processing methods, this project will provide a unified understanding of the noise sources and quantitative models for broadband noise. Conversely, we will be able to interpret seismic noise in terms of ocean wave properties. This first analysis step will lead to the identification and characterization of source events, which we will use to improve noise tomography, and seismic monitoring.

9.2.8. ANR LANDQUAKES (2012–2016)

Participant: Anne Mangeney.

Program: ANR Blanc "Mathématiques et interactions"

Project acronym: LANDQUAKES

Project title: Modélisation des glissements de terrain et des ondes sismiques générées pour détecter et comprendre les instabilités gravitaires

Coordinator: Anne Mangeney

Within the ANR domain "Mathematics and Interfaces", this ANR project (between Univ. Paris-Est – LAMA, Univ. Denis Diderot Paris 7 – IPGP, Univ. Nantes – LPGN, Univ. Strasbourg EOST, 180.000 euros) deals with the mathematical and numerical modelling of landslides and generated seismic waves.

A. Mangeney is also involved in the CARIB ANR program (2014–2017) entitled "Comprendre les processus de construction et de destruction des volcans de l'Arc des Petites Antilles".

9.2.9. GdR EGRIN (2013–2017)

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.

9.3. European Initiatives

9.3.1. ERC Consolidator Grant (2013-2018)

Participants: Anne Mangeney, Hugo Martin.

The project SLIDEQUAKES is about detection and understanding of landslides by observing and modelling gravitational flows and generated earthquakes and is funded by the European Research Council (2 million euros). More precisely, it deals with the mathematical, numerical and experimental modelling of gravitational flows and generated seismic waves coupled with field measurements to better understand and predict these natural hazards and their link with volcanic, seismic and climatic activities.

9.4. International Initiatives

9.4.1. Inria International Partners

9.4.1.1. Informal International Partners

The collaboration with IMUS (Institute of Mathematics of the university of Sevilla, Spain) was informaly launched in 2016 through several visits in Spain of members of ANGE and the writing of a paper. To go further, a submission was made to create an Inria Associate Team.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

In the framework of the collaboration with researchers at the university of Sevilla (Spain), Enrique Fernández-Nieto spent two weeks (weeks n. 13 and 41) at Inria. IPGP hosted several researchers who work with A. Mangeney: Pere Roig (PhD, Departamento de Geodinámica i Geofísica, University of Barcelona, Spain), Giulia Bossi (postdoc, ETH, Zürich), Andrea Wolter and Margherita Spreafico (permanent positions, ETH, Zürich).

9.5.2. Visits to International Teams

9.5.2.1. Research Stays Abroad

- Y. Penel spent three months (Jan.-Mar.) at the university of Sevilla (Spain) to collaborate with E. Fernández-Nieto.
- N. Aissiouene went to the university of Málaga for one month (Apr.) and was involved in the project EDANYA.
- C. Guichard, Y. Penel and J. Sainte-Marie were invited to the university of Sevilla for one week (week n. 42) to set up a forthcoming project.
- A. Mangeney went to Sevilla in November (week n. 47).

We also mention that M. Parisot spent one week (week n. 48) at the university of Toulouse (CERFACS).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

B. Di Martino, C. Guichard, A. Mangeney, Y. Penel and J. Sainte-Marie organised the workshop “Complex rheology for granular flows: challenges and deadlocks” that took place at IPGP on October 14th and that gathered 25 researchers.

M. Parisot organise's the Working group (GdT) of the Ange team.

Moreover, L. Boittin co-organise's the Junior Seminar at Inria–Paris.

10.1.2. Journal

10.1.2.1. Reviewer - Reviewing Activities

Member	Journal
E. Audusse	Advanced Water Resources, ESAIM:ProcS., SMAI Journal of Computational Mathematics
E. Godlewski	Computers and Fluids
C. Guichard	Numerische Mathematik, Mathematics and Computers in Simulation, Annali di Matematica Pura ed Applicata
B. Haspot	Analysis of PDE, JDE, JMAA, M3AS, ARMA, Siam Analysis, Mathematische Annalen, CMS, JFA, Acta Applicandae Mathematicae
A. Mangeney	J. Fluid Mech., J. Geophys. Res., Numeric. Analyt. Methods Geomech., Earth Surf. Processes and Landforms
M. Parisot	Continuum Mechanics and Thermodynamics, Hydrological Processes
Y. Penel	Hydrological Processes
J. Salomon	SIAM Journal on Numerical Analysis, M2AN, SIAM Journal of Scientific Computing
J. Sainte-Marie	M2AN, Nonlinearity, IJNMF, Applied mathematical modellling, Journal of Hydraulic Research, Computers and Fluids, European Journal of Applied Math, Numerische Mathematik

10.1.3. Invited Talks

Conference	Location	Month	Members involved
European Geosciences Union	Vienna (Austria)	April	A. Mangeney
Peril flood	Paris	May	M. Parisot
CANUM	Obernai	May	D. Kazerani
ECMI	Santiago de Compostela (Spain)	June	C. Guichard
CECAM	Lyon	June	A. Mangeney
Martian gullies and Earth analogues	London (UK)	June	A. Mangeney
Scientific computing and modelling	Amiens	June	J. Sainte-Marie
Asymptotic Behavior of systems of PDE	Lille	June	D. Kazerani
AIMS	Orlando (US)	July	N. Aissiouene
EGU FORM-OSE training school	Azores (Spain)	July	A. Mangeney
HYP	Aachen (Germany)	August	Y. Penel
Modelling of coastal hydrodynamics	Vannes	September	E. Audusse, J. Sainte-Marie
GDR Renewable marine energies	Nantes	October	J. Salomon
GDR Transnat	Roscoff	November	E. Audusse
Liquid-vapor interfaces in fluid flows	Paris	December	D. Kazerani

Seminars	Date	Member
IMUS (Sevilla, Spain)	February	Y. Penel
EDANYA (Málaga, Spain)	April	N. Aissiouene
Paris (LJLL)	May	M. Parisot
GIS HED[2]	June	D. Kazerani
Luminy (CEMRACS)	August	E. Audusse
Lille	September	M. Parisot

10.1.4. Leadership within the Scientific Community

Organisation	People	Duty
AMIES	E. Godlewski	Member of board
CFEM	E. Godlewski	Director
Comité d'Orientation Pour les Risques Naturels Majeurs du Ministère de l'Environnement	A. Mangeney	Member
EGRIN	E. Audusse B. di Martino N. Goutal C. Guichard A. Mangeney M. Parisot J. Sainte-Marie	Correspondent (Paris 13) Correspondent (Corse) Correspondent (EDF) Correspondent (UPMC) Member of board Correspondent (ANGE) Scientific head
HCERES	A. Mangeney	Expert
LJLL	E. Godlewski	Deputy director
Research commission of UPMC	D. Kazerani	Member
SMAI	Y. Penel	Member of board

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master's degree (M2) E. Godlewski and J. Sainte-Marie, Hyperbolic models for complex flows and energy applications, 25 hours (lectures), Univ. Pierre et Marie Curie Paris 6

Master's degree (M2) E. Godlewski, Numerical methods for nonstationary PDEs, 18 hours (example and programming classes), Univ. Pierre et Marie Curie Paris 6

Master's degree (M2) C. Guichard, Numerical methods for nonstationary PDEs, 6 hours (programming classes), Univ. Pierre et Marie Curie Paris 6

Master's degree (M2) H. Martin and J. Sainte-Marie, Numerical methods in fluid mechanics, 52 hours (lectures and programming classes), Univ. Paris Diderot Paris 7, IPGP

Master's degree (M2) B. Haspot, Linear partial differential equations, 15 hours (lectures), Univ. Paris Dauphine

Master's degree (M2) B. Di Martino, Mathematical modelling, 21 hours (lectures, example and programming classes), Univ. Corse

Master's degree (M2) J. Salomon, Scientific computing and numerical analysis, 30 hours (lectures, example and programming classes), Univ. Paris Dauphine

Engineering school (2nd year) E. Audusse, Hyperbolic systems, 30 hours (lectures and example classes), Univ. Paris 13

Engineering school (2nd year) E. Audusse, Finite difference method for PDEs, 30 hours (lectures and programming classes), Univ. Paris 13

Master's degree (M1) C. Guichard, Basis of numerical methods, 68 hours (programming classes), Univ. Pierre et Marie Curie Paris 6

Master's degree (M1) H. Martin and J. Sainte-Marie, Models in geosciences, 40 hours (lectures and programming classes), Univ. Paris Diderot Paris 7, IPGP

Master's degree (M1) B. Di Martino, Numerical methods, 27 hours (lectures and programming classes), Univ. Corse

Master's degree (M1) B. Di Martino, Mathematics, 33 hours (lectures, example and programming classes), Univ. Corse

Master's degree (M1) B. Haspot, Functional analysis, 85 hours (lectures and example classes), Univ. Paris Dauphine

Master's degree (M1) E. Godlewski, Numerical methods, 24 hours (lectures), Univ. Pierre et Marie Curie Paris 6

Engineering school (1st year) E. Audusse, Numerical analysis for differential equations, 30 hours (example classes), Univ. Paris 13

Bachelor's degree (L3) C. Guichard, Numerical linear algebra, 76.5 hours (example and programming classes), Univ. Pierre et Marie Curie Paris 6

Bachelor's degree (L3) M. Parisot, Hilbert analysis, 30 hours (lectures and example classes), Univ. Pierre et Marie Curie Paris 6

Bachelor's degree (L3) B. Di Martino, Numerical methods, 18 hours (lectures and programming classes), Univ. Corse

Bachelor's degree (L2) E. Audusse, Scientific computing, 36 hours (lectures), Univ. Paris 13

Bachelor's degree (L2) Y. Penel, Integration in 2 and 3 dimensions, 12 hours (lectures), Univ. Pierre et Marie Curie Paris 6

Bachelor's degree (L1) F. Wahl, Linear algebra, 54 hours (example classes), Univ. Pierre et Marie Curie Paris 6

Bachelor's degree (L1) F. Wahl, Analysis and algebra for sciences, 38.5 hours (example classes), Univ. Pierre et Marie Curie Paris 6

Bachelor's degree (L1] E. Nayir, Mathematics, 72 hours (example classes), Univ. Pierre et Marie Curie Paris 6

Bachelor's degree (L1) E. Godlewski and E. Nayir, Linear algebra, 48 hours (lectures and example classes), Univ. Pierre et Marie Curie Paris 6

Bachelor's degree (L1) L. Boittin, Calculus, 28 hours (example classes), Univ. Pierre et Marie Curie Paris 6

Some members are responsible of educational pathways:

- E. Audusse is the deputy director of the “Applied Mathematics and Scientific Computing” program of the SupGalilee engineering school.
- E. Godlewski is the head of the “Mathematics for Industry” M.Sc. program of Univ. Pierre et Marie Curie Paris 6.
- C. Guichard is the associated head of the “Mathematics and Programming” B. program of Univ. Pierre et Marie Curie Paris 6.
- A. Mangeney is the head of the “Telluric Natural Hazards” M. Sc. specialty of IPGP.

10.2.2. Supervision

PostDoc in progress El Hadji Kone, *Numerical modelling of shallow two-phase flows*, Institut de Physique du Globe (Univ. Paris 7), supervised by A. Mangeney, from 2016

PostDoc in progress Pierre-Olivier Lamarre, *Optimization of the hydrodynamic regime in a raceway and lagrangian trajectories of algae*, supervised by J. Sainte-Marie and N. Aïssiouene (in collaboration with O. Bernard, BIOCORE)

PhD Nora Aïssiouene, *Derivation and analysis of a non-hydrostatic Shallow water type model*, Univ. Pierre et Marie Curie Paris 6 (Inria grant), supervised by E. Godlewski and J. Sainte-Marie, defended on Dec. 16

PhD Dena Kazerani, *Simulation et modélisation de problèmes à frontière libre*, Univ. Pierre et Marie Curie Paris 6, supervised by N. Seguin (in collaboration with P. Frey and C. Audiard), defended on Nov. 16

PhD Clément Mifsud, *Analyse et approximation des systèmes de Friedrichs : application à la modélisation de l'élastoplasticité*, Univ. Pierre et Marie Curie Paris 6, supervised by N. Seguin (in collaboration with J.-F. Babadjian and B. Després), defended on Nov. 16

PhD Amandine Sergeant-Boy, *Detection and characterisation of seismic sources generated by glaciers: numerical modelling and analysis of seismic waves*, Institut de Physique du Globe (Univ. Paris 7), supervised by A. Mangeney (in collaboration with J.-P. Montagner, E. Stutzmann and O. Castelnau), defended on Nov. 16

PhD in progress Léa Boittin, *Modelling, analysis and efficient numerical resolution for erosion processes*, Univ. Pierre et Marie Curie Paris 6 (Inria grant), supervised by E. Audusse, M. Parisot and J. Sainte-Marie, from Jan. 16

PhD in progress Vincent Bachelet, *Granular flows and generated acoustic waves: a laboratory investigation*, Institut de Physique du Globe (Univ. Paris 7), supervised by A. Mangeney (in collaboration with J. De Rosny and R. Toussaint), from 2015

PhD in progress Do Minh Hieu, *Analyse mathématique et schémas volumes finis pour la simulation des écoulements quasi-géostrophiques à bas nombre de Froude*, Univ. Paris 13, supervised by E. Audusse and Y. Penel (in collaboration with S. Dellacherie and P. Omnes), from 2014

PhD in progress Virginie Durand, *Spatio-temporal measurement, analysis and simulation of gravitational flows and seismic signals*, Institut de Physique du Globe (Univ. Paris 7), supervised by A. Mangeney, from 2016

PhD in progress Julian Kühnert, *Simulation of high frequency seismic waves*, Institut de Physique du Globe (Univ. Paris 7), supervised by A. Mangeney, from 2016

PhD in progress Hugo Martin, *Simulation of the coupling between seismic waves and granular flows*, Institut de Physique du Globe (Univ. Paris 7), supervised by A. Mangeney (in collaboration with Y. Maday), from 2016

PhD in progress Hélène Miallot, *Numerical modelling of landquakes*, Institut de Physique du Globe (Univ. Paris 7), supervised by A. Mangeney (in collaboration with Y. Capdeville), from 2015

PhD in progress Ethem Nayir, *Approximation multi-vitesse des équations de Navier-Stokes hydrostatiques: Analyse mathématique et simulations numériques*, Univ. Pierre et Marie Curie Paris 6, supervised by E. Audusse, Y. Penel and J. Sainte-Marie, from 2014

PhD in progress Nourelhouda Omrane, *Mathematical analysis and control of free-surface flows in variable domains*, Univ. Corse, supervised by B. Di Martino, from 2016

PhD in progress Pierrick Quémar, *3D numerical simulations of environmental hydrolics: application to Telemac*, Univ. Paris 13, supervised by E. Audusse and N. Goutal (in collaboration with A. Decoene, O. Lafitte, A. Leroy and C. Tuân Phan), from 2016

PhD in progress Sebastian Reyes-Riffo, *Mathematical methods for recovering marine energies*, Univ. Paris Dauphine, supervised by J. Salomon, from 2016

PhD in progress Fabien Wahl, *Modelling and analysis of interactions between free surface flows and floating objects*, Univ. Pierre et Marie Curie Paris 6, supervised by C. Guichard, E. Godlewski, M. Parisot and J. Sainte-Marie, from 2015

M2 internship Marie Zehgdoudi, *Étude des flux de lave émis au Piton de la Fournaise à partir d'enregistrements sismiques*, IPGP, supervised by A. Mangeney, Summer 2016

L3 internship Albin Grataloup, *Perturbation theory applied to the shallow water equations*, Ecole Normale Supérieure de Lyon, supervised by E. Audusse and Y. Penel, Summer 2016

We also mention that N. Goutal and M. Parisot supervised a student during the 2016 session of CEMRACS during the summer.

10.2.3. Juries

Feb., PhD A. Mangeney (president): Antoine Frère (CEA, *Modélisation des tsunamis génér{es par écoulement gravitaire : application dans le golfe de Gascogne*)

Feb., PhD A. Mangeney (referee): Véronique Dansereau (Univ. Grenoble, *A Maxwell-Elasto-Brittle model for the drift of ice*)

July, PhD E. Godlewski: Khalil Haddaoui (Univ. Pierre et Marie Curie Paris 6, *Méthodes numériques de haute précision et calcul scientifique pour le couplage de modèles hyperboliques*)

Sept., PhD E. Godlewski (referee): Asma Toumi (Univ. Toulouse, *Méthode numérique asynchrone pour la modélisation de phénomènes multi-échelles*)

Oct., PhD E. Godlewski (referee): Hippolyte Lochon (Univ. Marseille, *Modélisation et simulation d'écoulements transitoires eau-vapeur en approche bi-fluide*)

Oct., PhD B. Di Martino (referee): Ralph Lteif (Univ. Chambéry, *Modélisation et analyse mathématique de modèles en océanographie*)

Nov., PhD J. Sainte-Marie (referee): Amina Nouhou-Bako (Univ. Orléans, *Modélisation numérique de l'érosion diffuse des sols. Interaction gouttes-ruissellement*)

Nov., PhD E. Godlewski and J. Sainte-Marie: Dena Kazerani (Univ. Pierre et Marie Curie Paris 6, *Études mathématiques de fluides à frontières libres en dynamique incompressible*)

Nov., PhD A. Mangeney: Amandine Sergeant-Boy (Univ. Paris 7, *Detection and characterisation of seismic sources generated by glaciers: numerical modelling and analysis of seismic waves*)

Nov., PhD A. Mangeney: Sébastien Lherminier (Univ. Lyon 1, *Dynamique des avalanches invariantes d'échelle*)

Dec., PhD E. Godlewski, A. Mangeney and J. Sainte-Marie: Nora Aïssiouene (Univ. Pierre et Marie Curie Paris 6, *A numerical method for a dispersive shallow water system*)

Dec., HdR J. Sainte-Marie: Carine Lucas (Univ. Orléans, *Analyse mathématique et numérique de quelques modèles d'érosion*)

M. Parisot participated to the intermediate evaluation of Cécile Taing's PhD thesis (Univ. Pierre et Marie Curie Paris 6). A. Mangeney was a member of the jury evaluating M2 internships at IPGP.

10.3. Popularization

March E. Audusse intervened in a middle school on the occasion of the French week of Maths. J. Salomon went to a high school at Limay.

May N. Aïssiouene, E. Audusse, E. Godlewski (organiser), Y. Penel and F. Wahl ran a stand on the occasion of the "salon de la culture et des jeux mathématiques".

July M. Parisot gave a vulgarisation talk at Inria ("demi-heure de la science").

October E. Audusse intervened in a high school on the occasion of Savantes Banlieues.

November E. Nayir, Y. Penel and F. Wahl ran a stand during the ONISEP exhibition.

December E. Godlewski managed a group of middle school students at the Jacques-Louis Lions lab.

December F. Wahl helped the organisation at the "Math. Employment" show.

A. Sergeant-Boy, A. Mangeney, E. Stutzmann, J.-P. Montagner, F. Walter, L. Moretti, and O. Castelnau wrote an article entitled "La sismologie pour ausculter les pertes des glaciers des calottes polaires, lors du vêlage d'icebergs" in the CNRS–INSU newspaper (Apr.).

A. Mangeney is coaching high-school students from disadvantaged areas to manage scientific projects.

11. Bibliography

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, B. DI MARTINO, C. GUICHARD, J. SAINTE-MARIE.*Layer-averaged Euler and Navier-Stokes equations*, Sep 2015, working paper or preprint, https://hal.inria.fr/hal-01202042.

[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° 3, p. 459-490, http://dx.doi.org/10.1142/S0218202511005118.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[5] N. AISSIOUENE.*Numerical analysis and discrete approximation of a dispersive shallow water model*, Pierre et Marie Curie, Paris VI, December 2016, https://hal.archives-ouvertes.fr/tel-01418676.

Articles in International Peer-Reviewed Journal

[6] N. AISSIOUENE, M.-O. BRISTEAU, E. GODLEWSKI, J. SAINTE-MARIE.*A combined finite volume - finite element scheme for a dispersive shallow water system*, in "Networks and Heterogeneous Media (NHM)", January 2016, vol. 11, n° 1, p. 1-27, https://hal.inria.fr/hal-01160718.

[7] E. AUDUSSE, F. BOUCHUT, M.-O. BRISTEAU, J. SAINTE-MARIE.*Kinetic entropy inequality and hydrostatic reconstruction scheme for the Saint-Venant system*, in "Mathematics of Computation", November 2016, vol. 85, p. 2815-2837 [*DOI :* 10.1090/MCOM/3099], https://hal.inria.fr/hal-01063577.

[8] J.-F. BABADJIAN, C. MIFSUD, N. SEGUIN.*Relaxation approximation of Friedrich's systems under convex constraints*, in "Networks and Heterogeneous Media", 2016, vol. 11, n° 2, p. 223 - 237 [*DOI :* 10.3934/NHM.2016.11.223], https://hal.archives-ouvertes.fr/hal-01157484.

[9] F. BOUCHUT, I. R. IONESCU, A. MANGENEY.*An analytic approach for the evolution of the static/flowing interface in viscoplastic granular flows*, in "Communications in Mathematical Sciences", 2016, vol. 14, n° 8, p. 2101-2126 [*DOI :* 10.4310/CMS.2016.V14.N8.A2], https://hal-upec-upem.archives-ouvertes.fr/hal-01081213.

[10] C. CANCÈS, F. COQUEL, E. GODLEWSKI, H. MATHIS, N. SEGUIN.*Error analysis of a dynamic model adaptation procedure for nonlinear hyperbolic equations*, in "Communications in Mathematical Sciences", 2016, vol. 14, n° 1, p. 1-30, https://hal.archives-ouvertes.fr/hal-00852101.

[11] D. KAZERANI.*Global existence for small data of the viscous Green-Naghdi type equations*, in "Journal of Differential Equations", July 2016, vol. 261, n$^{\text{o}}$ 1, p. 762-796 [*DOI :* 10.1016/J.JDE.2016.03.022], https://hal.archives-ouvertes.fr/hal-01111941.

[12] D. KAZERANI.*The symmetric structure of the Green-Naghdi type equations*, in "Communications in Mathematical Sciences", August 2016, vol. 14, n$^{\text{o}}$ 7, p. 1925-1946 [*DOI :* 10.4310/CMS.2016.V14.N7.A7], https://hal.archives-ouvertes.fr/hal-01074488.

[13] M. LACHOWICZ, H. LESZCZYŃSKI, M. PARISOT.*A simple kinetic equation of swarm formation: blow–up and global existence*, in "Applied Mathematics Letters", January 2016 [*DOI :* 10.1016/J.AML.2016.01.008], https://hal.inria.fr/hal-01241998.

[14] M. LACHOWICZ, M. PARISOT, Z. SZYMAŃSKA.*Intracellular protein dynamics as a mathematical problem*, in "Discrete and Continuous Dynamical Systems - Series B (DCDS-B)", February 2016, https://hal.inria.fr/hal-01152399.

[15] C. MIFSUD, B. DESPRÉS, N. SEGUIN.*Dissipative formulation of initial boundary value problems for Friedrichs' systems*, in "Communications in Partial Differential Equations", January 2016, vol. 41, n$^{\text{o}}$ 1 [*DOI :* 10.1080/03605302.2015.1103750], https://hal.archives-ouvertes.fr/hal-01074542.

[16] M. PARISOT, M. LACHOWICZ.*A Kinetic Model for the formation of Swarms with nonlinear interactions*, in "Kinetic and Related Models ", March 2016, vol. 9, n$^{\text{o}}$ 1, 33 [*DOI :* 10.3934/KRM.2016.9.131], https://hal.inria.fr/hal-01152397.

[17] M. PARISOT, J.-P. VILA.*Centered-potential regularization for the advection upstream splitting method*, in "SIAM Journal on Numerical Analysis", 2016, https://hal.inria.fr/hal-01152395.

Other Publications

[18] N. AISSIOUENE, T. AMTOUT, M. BRACHET, E. FRÉNOD, R. HILD, C. PRUD 'HOMME, A. ROUSSEAU, S. SALMON.*Hydromorpho: A coupled model for unsteady Stokes/Exner equations and numerical results with Feel++ library*, February 2016, working paper or preprint, https://hal.inria.fr/hal-01266223.

[19] E. AUDUSSE, C. CHALONS, P. UNG.*A simple three-wave Approximate Riemann Solver for the Saint-Venant–Exner equations*, August 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01204754.

[20] M.-O. BRISTEAU, D. FROGER, R. HAMOUDA, A. MANGENEY, J. SAINTE-MARIE, M. VALLÉE.*Numerical simulations of 3d hydrostatic Navier-Stokes system with free surface*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01393147.

[21] C. CANCÈS, C. GUICHARD.*Numerical analysis of a robust free energy diminishing Finite Volume scheme for parabolic equations with gradient structure*, 2016, to appear in Foundations of Computational Mathematics, https://hal.archives-ouvertes.fr/hal-01119735.

[22] B. DI MARTINO, B. HASPOT, Y. PENEL.*Global stability of weak solutions for a multilayer Saint-Venant model with interactions between the layers*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01407886.

[23] J. DRONIOU, R. EYMARD, T. GALLOUËT, C. GUICHARD, R. HERBIN.*The gradient discretisation method : A framework for the discretisation and numerical analysis of linear and nonlinear elliptic and parabolic problems*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01382358.

[24] R. EYMARD, P. FERON, C. GUICHARD.*Family of convergent numerical schemes for the incompressible Navier-Stokes equations*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01382924.

[25] E. D. FERNANDEZ-NIETO, M. PARISOT, Y. PENEL, J. SAINTE-MARIE.*Layer-averaged approximations for inviscid flow models*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01324012.

[26] E. GODLEWSKI, M. PARISOT, J. SAINTE-MARIE, F. WAHL.*Congested shallow water type model : roof modelling in free surface flow*, September 2016, working paper or preprint, https://hal.inria.fr/hal-01368075.

[27] M. LACHOWICZ, H. LESZCZYŃSKI, M. PARISOT.*Blow-up and global existence for a kinetic equation of swarm formation*, 2016, working paper or preprint, https://hal.inria.fr/hal-01370006.

[28] C. LUSSO, F. BOUCHUT, A. ERN, A. MANGENEY.*A free interface model for static/flowing dynamics in thin-layer flows of granular materials with yield: simple shear simulations and comparison to experiments*, December 2016, working paper or preprint, https://hal-upec-upem.archives-ouvertes.fr/hal-00992309.

[29] C. LUSSO, F. BOUCHUT, A. ERN, A. MANGENEY.*Explicit solutions to a free interface model for the static/flowing transition in thin granular flows*, August 2016, working paper or preprint, https://hal-upec-upem.archives-ouvertes.fr/hal-01180686.

[30] C. LUSSO, A. ERN, F. BOUCHUT, A. MANGENEY, M. FARIN, O. ROCHE.*Two-dimensional simulation by regularization of free surface viscoplastic flows with Drucker-Prager yield stress and application to granular collapse*, December 2016, working paper or preprint [*DOI :* 10.1016/J.JCP.2016.12.036], https://hal-upec-upem.archives-ouvertes.fr/hal-01133786.

References in notes

[31] 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.

[32] F. BOUCHUT, J. SOMMER, V. ZEITLIN.*Frontal geostrophic adjustment and nonlinear wave phenomena in one-dimensional rotating shallow water. part 2. high-resolution numerical simulations*, in "J. Fluid Mech.", 2004, vol. 514, p. 35–63.

[33] 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.

[34] 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.

[35] S. DELLACHERIE.*Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number*, in "J. Comput. Phys.", 2010, vol. 229, n^{o} 4, p. 978–1016.

[36] B. DESPRÉS, C. BUET.*The structure of well-balanced schemes for Friedrichs systems with linear relaxation*, in "Appl. Math. Comput.", 2016, vol. 272, p. 440–459.

[37] N. GOUTAL, M. PARISOT, F. ZAOUI.*A 2D reconstruction for the transverse coupling of shallow water models*, in "Int. J. Numer. Methods Fluids", 2014, vol. 75, n° 11, p. 775–799.

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:

2. - Software
2.1. - Programming Languages
2.1.1. - Semantics of programming languages
2.2.1. - Static analysis
2.3.1. - Embedded systems
2.4. - Verification, reliability, certification
2.4.1. - Analysis
2.4.2. - Model-checking
2.4.3. - Proofs
4.4. - Security of equipment and software
4.5. - Formal methods for security

Other Research Topics and Application Domains:

1.1. - Biology
1.1.10. - Mathematical biology
1.1.11. - Systems biology
5.2. - Design and manufacturing
5.2.1. - Road vehicles
5.2.2. - Railway
5.2.3. - Aviation
5.2.4. - Aerospace
6.1. - Software industry
6.1.1. - Software engineering
6.1.2. - Software evolution, maintenance
6.6. - Embedded systems

1. Members

Research Scientists

Xavier Rival [Team leader, Inria, Senior Researcher, HDR]
Patrick Cousot [ENS Paris, Professor Emeritus, HDR]
Vincent Danos [CNRS, Senior Researcher, HDR]
Cezara Dragoi [Inria, Researcher]
Jérôme Feret [Inria, Researcher]

Technical Staff

Francois Berenger [Inria]
Tie Cheng [Inria, until May 2016]
Kim Quyen Ly [Inria]

PhD Students

Mehdi Bouaziz [Inria, until Nov 2016]
Ferdinanda Camporesi [Inria]
Huisong Li [Inria]
Jiangchao Liu [Inria]
Hugo Illous [CEA, ENS]
Thibault Suzanne [ENS Paris]
Andreea Beica [ENS Paris]
Patric Fulop [University of Edinburgh and ENS Paris]

Post-Doctoral Fellows
Ilias Garnier [ENS Paris and University of Edinburgh]
Nicolas Behr [ENS Paris and University of Edinburgh]
Ricardo Honorato-Zimmer [ENS Paris and University of Edinburgh]

Visiting Scientist
Kwangkeun Yi [Seoul National University, until Oct 2016]

Administrative Assistant
Nathalie Gaudechoux [Inria]

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 became 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), and medical systems (pacemakers, surgery and patient monitoring systems) rely on complex software, which should be verified.

Beyond the field of embedded systems, other pieces of software may cause very significant harm in 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 be 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 expressing 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* [13]. 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* [13], which over-approximates 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.

Yet, the same principles can also be applied successfully to other goals. In particular, the abstract interpretation framework should be viewed 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 [12], 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. The analysis of biological models

One of our application domains, the analysis of biological models, is not a classical target of static analysis because it aims at analyzing models instead of programs. Yet, the analysis of biological models is closely intertwined with the other application fields of our group. Firstly, abstract interpretation provides a formal understanding of the abstraction process which is inherent to the modeling process. Abstract interpretation is also used to better understand the systematic approaches which are used in the systems biology field to capture the properties of models, until getting formal, fully automatic, and scalable methods. Secondly, abstract interpretation is used to offer various semantics with different grains of abstraction, and, thus, new methods to apprehend the overall behavior of the models. Conversely, some of the methods and abstractions which are developed for biological models are inspired by the analysis of concurrent systems and by security analysis. Lastly, the analysis of biological models raises issues about differential systems, stochastic systems, and hybrid systems. Any breakthrough in these directions will likely be very important to address the important challenge of the certification of critical systems in interaction with their physical environment.

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 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, monitoring typically rely on a *parallel* structure, where several threads are executed in parallel, 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. Biological systems

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 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 abstract interpretation-based analyses, 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 so as to identify the key elements in the dynamics of these models. The results of these analyses are sent to another tool, which is used to automatically simplify the 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.

5. Highlights of the Year

5.1. Highlights of the Year

The team obtained several strong results published in excellent international conferences, with high theoretical and applied impact(see detailed results). Among the theoretical results we underline those presented in conferences like Principles of programming languages POPL 2016, and among the applied results we underline the release of MemCad, the first analyzer that can handle the analysis of various data structures.

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

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 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: Patrick Cousot, Radhia Cousot, Jérôme Feret, Laurent Mauborgne, Antoine Miné and Xavier Rival
- Partner: CNRS
- Contact: Patrick Cousot
- URL: http://www.astree.ens.fr/

6.3. AstréeA

The AstréeA Static Analyzer of Asynchronous Software

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: Patrick Cousot, Radhia Cousot, Jérôme Feret, Antoine Miné and Xavier Rival est toujours membre de Inria. logiciels Inria): https://bil.inria.fr/
- Contact: Patrick Cousot
- URL: http://www.astreea.ens.fr/

6.4. ClangML

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: François Berenger, Pippijn Van Steenhoven and Devin Mccoughlin toujours membre de Inria. Inria): https://bil.inria.fr/
- Contact: François Berenger
- 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: Caterina Urban and Antoine Miné
- Contact: Caterina Urban
- URL: http://www.di.ens.fr/~urban/FuncTion.html

6.6. MemCAD

The MemCAD static analyzer
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 350 small size test cases that are used as regression tests.

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

6.7. OPENKAPPA

La platte-forme de modélisation OpenKappa
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.

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

6.8. 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

6.9. Translation Validation

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 guara 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
- Contact: Xavier Rival

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

6.11. CELIA

The MemCAD static analyzer

FUNCTIONAL DESCRIPTION

CELIA is a tool for the static analysis and verification of C programs manipulating dynamic lists. The static analyzer computes for each control point of a C program the assertions which are true (i.e., invariant) at this control point. The specification language is a combination of Separation Logic with a first order logic over sequences of integers. The inferred properties describe the shape of the lists, their size, the relations between the data (or the sum, or the multiset of data) in list cells. The analysis is inter-procedural, i.e., the assertions computed relate the procedure local heap on entry to the corresponding local heap on exit of the procedure. The results of the analysis can provide insights about equivalence of procedures on lists or null pointer dereferencing. The analysis is currently extended to programs manipulating concurrent data structures.

- Participants: Ahmed Bouajjani, Cezara Drăgoi, Constantin Enea, Mihaela Sighireanu
- Contact: Cezara Drăgoi
- URL: http://www.liafa.jussieu.fr/celia/

6.12. DAFT

DAFT

FUNCTIONAL DESCRIPTION

DAFT is a distributed file management system in user-space, with a command-line interface. DAFT is intended at computational scientists, involved in data-intensive, distributed experiments and when no distributed file-system is available on computing nodes. DAFT is secure; all messages are cryptographically signed and encrypted by default.

- Participants: Francois Berenger and Camille Coti.
- Contact: Francois Berenger and Camille Coti.
- URL: https://github.com/UnixJunkie/daft.

7. New Results

7.1. Memory Abstraction

7.1.1. Abstraction of arrays based on non contiguous partitions

Participants: Jiangchao Liu, Xavier Rival [correspondant].

Abstract interpretation, Memory abstraction, Array abstract domains. In [2], we studied array abstractions.

Array partitioning analyses split arrays into contiguous partitions to infer properties of cell sets. Such analyses cannot group together non contiguous cells, even when they have similar properties. We proposed an abstract domain which utilizes semantic properties to split array cells into groups. Cells with similar properties will be packed into groups and abstracted together. Additionally, groups are not necessarily contiguous. This abstract domain allows to infer complex array invariants in a fully automatic way. Experiments on examples from the Minix 1.1 memory management demonstrated its effectiveness.

7.2. Rule-based modeling

7.2.1. Reachability analysis via orthogonal sets of patterns

Participants: Kim Quyên Ly, Jérôme Feret [correspondant].

Rule-based modeling languages, as Kappa, allow for the description of very detailed mechanistic models. Yet, as the rules become more and more numerous, there is a need for formal methods to enhance the level of confidence in the models that are described with these languages. We develop abstract interpretation tools to capture invariants about the biochemical structure of bio-molecular species that may occur in a given model. In previous works, we have focused on the relationships between the states of the sites that belong to a same instance of a protein. This comes down to detect for a specific set of patterns, which ones may be reachable during the execution of the model. This paper [6], we generalize this approach to a broader family of abstract domains, that we call orthogonal sets of patterns. More precisely, an orthogonal set of patterns is obtained by refining recursively the information about some patterns containing a given protein, so as to partition of the set of occurrences of this protein in any mixture.

7.2.2. *Local traces: an over-approximation of the behaviour of the proteins in rule-based models*

Participants: Kim Quyên Ly, Jérôme Feret [correspondant].

Thanks to rule-based modelling languages, we can assemble large sets of mechanistic protein-protein interactions within integrated models. Our goal would be to understand how the behaviour of these systems emerges from these low-level interactions. Yet this is a quite long term challenge and it is desirable to offer intermediary levels of abstraction, so as to get a better understanding of the models and to increase our confidence within our mechanistic assumptions. In this paper [5], we propose an abstract interpretation of the behaviour of each protein, in isolation. Given a model written in Kappa, this abstraction computes for each kind of protein a transition system that describes which conformations this protein can take and how a protein can pass from one conformation to another one. Then, we use simplicial complexes to abstract away the interleaving order of the transformations between conformations that commute. As a result, we get a compact summary of the potential behaviour of each protein of the model.

7.3. Formal Derivation of Qualitative Dynamical Models from Biochemical Networks

Participants: Wassim Abou-Jaoudé, Denis Thieffry, Jérôme Feret [correspondant].

As technological advances allow a better identification of cellular networks, more and more molecular data are produced allowing the construction of detailed molecular interaction maps. One strategy to get insights into the dynamical properties of such systems is to derive compact dynamical models from these maps, in order to ease the analysis of their dynamics. Starting from a case study, we present in [1] a methodology for the derivation of qualitative dynamical models from biochemical networks. Properties are formalised using abstract interpretation. We first abstract states and traces by quotienting the number of instances of chemical species by intervals. Since this abstraction is too coarse to reproduce the properties of interest, we refine it by introducing additional constraints. The resulting abstraction is able to identify the dynamical properties of interest in our case study.

7.4. Taking Static Analysis to the Next Level: Proving the Absence of Run-Time Errors and Data Races with Astrée

Participants: Antoine Miné, Laurent Mauborgne, Xavier Rival, Jérôme Feret [correspondant], Patrick Cousot, Daniel Kästner, Stephan Wilhelm, Christian Ferdinand.

In [9], we present an extension of Astrée to concurrent C software. Astrée is a sound static analyzer for run-time errors previously limited to sequential C software. Our extension employs a scalable abstraction which covers all possible thread interleavings, and soundly reports all run-time errors and data races: when the analyzer does not report any alarm, the program is proven free from those classes of errors. We show how this extension is able to support a variety of operating systems (such as POSIX threads, ARINC 653, OSEK/AUTOSAR) and report on experimental results obtained on concurrent software from different domains, including large industrial software.

7.5. Stochastic mechanics of graph rewriting

Participants: Nicolas Behr, Vincent Danos, Ilias Garnier [correspondant].

We propose an algebraic approach to stochastic graph-rewriting which extends the classical construction of the Heisenberg-Weyl algebra and its canonical representation on the Fock space. Rules are seen as particular elements of an algebra of “diagrams": the diagram algebra D. Diagrams can be thought of as formal computational traces represented in partial time. They can be evaluated to normal diagrams (each corresponding to a rule) and generate an associative unital non-commutative algebra of rules: the rule algebra R. Evaluation becomes a morphism of unital associative algebras which maps general diagrams in D to normal ones in R. In this algebraic reformulation, usual distinctions between graph observables (real-valued maps on the set of graphs defined by counting subgraphs) and rules disappear. Instead, natural algebraic substructures of R arise: formal observables are seen as rules with equal left and right hand sides and form a commutative subalgebra, the ones counting subgraphs forming a sub-subalgebra of identity rules. Actual graph-rewriting is recovered as a canonical representation of the rule algebra as linear operators over the vector space generated by (isomorphism classes of) finite graphs. The construction of the representation is in close analogy with and subsumes the classical (multi-type bosonic) Fock space representation of the Heisenberg-Weyl algebra.

This shift of point of view, away from its canonical representation to the rule algebra itself, has unexpected consequences. We find that natural variants of the evaluation morphism map give rise to concepts of graph transformations hitherto not considered. These will be described in a separate paper [2]. In this extended abstract we limit ourselves to the simplest concept of double-pushout rewriting (DPO). We establish “jump-closure", i.e. that the sub-space of representations of formal graph observables is closed under the action of any rule set. It follows that for any rule set, one can derive a formal and self-consistent Kolmogorov backward equation for (representations of) formal observables.

This result and the fallowing ones, co-authored by Vincent Danos, were published in peer-reviewed international conferences and journals. Although the papers are on HAL, they are not imported in the bibtex file so we can’t cite them properly.

7.6. Giry and the machine

Participants: Fredrik Dahlqvist, Vincent Danos, Ilias Garnier [correspondant].

We present a general method – the Machine – to analyse and characterise in finitary terms natural transformations between well-known functors in the category Pol of Polish spaces. The method relies on a detailed analysis of the structure of Pol and a small set of categorical conditions on the domain and codomain functors. We apply the Machine to transformations from the Giry and positive measures functors to combinations of the Vietoris, multiset, Giry and positive measures functors. The multiset functor is shown to be defined in Pol and its properties established. We also show that for some combinations of these functors, there cannot exist more than one natural transformation between the functors, in particular the Giry monad has no natural transformations to itself apart from the identity. Finally we show how the Dirichlet and Poisson processes can be constructed with the Machine.

7.7. Robustly Parameterised Higher-Order Probabilistic Models

Participants: Fredrik Dahlqvist, Vincent Danos, Ilias Garnier [correspondant].

We present a method for constructing robustly parameterised families of higher-order probabilistic models. Parameter spaces and models are represented by certain classes of functors in the category of Polish spaces. Maps from parameter spaces to models (parameterisations) are continuous and natural transformations between such functors. Naturality ensures that parameterised models are invariant by change of granularity – i.e. that parameterisations are intrinsic. Continuity ensures that models are robust with respect to their parameterisation. Our method allows one to build models from a set of basic functors among which the Giry probabilistic functor, spaces of cadlag trajectories (in continuous and discrete time), multisets and compact powersets. These functors can be combined by guarded composition, product and coproduct. Parameter spaces range over the polynomial closure of Giry-like functors. Thus we obtain a class of robust parameterised models which includes the Dirichlet process, various point processes (random sequences with values in Polish spaces) and other classical objects of probability theory. By extending techniques developed in prior work, we show how to reduce the questions of existence, uniqueness, naturality, and continuity of a parameterised model to combinatorial questions only involving finite spaces.

7.8. Bayesian inversion by ω-complete cone duality

Participants: Fredrik Dahlqvist, Vincent Danos, Ilias Garnier [correspondant], Ohad Kammar.

The process of inverting Markov kernels relates to the important subject of Bayesian modelling and learning. In fact, Bayesian update is exactly kernel inversion. In this paper, we investigate how and when Markov kernels (aka stochastic relations, or probabilistic mappings, or simply kernels) can be inverted. We address the question both directly on the category of measurable spaces, and indirectly by interpreting kernels as Markov operators: For the direct option, we introduce a typed version of the category of Markov kernels and use the so-called 'disintegration of measures'. Here, one has to specialise to measurable spaces borne from a simple class of topological spaces -e.g. Polish spaces (other choices are possible). Our method and result greatly simplify a recent development in Ref. [4]. For the operator option, we use a cone version of the category of Markov operators (kernels seen as predicate transformers). That is to say, our linear operators are not just continuous, but are required to satisfy the stronger condition of being ω-chain-continuous.1 Prior work shows that one obtains an adjunction in the form of a pair of contravariant and inverse functors between the categories of L1- and L^∞-cones [3]. Inversion, seen through the operator prism, is just adjunction.2 No topological assumption is needed. We show that both categories (Markov kernels and ω-chain-continuous Markov operators) are related by a family of contravariant functors Tp for $1 \leq$ p $\leq \infty$. The Tp's are Kleisli extensions of (duals of) conditional expectation functors introduced in Ref. [3]. With this bridge in place, we can prove that both notions of inversion agree when both defined: if f is a kernel, and f† its direct inverse, then T∞(f)† = T1(f†).

7.9. Continuous-time Markov chains as transformers of unbounded observables

Participants: Vincent Danos, Ilias Garnier [correspondant], Tobias Heindel, Jakob Simonsen.

We provide broad sufficient conditions for the com-putability of time-dependent averages of stochastic processes of the form f (Xt) where Xt is a continuous-time Markov chain (CTMC), and f is a real-valued function (aka an observable). We consider chains with values in a countable state space S, and possibly unbounded f s. Observables are seen as generalised predicates on S and chains are interpreted as transformers of such generalised predicates, mapping each observable f to a new observable Ptf defined as (Ptf)(x) = Ex(f (Xt)), which represents the mean value of f at time time t as a function of the initial state x. We obtain three results. First, the well-definedness of this operator interpretation is obtained for a large class of chains and observables by restricting Pt to judiciously chosen rescalings of the basic Banach space C0(S) of S-indexed sequences which vanish at infinity. We prove, under appropriate assumptions, that the restricted family Pt forms a strongly continuous operator semigroup (equivalently the time evolution map t $\rightarrow$ Pt is continuous w.r.t. the usual topology on bounded operators). The computability of the time evolution map follows by generic arguments of constructive analysis. A key point here is that the assumptions are flexible enough to accommodate unbounded observables, and we give explicit examples of such using stochastic Petri nets and stochastic string rewriting. Thirdly, we show that if the rate matrix (aka the q-matrix) of the CTMC is locally algebraic on a subspace containing f , the time evolution of projections t $\rightarrow$ (Ptf)(x) is PTIME computable for each x. These results provide a functional analytic alternative to Monte Carlo simulation as test bed for mean-field approximations, moment closure, and similar techniques that are fast, but lack absolute error guarantees.

7.10. Communities in socio-cognitive networks.

Participants: Vincent Danos, Ricardo Honorato-Zimmer [correspondant].

We investigate a recent network model which combines social and cognitive features. Each node in the social network holds a (possibly different) cognitive network that represent its beliefs. In this internal cognitive network a node denotes a concept and a link indicates whether the two linked concepts are taken to be of a similar or opposite nature. We show how these networks naturally organise into communities and use this to develop a method that detects communities in social networks. How they organise depends on the social structure and the ratio between the cognitive and social forces driving the propagation of beliefs.

7.11. Synchronous Balanced Analysis

Participants: Andreea Beica [correspondant], Vincent Danos.

When modeling Chemical Reaction Networks, a commonly used mathematical formalism is that of Petri Nets, with the usual interleaving execution semantics. We aim to substitute to a Chemical Reaction Network, especially a “growth” one (i.e., for which an exponential stationary phase exists), a piecewise synchronous approximation of the dynamics: a resource-allocation-centered Petri Net with maximal-step execution semantics. In the case of unimolecular chemical reactions, we prove the correctness of our method and show that it can be used either as an approximation of the dynamics, or as a method of constraining the reaction rate constants (an alternative to flux balance analysis, using an emergent formally defined notion of “growth rate” as the objective function), or a technique of refuting models.

7.12. Pointless learning

Participants: Florence Clerc, Fredrik Dahlqvist, Vincent Danos, Ilias Garnier [correspondant].

Bayesian inversion is at the heart of probabilistic programming and more generally machine learning. Understanding inversion is made difficult by the pointful (kernel-centric) point of view usually taken in the literature. We develop in a pointless (kernel-free) approach to inver- sion. While doing so, we revisit some foundational objects of probability theory, unravel their category-theoretical underpinnings and show how pointless Bayesian inversion sits naturally at the centre of this construction.

7.13. Survival of the fattest.

Participants: Andreea Beica [correspondant], Vincent Danos, Guillaume Terradot, Andrea Weisse.

Cells derive resources from their environments and use them to fuel the biosynthetic processes that determine cell growth. Depending on how responsive the biosynthetic processes are to the availability of intracellular resources, cells can build up different levels of resource storage. Here we use a recent mathematical model of the coarse-grained mechanisms that drive cellular growth to investigate the effects of cellular resource storage on growth. We show that, on the one hand, there is a cost associated with high levels of storage resulting from the loss of stored resources due to dilution. We further show that, on the other hand, high levels of storage can benefit cells in variable environments by increasing biomass production during transitions from one medium to another. Our results thus suggest that cells may face trade-offs in their maintenance of resource storage based on the frequency of environmental change.

7.14. The algebras of graph rewriting

Participants: Nicolas Behr, Vincent Danos, Ilias Garnier [correspondant], Tobias Heindel.

The concept of diagrammatic combinatorial Hopf algebras in the form introduced for describing the Heisenberg-Weyl algebra is extended to the case of so-called rule diagrams that present graph rewriting rules and their composites. The resulting rule diagram algebra may then be suitably restricted in four different ways to what we call the rule algebras, which are non-commutative, unital associative algebras that implement the algebra of compositions of graph rewriting rules. Notably, our framework reveals that there exist two more types of graph rewriting systems than previously known in the literature, and we present an analysis of the structure of the rule algebras as well as a form of Poincaré-Birkhoff-Witt theorem for the rule diagram algebra. Our work lays the foundation for a fundamentally new way of analyzing graph transformation systems, and embeds this very important concept from theoretical computer science firmly into the realm of mathematical combinatorics and statistical physics.

7.15. PSYNC: A partially synchronous language for fault-tolerant distributed algorithms

Participants: Cezara Drăgoi [correspondant], Thomas Henzinger [IST Austria, Austria], Damien Zufferey [MIT, CSAIL, USA].

Fault-tolerant distributed systems, Programming languages, Verification Fault-tolerant distributed algorithms play an important role in many critical/high-availability applications. These algorithms are notori- ously difficult to implement correctly, due to asynchronous com- munication and the occurrence of faults, such as the network drop- ping messages or computers crashing. We introduce PSYNC in [4], a domain specific language based on the Heard-Of model, which views asynchronous faulty systems as syn- chronous ones with an adversarial environment that simulates asyn- chrony and faults by dropping messages. We define a runtime sys- tem for PSYNC that efficiently executes on asynchronous networks. We formalize the relation between the runtime system and PSYNC in terms of observational refinement. The high-level lockstep ab- straction introduced by PSYNC simplifies the design and imple- mentation of fault-tolerant distributed algorithms and enables auto- mated formal verification. We have implemented an embedding of PSYNC in the SCALA programming language with a runtime system for asynchronous networks. We show the applicability of PSYNC by implementing several important fault-tolerant distributed algorithms and we com- pare the implementation of consensus algorithms in PSYNC against implementations in other languages in terms of code size, runtime efficiency, and verification.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. 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 - December 2018

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.

8.1.2. 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.

8.1.3. VerAsCo

Title: Formally-verified static analyzers and compilers

Type: ANR Ingénierie Numérique Sécurité 2011

Instrument: ANR grant

Duration: September 2011 - June 2016

Coordinator: Inria (France)

Others partners: Airbus France (France), IRISA (France), Inria Saclay (France)

See also: http://www.systematic-paris-region.org/fr/projets/verasco

Abstract: The usefulness of verification tools in the development and certification of critical software is limited by the amount of trust one can have in their results. A first potential issue is *unsoundness* of a verification tool: if a verification tool fails (by mistake or by design) to account for all possible executions of the program under verification, it can conclude that the program is correct while it actually misbehaves when executed. A second, more insidious, issue is *miscompilation*: verification tools generally operate at the level of source code or executable model; a bug in the compilers and code generators that produce the executable code that actually runs can lead to a wrong executable being generated from a correct program.

The project VerAsCoadvocates a mathematically-grounded solution to the issues of formal verifying compilers and verification tools. We set out to develop a generic static analyzer based on abstract interpretation for the C language, along with a number of advanced abstract domains and domain combination operators, and prove the soundness of this analyzer using the Coq proof assistant. Likewise, we will continue our work on the CompCert C formally-verified compiler, the first realistic C compiler that has been mechanically proved to be free of any miscompilation will be continued. Finally, the tool qualification issues that must be addressed before formally-verified tools can be used in the aircraft industry, will be investigated.

8.1.4. AstréeA

Title: Static Analysis of Embedded Asynchronous Real-Time Software

Type: ANR Ingénierie Numérique Sécurité 2011

Instrument: ANR grant

Duration: January 2012 - November 2016

Coordinator: Airbus France (France)

Others partners: École normale supérieure (France)

Inria contact: Antoine Miné

See also: http://www.astreea.ens.fr

Abstract: The focus of the ASTRÉEA project is on the development of static analysis by abstract interpretation to check the safety of large-scale asynchronous embedded software. During the THESEE ANR project (2006–2010), we developed a concrete and abstract models of the ARINC 653 operating system and its scheduler, and a first analyzer prototype. The gist of the ASTRÉEA project is the continuation of this effort, following the recipe that made the success of ASTRÉE: an incremental refinement of the analyzer until reaching the zero false alarm goal. The refinement concerns: the abstraction of process interactions (relational and history-sensitive abstractions), the scheduler model (supporting more synchronisation primitives and taking priorities into account), the memory model (supporting volatile variables), and the abstraction of dynamical data-structures (linked lists). Patrick Cousot is the principal investigator for this project.

8.1.5. VeriFault

This was a PEPS project for one year, coordinated by Cezara Drăgoi, on the topic of fault-tolerant distributed algorithms. These algorithms are notoriously difficult to implement correctly, due to asynchronous communication and the occurrence of faults, such as the network dropping messages or computers crashing. Although fault-tolerant algorithms are at the core of critical applications, there are no automated verification techniques that can deal with their complexity. Due to the complexity distributed systems have reached, we believe it is no longer realistic nor efficient to assume that high level specifications can be proved when development and verification are two disconnected steps in the software production process. Therefore we propose to introduce a domain specific language that has a high-level control structure which focuses on the algorithmic aspects rather than on low-level network and timer code, and makes programs amendable to automated verification.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

ASSUME, ITEA 3 project (Affordable Safe & Secure Mobility Evolution). Affordable Safe & Secure Mobility Evolution

Future mobility solutions will increasingly rely on smart components that continuously monitor the environment and assume more and more responsibility for a convenient, safe and reliable operation. Currently the single most important roadblock for this market is the ability to come up with an affordable, safe multi-core development methodology that allows industry to deliver trustworthy new functions at competitive prices. ASSUME will provide a seamless engineering methodology, which addresses this roadblock on the constructive and analytic side.

8.3. International Research Visitors

8.3.1. Visits of International Scientists

Prof. Kwangkeun Yi Visiteur from Seoul National University, was an invited visitor until Oct 2016.

8.3.1.1. Internships

- Ken Chanseau Saint-Germain, ENS Paris, until Aug 2016
- Marc Chevalier, ENS Lyon, since Sept 2016
- Anton Kulaga, Jul and Aug 2016
- Yoon Seok Ko, Inria, until Jun 2016
- David Romero Suarez, Inria, from Feb 2016 until May 2016]
- Gaelle Candel, Chimie ParisTech

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Jérôme Feret is a member of the editorial board of the Frontiers in Genetics journal and the Open Journal of Modeling and Simulation.

9.1.1.2. Member of the Organizing Committees

Jérôme Feret organized the 40th of Abstract Interpretation at POPL2017 January 21, 2017, Paris, France (co-organizer).

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

Xavier Rival was chair of Static Analysis Symposium SAS 2016, Edinburgh.

Jérôme Feret co-chaired the fifteenth Conference on Computational Methods in Systems Biology - CMSB 2017, September 27–29, 2017, Darmstadt, Germany.

9.1.2.2. Member of the Conference Program Committees

Vincent Danos served on the PC of Computational Methods in Systems Biology, CMSB'16, Cambridge and Complexis'17.

Xavier Rival served on the PC of the 26th European Symposium on Programming (ESOP 2017).

Cezara Drăgoi served on the PC of

- 28th International Conference on Computer-Aided Verification (ERC), CAV 2016,
- 37th INTERNATIONAL CONFERENCE ON APPLICATIONS AND THEORY OF PETRI NETS AND CONCURRENCY, ACSD 2016,
- 23rd International Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS 2017,
- 18th International Conference on Verification, Model Checking, and Abstract Interpretation, VM-CAI 2017,
- ACM SIGPLAN Symposium on Programming Language Design & Implementation, PLDI 2017.

Jérôme Feret served on the PC of

- the 8th International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies - BIOTECHNO 2016,
- the 26th International Symposium on Logic-Based Program Synthesis and transformation - LOPSTR 2016,
- the 23rd Static Analysis Symposium Sept 8-10 2016, Edinburgh,
- 7th International Workshop on Static Analysis and Systems Biology - SASB 2016,
- 14th International Conference on Computational Methods in Systems Biology - CMSB 2016 Sept 21-23 2016, Cambridge, UK,
- Fourth International Conference on Tools and Methods for Program Analysis - TMPA 2017 March 3–4, 2017, Moscow, Russia, JOBIM 2017 July 2-6 2017, Lille, France.

9.1.2.3. Reviewer

Vincent Danos was a reviewer for the 19th International Conference on Foundations of Software Science and Computation Structures (FoSSaCS) 2017, the 26th European Symposium on Programming (ESOP 2017) 2017, LMCS, MSCS.

Cezara Drăgoi was a reviewer for 19th International Conference on Foundations of Software Science and Computation Structures (FoSSaCS) 2017, the 27th International Conference on Concurrency Theory CONCUR 2016, and the 26th European Symposium on Programming (ESOP 2017) 2017.

Xavier Rival was a reviewer for 23rd International Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS 2017 and ACM SIGPLAN Symposium on Programming Language Design & Implementation, PLDI 2017.

Jérôme Feret was a reviewer for the 17th International Conference on Verification, Model Checking, and Abstract Interpretation, VMCAI 2016, the 27th International Conference on Concurrency Theory CONCUR 2016, the 43rd International Colloquium on Automata, Languages and Programming 2016, the 31st ACM/IEEE Symposium on Logic in Computer Science, LICS 2016.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Jérôme Feret is a member of the editorial board of the Frontiers in Genetics journal and the Open Journal of Modeling and Simulation .

9.1.3.2. Reviewer - Reviewing Activities

Xavier Rival was a reviewer for ACM Transactions on Programming Languages and Systems TOPLAS.

Jerome Feret was a reviewer for Theoretical Computer Science 2016.

9.1.4. Invited Talks

Jérôme Feret gave “An overview of the Astrée/AstréeA analyzer. " at Journées scientifiques Inria Rennes, 20-22 june 2016 and at the workshop « Verified Trustworthy Software Systems » Imperial College, 6-7 April 2016.

Vincent Danos talked about ‘Residence: Simons Institute Program Logical Structures and Computations" at CONCUR 2016, Quebec, Aug 25-2 and at Berkeley, Aug 17-Dec 16. He also gave invited talks at SysMod SIG 2016, ISMB, Orlando, Jul 9, Xenobiology 2, XB2, May 24-26, Berlin, IPM Formal Methods Day, Teheran, Jan 10.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence :

- Xavier Rival, “Semantics and Application to Verification",20h, Undergraduate course (L3), at Ecole Normale Supérieure
- Xavier Rival, “Introduction to Static Analysis", 8h, Course at Ecole des Mines de Paris, L3
- Cezara Drăgoi, “Programation concurrente et distribuée", Ecole Politechnique, L2
- Cezara Drăgoi, “Les principes des langages de programmation", Ecole Politechnique, L1
- Jérôme Feret, and Cezara Drăgoi, Mathematics, 40h, L1, FDV Bachelor program (Frontiers in Life Sciences (FdV)), Université Paris-Descartes, France.

Master :

- Vincent, Disruptive technologies and public policies, MSc Public affairs, Sciences Po, France.
- Xavier Rival, Protocol Safety and Verification, Master Course (M2) in the Advanced Communication Networks Master (12h hours), at Polytechnique and Ecole Nationale Supérieure des Telecoms
- Xavier Rival, “Verification" Lab Course at Ecole Polytechnique (M1, 20h)
- Vindent Danos and Jérôme Feret (with Jean Krivine), Computational Biology, 24h, M1. Interdisciplinary Approaches to Life Science (AIV), Master Program, Université Paris-Descartes, France.
- Cezara Drăgoi, Jérôme Feret, Antoine Miné, and Xavier Rival, Abstract Interpretation: application to verification and static analysis, 72h ETD, M2. Parisian Master of Research in Computer Science (MPRI). École normale supérieure. France.

Doctorat : Jérôme Feret, “Interprétation abstraite de modèles de voies de signalisation intracellulaire ", Lectures (3 hours) in the summer school “Modélisation Formelle de Réseaux de Régulation Biologique”, Porquerolles, June 2016 France.

9.2.2. Juries

Jérôme Feret was a member of the recruitment committee for an assistant professor in Paris-Diderot University 2016.

Vincent Danos was examiner and reviewer for the HDR of Sylvain Soliman (Ecole Polytechnique, 7th of December 2016).

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] W. ABOU-JAOUDÉ, D. THIEFFRY, J. FERET.*Formal Derivation of Qualitative Dynamical Models from Biochemical Networks*, in "BioSystems", September 2016, 100 [*DOI :* 10.1016/J.BIOSYSTEMS.2016.09.001], https://hal.inria.fr/hal-01379733.

[2] J. LIU, X. RIVAL.*An array content static analysis based on non-contiguous partitions*, in "Computer Languages, Systems and Structures", 2017, vol. 47, n^{o} 1, p. 104–129 [*DOI :* 10.1016/J.CL.2016.01.005], https://hal.inria.fr/hal-01399837.

[3] A. OUADJAOUT, A. MINÉ, N. LASLA, N. BADACHE.*Static analysis by abstract interpretation of functional properties of device drivers in TinyOS*, in "Journal of Systems and Software", 2016, vol. 120, p. 114–132 [*DOI :* 10.1016/J.JSS.2016.07.030], http://hal.upmc.fr/hal-01350646.

International Conferences with Proceedings

[4] C. DRĂGOI, T. HENZINGER, D. ZUFFEREY.*PSYNC: A Partially Synchronous Language for Fault-Tolerant Distributed Algorithms*, in "POPL", Saint Petersburg , United States, January 2017 [*DOI :* 10.1145/NNNNNNN.NNNNNNN], https://hal.inria.fr/hal-01434325.

[5] J. FERET, K. Q. LY.*Local traces: an over-approximation of the behaviour of the proteins in rule-based models*, in "CMSB 2016 - Fourteenth Conference on Computational Method in Systems Biology", Cambridge, United Kingdom, E. BARTOCCI, P. LIO', N. PAOLETTI (editors), Computational Methods in Systems Biology, Springer, September 2016, vol. 9859, p. 116-131 [*DOI :* 10.1007/978-3-319-45177-0_8], https://hal.inria.fr/hal-01379897.

[6] J. FERET, K. Q. LY.*Reachability analysis via orthogonal sets of patterns*, in "7th International Workshop on Static Analysis and Systems Biology, (SASB 2016)", Edinburgh, United Kingdom, D. SAFRANEK, G. SANGUINETTI (editors), Static Analysis and Systems Biology, Elsevier, September 2016, https://hal.inria.fr/hal-01379902.

[7] T. SUZANNE, A. MINÉ.*From Array Domains to Abstract Interpretation Under Store-Buffer-Based Memory Models*, in "SAS 2016 - 23rd Static Analysis Symposium", Edinburgh, United Kingdom, Lecture Notes in Computer Science, Springer, September 2016, vol. 9837, p. 469-488 [*DOI :* 10.1007/978-3-662-53413-7_23], http://hal.upmc.fr/hal-01360566.

Conferences without Proceedings

[8] C. DRĂGOI, T. HENZINGER, D. ZUFFEREY.*PSYNC: A partially synchronous language for fault-tolerant distributed algorithms*, in "Proceedings of the 43nd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages", Saint Petersburg, Florida, United States, January 2016 [*DOI :* 10.1145/NNNNNNN.NNNNNNN], https://hal.inria.fr/hal-01251199.

[9] A. MINÉ, L. MAUBORGNE, X. RIVAL, J. FERET, P. COUSOT, D. KÄSTNER, S. WILHELM, C. FERDINAND.*Taking Static Analysis to the Next Level: Proving the Absence of Run-Time Errors and Data Races with Astrée*, in "8th European Congress on Embedded Real Time Software and Systems (ERTS 2016)", Toulouse, France, January 2016, https://hal.archives-ouvertes.fr/hal-01271552.

Books or Proceedings Editing

[10] X. RIVAL (editor). *Static Analysis: 23rd International Symposium, (SAS 2016), Edinburgh, UK, September 8-10, 2016, Proceedings*, Springer, Edinburgh, United Kingdom, 2016, vol. LNCS, n° 9837 [*DOI :* 10.1007/978-3-662-53413-7], https://hal.inria.fr/hal-01388205.

[11] C. ZHANG, X. RIVAL (editors). *State Of the Art in Program Analysis: International Workshop, (SOAP 2016), SOAP@PLDI 2016, Santa Barbara, CA, USA, June 14, 2016Proceedings of the 5th ACM SIGPLAN* , ACM, Santa Barbara, United States, 2016 [*DOI :* 10.1145/2931021], https://hal.inria.fr/hal-01388271.

References in notes

[12] 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.

[13] 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.

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

Université Pierre et Marie Curie (Paris 6)

RESEARCH CENTER
Paris

THEME
Computational Neuroscience and Medecine

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:

3.4. - Machine learning and statistics
3.4.1. - Supervised learning
3.4.2. - Unsupervised learning
3.4.4. - Optimization and learning
3.4.5. - Bayesian methods
3.4.7. - Kernel methods
5.3. - Image processing and analysis
5.3.2. - Sparse modeling and image representation
5.3.3. - Pattern recognition
5.3.4. - Registration
5.4.4. - 3D and spatio-temporal reconstruction
5.9. - Signal processing
5.9.4. - Signal processing over graphs
8.2. - Machine learning
8.3. - Signal analysis
8.6. - Decision support

Other Research Topics and Application Domains:

2. - Health
2.2.6. - Neurodegenerative diseases
2.6.1. - Brain imaging

1. Members

Research Scientists

Olivier Colliot [Team leader, CNRS, Senior Researcher, HDR]
Mario Chavez [CNRS, Researcher]
Fabrizio de Vico Fallani [Inria, Starting Research Position, HDR]
Stanley Durrleman [Inria, Researcher, ingénieur des Mines en détachement]

Faculty Members

Didier Dormont [Université Pierre et Marie Curie/AP-HP, Professor, Hospital Neuroradiologist, HDR]
Damien Galanaud [Université Pierre et Marie Curie/AP-HP, Professor, Hospital Neuroradiologist, HDR]
Dominique Hasboun [Université Pierre et Marie Curie, Associate Professor]

Technical Staff

Marie Chupin [CNRS, Permanent, Research Engineer]
Michael Bacci [Inria]
Sabrina Fontanella [IHU-A-ICM, from Sep 2016]
Hugo Dary [Université Pierre et Marie Curie]
Ludovic Fillon [Université Pierre et Marie Curie]
Mathieu Dubois [ICM]

Chabha Azouani [ICM, Clinical Research Associate]
Sonia Djobeir [Université Pierre et Marie Curie, Clinical Research Associate]
Kelly Martineau [ICM, Clinical Research Associate]

PhD Students

Jean-Baptiste Schiratti [Ecole Polytechnique, until Oct 2016]
Barbara Gris [ENS de Cachan]
Géraldine Rousseau [Université Pierre et Marie Curie]
Marika Rudler [Université Pierre et Marie Curie]
Jeremy Guillon [Université Pierre et Marie Curie]
Catalina Obando Forero [Inria]
Alexandre Routier [Université Pierre et Marie Curie]
Jorge Samper Gonzalez [Inria]
Junhao Wen [Université Pierre et Marie Curie]
Wen Wei [Inria, from Oct 2016]
Manon Ansart [Inserm, from Oct 2016]
Fanny Grosselin [MyBrainTechnologies]
Igor Koval [Inserm, from Oct 2016]
Alexandre Bône [Inserm, from May 2016]

Post-Doctoral Fellows

Ana Fouquier [Inria, until Apr 2016]
Xavier Navarro [UPMC]
Soledad Fernandez Garcia [Inria, until Jun 2016]
Takoua Kaaouana [IHU-A-ICM]
Pietro Gori [Inria, until Aug 2016]
Marie-Constance Corsi [Inria, from Jun 2016]
Federico Battiston [Inria/CNRS, from Oct 2016]

Administrative Assistant

Virginie Collette [Inria]

Others

Anne Bertrand [Université Pierre et Marie Curie, Hospital Neuroradiologist - Praticien Hospitalier, poste d'accueil Inria]
Kuldeep Kumar [Inria, Visiting Scientist, from Oct 2016]

2. Overall Objectives

2.1. Introduction

Understanding brain function and its alterations requires the integration of multiple levels of organization, operating at different spatial and temporal scales. The integration of such a large variety of data is now possible thanks to the recent emergence of large-scale multimodal datasets (e.g. Alzheimer's disease neuroimaging initiative [ADNI], gene expression atlases from the Allen Institute...). In this context, mathematical and computational approaches are becoming increasingly important because: i) they provide formalized, operational and flexible frameworks for integrating multiple processes and scales; ii) they allow automated processing and analysis of massive datasets. These approaches can then be used to find biomarkers of a disease, for genotype/phenotype correlations, or to characterize functional responses for instance.

3. Research Program

3.1. General aim

The overall aim of our project is to design new computational and mathematical approaches for studying brain structure (based on anatomical and diffusion MRI) and functional connectivity (based on EEG, MEG and intracerebral recordings). The goal is to transform raw unstructured images and signals into formalized, operational models such as geometric models of brain structures, statistical population models, and graph-theoretic models of brain connectivity. This general endeavor is addressed within the three following main objectives.

3.2. Modeling brain structure: from imaging to geometric models

Structural MRI (anatomical or diffusion-weighted) allows studying in vivo the anatomical architecture of the brain. Thanks to the constant advance of these imaging techniques, it is now possible to visualize various anatomical structures and lesions with a high spatial resolution. Computational neuroanatomy aims at building models of the structure of the human brain, based on MRI data. This general endeavor requires addressing the following methodological issues: i) the extraction of geometrical objects (anatomical structures, lesions, white matter tracks...) from anatomical and diffusion-weighted MRI; ii) the design of a coherent mathematical framework to model anatomical shapes and compare them across individuals. Within this context, we pursue the following objectives.

First, we aim to develop new methods to segment anatomical structures and lesions. We are most specifically interested in the hippocampus, a structure playing a crucial role in Alzheimer's disease, and in lesions of vascular origin (such as white matter hyperintensities and microbleeds). We pay particular attention to the robustness of the approaches with respect to normal and pathological anatomical variability and with respect to differences in acquisition protocols, for application to multicenter studies. We dedicate specific efforts to the validation on large populations of coming from patients data acquired in multiple centers.

Then, we develop approaches to estimate templates from populations and compare anatomical shapes, based on a diffeomorphic deformation framework and matching of distributions. These methods allow the estimation of a prototype configuration (called template) that is representative of a collection of anatomical data. The matching of this template to each observation gives a characterization of the anatomical variability within the population, which is used to define statistics. In particular, we aim to design approaches that can integrate multiple objects and modalities, across different spatial scales.

3.3. Modeling dynamical 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 and perceptual tasks. Indeed, mounting evidence suggests that impairment of such mechanisms might be the first step of a chain of events triggering several neurological disorders, such as the abnormal synchronization of epileptic activities. Naturally, neuroimaging studies investigating functional connectivity in the brain have become increasingly prevalent.

Our team develops a framework for the characterization of brain connectivity patterns, based on connectivity descriptors from the theory of complex networks. The description of the connectivity structure of neural networks is able to characterize for instance, the configuration of links associated with rapid/abnormal synchronization and information transfer, wiring costs, resilience to certain types of damage, as well as the balance between local processing and global integration. Furthermore, we propose to extend this framework to study the reconfiguration of networks over time. Indeed, neurophysiological data are often gathered from longitudinal recording sessions of the same subject to study the adaptive reconfiguration of brain connectivity. Finally, connectivity networks are usually extracted from different brain imaging modalities (MEG, EEG, fMRI or DTI) separately. Methods for combining the information carried by these different networks are still missing. We thus propose to combine connectivity patterns extracted from each modality for a more comprehensive characterization of networks.

3.4. Methodologies for large-scale datasets

Until recently, neuroimaging studies were often restricted to series of about 20-30 patients. As a result, such studies had a limited statistical power and could not adequately model the variability of populations. Thanks to wider accessibility of neuroimaging devices and important public and private funding, large-scale studies including several hundreds of patients have emerged in the past years. In the field of Alzheimer's disease (AD) for instance, one can cite the Alzheimer's Disease Neuroimaging Initiative (ADNI) including about 800 subjects (patients with AD or mild cognitive impairment (MCI) and healthy controls) or the French cohort MEMENTO including about 2000 subjects with memory complaint. These are most often multicenter studies in which patients are recruited over different centers and images acquired on different scanners. Moreover, cohort studies include a longitudinal component: for each subject, multiple images are acquired at different time points. Finally, such datasets often include multimodal data: neuroimaging, clinical data, cognitive tests and genomics data. These datasets are complex, high-dimensional and often heterogeneous, and thus require the development of new methodologies to be fully exploited.

In this context, our objectives are:

- to develop methodologies to acquire and standardize multicenter neuroimaging data;
- to develop imaging biomarkers based on machine learning and longitudinal models;
- to design multimodal analysis approaches for bridging anatomical models and genomics.

The first two aspects focus on neuroimaging and are tightly linked with the CATI project. The last one builds on our previous expertise in morphometry and machine learning, but aims at opening new research avenues combining imaging and "omics" data. This is developed in strong collaboration with the new biostatistics/bioinformatics platform of the IHU-A-ICM.

4. Application Domains

4.1. Introduction

We develop different applications of our new methodologies to brain pathologies, mainly neurodegenerative diseases, epilepsy and cerebrovascular disorders. These applications aim at:

- better understanding the pathophysiology of brain disorders;
- designing biomarkers of pathologies for diagnosis, prognosis and assessment of drug efficacy;
- developping brain computer interfaces for clinical applications;
- improving the localisation of stimulation targets in Deep Brain Stimulation protocol.

These applications are developed in close collaboration with biomedical researchers of the ICM and clinicians of the Pitié-Salpêtrière hospital.

4.2. Understanding brain disorders

The approaches that we develop allow to characterize anatomical and functional alterations, thus making it possible to study these alterations in different clinical populations. This can provide provide new insights into the mechanisms and progression of brain diseases. This typically involves the acquisition of neuroimaging data in a group of patients with a given pathology and in a group of healthy controls. Measures of anatomical and functional alterations are then extracted in each subject (for instance using segmentation of anatomical structures, shape models or graph-theoretic measures of functional connectivity). Statistical analyses are then performed to identify: i) significant differences between groups, ii) correlations between anatomical/functional alterations on the one hand, and clinical, cognitive or biological measures on the other hand, iii) progression of alterations over time.

We propose to apply our methodologies to study the pathophysiology of neurodegenerative diseases (mostly Alzheimer's disease and fronto-temporal dementia), epilepsy, cerebrovascular pathologies and neurodevelopmental disorders (Gilles de la Tourette syndrome). In neurodegenerative diseases, we aim at establishing the progression of alterations, starting from the early and even asymptomatic phases. In Gilles de la Tourette syndrome, we study the atypical anatomical patterns that may contribute to the emergence of symptoms. In epilepsy, we aim at studying the relationships between the different functional and structural components of epileptogenic networks.

4.3. Biomarkers for diagnosis, prognosis and clinical trials

Currently, the routine diagnosis of neurological disorders is mainly based on clinical examinations. This is also true for clinical trials, aiming to assess the efficacy of new treatments. However, clinical diagnoses only partially overlap with pathological processes. For instance, the sensitivity and specificity of clinical diagnosis of Alzheimer's disease (AD) based on established consensus criteria are of only about 70-80% compared to histopathological confirmation. Furthermore, the pathological processes often begin years before the clinical symptoms. Finally, clinical measures embed subjective aspects and have a limited reproducibility and are thus not ideal to track disease progression. It is thus crucial to supplement clinical examinations with biomarkers that can detect and track the progression of pathological processes in the living patient. This has potentially very important implications for the development of new treatments as it would help: i) identifying patients with a given pathology at the earliest stage of the disease, for inclusion in clinical trials; ii) providing measures to monitor the efficacy of treatments.

The derivation of biomarkers from image analysis approaches requires large-scale validation in well-characterized clinical populations. The ARAMIS team is strongly engaged in such efforts, in particular in the field of neurodegenerative disorders. To that purpose, we collaborate to several national studies (see section Partnerships) that involve multicenter and longitudinal acquisitions. Moreover, ARAMIS is strongly involved in the CATI which manages over 15 multicenter studies, including the national cohort MEMENTO (2000 patients).

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 base 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. The BCI approaches that we develop are based on the characterization of the information contained in the functional connectivity patterns. We expect to significantly increase the performance of the BCI system with respect to the sole use of standard power spectra of the activity generated by single local brain areas. Such an improvement will concretely provide the user with a more precise control of the external environment in open-loop BCI tasks and a more coherent feedback in the closed-loop BCI schemes.

4.5. Deep Brain Stimulation

Deep Brain Stimulation (DBS) is a surgical technique, which consists in sending electrical impulses, through implanted electrodes, to specific parts of the brain for the treatment of movement and affective disorders. The technique has been initially developped for otherwise-treatment-resistant patients with essential tremors or Parkinson's disease. Its benefit in other affections, such as dystonia, obsessive-compulsive disorders, Tourette syndrome is currently investigated. The localisation of the stimulation target in specific nucleus in deep brain regions is key to the success of the surgery. This task is difficult since the target nucleus, or the precise sub-territory of a given nucleus is rarely visible in the Magnetic Resonance Image (MRI) of the patients. To address this issue, a possible technique is to personalize a high-resolution histological atlas of the brain to each patient. This personalization is achieved by registering the histological atlas, which consists of an image and meshes of deep brain structures, to the pre-operative MRI of each patient. The registration is currently done by optimally aligning image intensities in the atlas and patient's MRI using a block-matching algorithm. The linear nature of the transform makes the technique robust at the cost of a lack of precision, especially for elderly patients with expanded ventricles. We investigate the use of non-linear registration techniques to optimally align both image intensities and contours of visible structures surrounding the target. We expect to improve the localisation of the target for patients with large ventricles while keeping the method robust in all cases.

5. Highlights of the Year

5.1. Highlights of the Year

- Stanley Durrleman's ERC Starting Grant "LEASP" has started.
- H2020 project EuroPOND, under societal challenge "Personalizing Health and Care" has started.
- ANR-NIH project NETBCI, under the "Collaborative Research in Computational Neuroscience" program (CRCNS) has started.
- The team has been awarded the ANR-NIH project HIPLAY7, under the "Collaborative Research in Computational Neuroscience" program (CRCNS)
- The team has been awarded the ANR project BRANDY, under the generic call programme "Vie, Sante et Bien-etre", Project duration: 2017-2020
- ARAMIS participates to the Human Brain Project (European Flagship).
- Anne Bertrand was awarded a one year Inria-APHP interface contract (i.e., "poste d'accueil"), allowing her to work half-time in the ARAMIS project team, from november 2016 to november 2017.
- Pietro Gori and Barbara Gris successfully defended their PhD.
- S. Durrleman has been appointed associate editor of IEEE Transactions on Medical Imaging (TMI).

6. New Software and Platforms

6.1. Clinica

KEYWORDS: Multimodal neuroimaging - anatomical MRI - diffusion MRI - functional MRI - PET - EEG/MEG

FUNCTIONAL DESCRIPTION

Clinica is a software platform for multimodal brain image analysis in clinical research studies. It aims at integrating a comprehensive set of processing tools for the main neuroimaging modalities: MRI (anatomical, functional, diffusion), PET and EEG/MEG. For each modality, it allows to easily extract various types of features (regional measures, parametric maps, surfaces, curves, networks) that can be 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 and in-house developments. It provides an integrated data management system to store raw and processing data.

- Participants: Olivier Colliot, Stanley Durrleman, Fabrizio De Vico Fallani, Michael Bacci, Alexandre Routier, Jorge Samper-Gonzalez, Junhao Wen, Jérémy Guillon, Sabrina Fontanella, Thomas Jacquemont
- Contact: Olivier Colliot

6.2. Brain Networks Toolbox

KEYWORDS: Neuroimaging - Medical imaging
FUNCTIONAL DESCRIPTION

Brain Networks Toolbox is a collection of Matlab routines developed to quantify topological metrics of complex brain networks.

- Participants: Mario Chavez and Fabrizio De Vico Fallani
- Contact: Mario Chavez
- URL: https://sites.google.com/site/fr2eborn/download

6.3. 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!
FUNCTIONAL 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 optimal 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: Stanley Durrleman, Alexandre Routier, Pietro Gori, Marcel Prastawa, Ana Fouquier, Joan Alexis Glaunès, Benjamin Charlier, Cedric Doucet, Michael Bacci and Barbara Gris
- Partners: Université de Montpellier 2 - Université Paris-Descartes - University of Utah
- Contact: Stanley Durrleman
- URL: http://www.deformetrica.org/

6.4. SACHA

Segmentation Automatisée Compétitive de l'Hippocampe et de l'Amygdale
KEYWORDS: Neuroimaging - 3D - Hippocampus - Amygdala - Brain scan - Medical imaging
SCIENTIFIC DESCRIPTION

The current stable version is fully automatic and focused on cross-sectional segmentation. The software can be used both as a command-line program or through a graphical user interface (GUI). The core of the program is coded in C++. It has a dependency to the AIMS library and preprocessing steps rely on processes in Matlab from SPM. The GUI is coded in Python and is based on BrainVISA.

FUNCTIONAL DESCRIPTION

SACHA is a software for the fully automatic segmentation of the hippocampus and the amygdala from MRI 3D T1 brain scans. It has been validated in various populations including healthy controls and patients with Alzheimer's disease, epilepsy and depression. It has been successfully applied to over 3,000 subjects, both controls, from adolescents to elderly subjects, and patients with different types of pathologies.

- Participants: Marie Chupin and Ludovic Fillon
- Contact: Marie Chupin

6.5. WHASA

White matter Hyperintensity Automatic Segmentation Algorithm
KEYWORDS: Health - Neuroimaging - Biomedical imaging
SCIENTIFIC DESCRIPTION

The current stable version is fully automatic and focused on cross-sectional segmentation. The software can be used both as a Matlab command-line or through a graphical user interface (GUI). The core of the program is coded in Matlab. It has a dependency to the SPM environment. The GUI is coded in Python and is based on BrainVISA.
FUNCTIONAL DESCRIPTION

WHASA ("White matter Hyperintensity Automatic Segmentation Algorithm") is a software for the fully automatic segmentation of age-related white matter hyperintensities from MRI FLAIR and 3D T1 brain scans. It has been validated on a population showing a wide range of lesion load, and is being further evaluated on elderly subjects with few clinical abnormalisties and with different acquisition characteristics.

- Participants: Marie Chupin, Ludovic Fillon and Thomas Samaille
- Contact: Marie Chupin

6.6. QualiCATI

KEYWORDS: Health - Neuroimaging - Medical imaging
SCIENTIFIC DESCRIPTION

QualiCATI requires training for the visual parts, and is closely linked with a team of clinical research assistants. It has been used to analyse about 5000 subjects from about 15 multi centre research projects initiated before or after the CATI started. Other modules will be added in the future to embed new aspects of the MRI protocol proposed by the CATI. The Aramis team is in charge of the second and third modules and jointly in charge of the first module. The software is centered on a graphical user interface (GUI). The whole program is coded in Python within the pyPTK environment. It has dependencies to SPM and brainVISA environments as well as specific tools for DICOM management.
FUNCTIONAL DESCRIPTION

QualiCATI is a software designed for comprehensive quality control of multimodal MRI data acquisition in large multicentre clinical studies. The software is built as a platform receiving several modules, developped by several CATI engineers. The first module is dedicated to acquisition requirement checking and conversion to nifti format. The second module aims at making 3DT1 acquisition quality check more systematic, and relies both on visual inspection and quantitative indices. The third module allows a simultaneous evaluation of the clinical part of the CATI acquisition protocol. The fourth module embeds automatic indices to evaluate resting state fMRI acquisition. The fifth module is dedicated to first preprocessings and quality indices for dMRI. The sixth module is dedicated to qMRI, with visual and automated quality control together with preprocessings. The last module is dedicated to data and project management.

- Participants: Marie Chupin and Hugo Dary
- Contact: Marie Chupin

7. New Results

7.1. A Bayesian Framework for Joint Morphometry of Surface and Curve meshes in Multi-Object Complexes

Participants: Pietro Gori [Correspondant], Olivier Colliot, Linda Marrakchi-Kacem, Yulia Worbe, Alexandre Routier, Cyril Poupon, Andreas Hartmann, Nicholas Ayache, Stanley Durrleman.

We present a Bayesian framework for atlas construction of multi-object shape complexes comprised of both surface and curve meshes (Figure 1). It is general and can be applied to any parametric deformation framework and to all shape models with which it is possible to define probability density functions (PDF). Here, both curve and surface meshes are modelled as Gaussian random varifolds, using a finite-dimensional approximation space on which PDFs can be defined. Using this framework, we can automatically estimate the parameters balancing data-terms and deformation regularity, which previously required user tuning. Moreover, it is also possible to estimate a well-conditioned covariance matrix of the deformation parameters. We also extend the proposed framework to data-sets with multiple group labels. Groups share the same template and their deformation parameters are modelled with different distributions. We can statistically compare the groups'distributions since they are defined on the same space. We test our algorithm on 20 Gilles de la Tourette patients and 20 control subjects, using three sub-cortical regions and their incident white matter fiber bundles. We compare their morphological characteristics and variations using a single diffeomorphism in the ambient space. The proposed method will be integrated with the Deformetrica software package.

More details in [15].

7.2. Parsimonious Approximation of Streamline Trajectories in White Matter Fiber Bundles

Participants: Pietro Gori [Correspondant], Olivier Colliot, Linda Marrakchi-Kacem, Fabrizio de Vico Fallani, Mario Chavez, Yulia Worbe, Alexandre Routier, Cyril Poupon, Andreas Hartmann, Nicholas Ayache, Stanley Durrleman.

Fiber bundles stemming from tractography algorithms contain many streamlines. They require therefore a great amount of computer memory and computational resources to be stored, visualised and processed. We propose an approximation scheme for fiber bundles which results in a parsimonious representation of weighted prototypes. Prototypes are chosen among the streamlines and they represent groups of similar streamlines. Their weight is related to the number of approximated streamlines. Both streamlines and prototypes are modelled as weighted currents. This computational model does not need point-to-point correspondences and two streamlines are considered similar if their endpoints are close to each other and if their pathways follow similar trajectories. Moreover, the space of weighted currents is a vector space with a closed-form metric. This permits easy computation of the approximation error and the selection of the prototypes is based on the minimisation of this error. We propose an iterative algorithm which approximates independently and simultaneously all the fascicles of the bundle in a fast and accurate way. We show that the resulting representation preserves the shape of the bundle and it can be used to accurately reconstruct the original structural connectivity (Figure 2). We evaluate our algorithm on bundles obtained from both deterministic and probabilistic tractography algorithms. The resulting approximations use on average only 2% of the original streamlines as prototypes. This drastically reduces the computational burden of the processes where the geometry of the streamlines is considered. We demonstrate its effectiveness using as example the registration between two fiber bundles.

More details in [14].

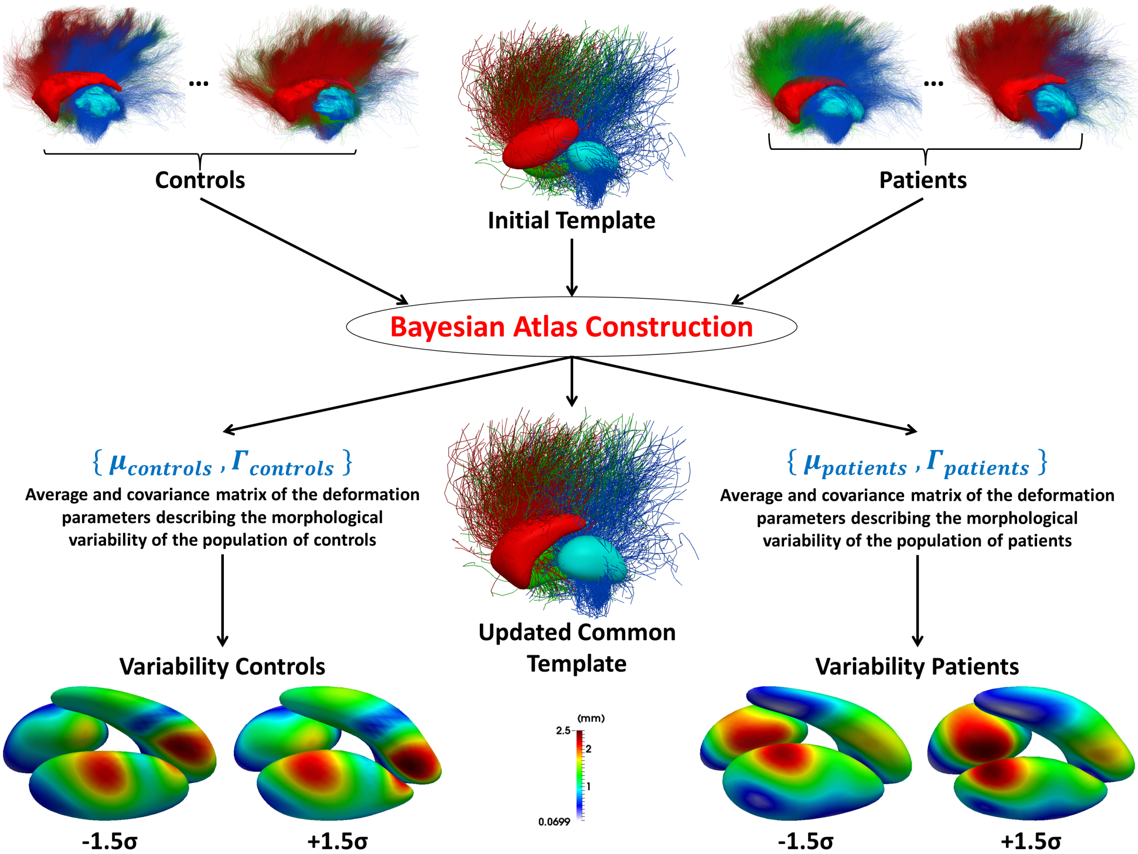

Figure 1. Bayesian framework for atlas construction of multi-object shape complexes comprised of both surface and curve meshes.

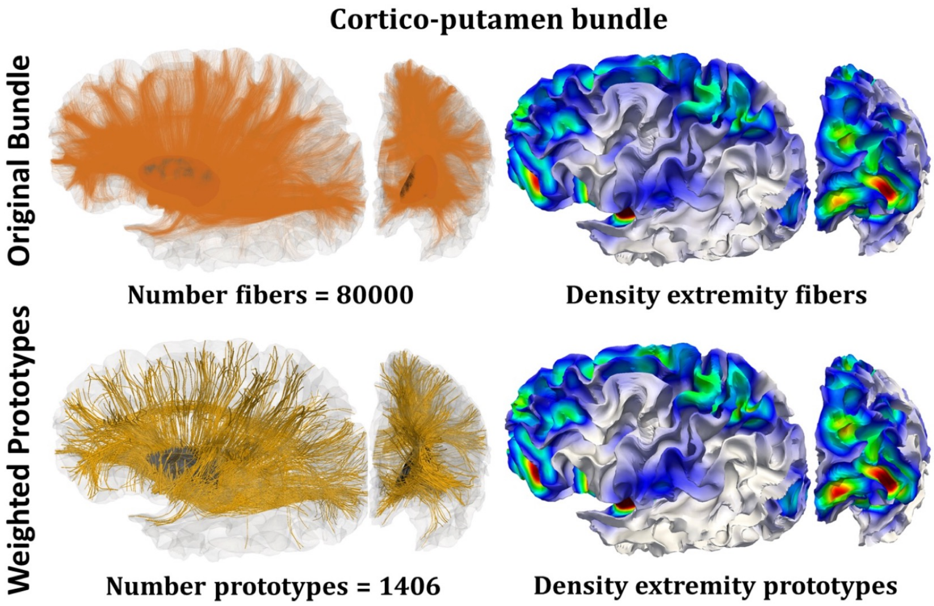

Figure 2. Weighted prototype approximations of a fiber bundle. As it is possible to notice, our approximation alters neither the global shape of the bundle nor the densities of the endpoints onto the cortical surface.

7.3. White matter lesions in FTLD: distinct phenotypes characterize GRN and C9ORF72 mutations

Participants: Fatima Ameur, Olivier Colliot, Didier Dormont, Alexis Brice, Isabelle Le Ber, Anne Bertrand [Correspondant].

Frontotemporal lobar degeneration (FTLD) has a high frequency of genetic forms; the 2 most common are GRN (progranulin) and C9ORF72 mutations. Recently, our group reported extensive white matter (WM) lesions in 4 patients with FTLD caused by GRN mutation, in the absence of noteworthy cardiovascular risk factors in line with other studies in GRN mutation carriers. Here we compared the characteristics of frontal WM lesions in patients with behavioral variant of FTLD (bv-FTLD) caused by GRN and C9ORF72 mutations. We found that WM lesions were more frequent and more atypical on both sides in the GRN group than in the control group and the C9ORF72 group.

More details in [3].

7.4. Riemannian geometry applied to detection of respiratory states from EEG signals: the basis for a brain-ventilator interface

Participants: Xavier Navarro-Sune, Anna Hudson, Fabrizio de Vico Fallani, Jacques Martinerie, Adrien Witon, Pierre Pouget, Mathieu Raux, Thomas Similowski, Mario Chavez [Correspondant].

During mechanical ventilation, patient-ventilator disharmony is frequently observed and may result in increased breathing effort, compromising the patient's comfort and recovery. This circumstance requires clinical intervention and becomes challenging when patients are sedated or verbal communication is difficult. In this work, we propose a brain computer interface (BCI) to automatically and non-invasively detect patient-ventilator disharmony from electroencephalographic (EEG) signals: a brain-ventilator interface. Our framework exploits the cortical activation provoked by the inspiratory compensation when the subject and the ventilator are desynchronized (Figure 3). Use of a one-class approach and Riemannian geometry of EEG covariance matrices allows effective classification of respiratory states. The BVI is validated on nine healthy subjects that performed different respiratory tasks that mimic a patient-ventilator disharmony. Results evidence that classification performances, in terms of areas under ROC curves, are significantly improved using EEG signals compared to detection based on air flow. Reduction in the number of electrodes that can achieve discrimination can often be desirable (e.g. for portable BCI systems). By using an iterative channel selection technique, the Common Highest Order Ranking (CHOrRa), we find that a reduced set of electrodes (n=6) can slightly improve for an intra-subject configuration, and it still provides fairly good performances for a general inter-subject setting. Results support the discriminant capacity of our approach to identify anomalous respiratory states, by learning from a single training set containing only normal respiratory epochs. The proposed framework opens the door to brain-ventilator interfaces for monitoring patient's breathing comfort and adapting ventilator parameters to patient respiratory needs.

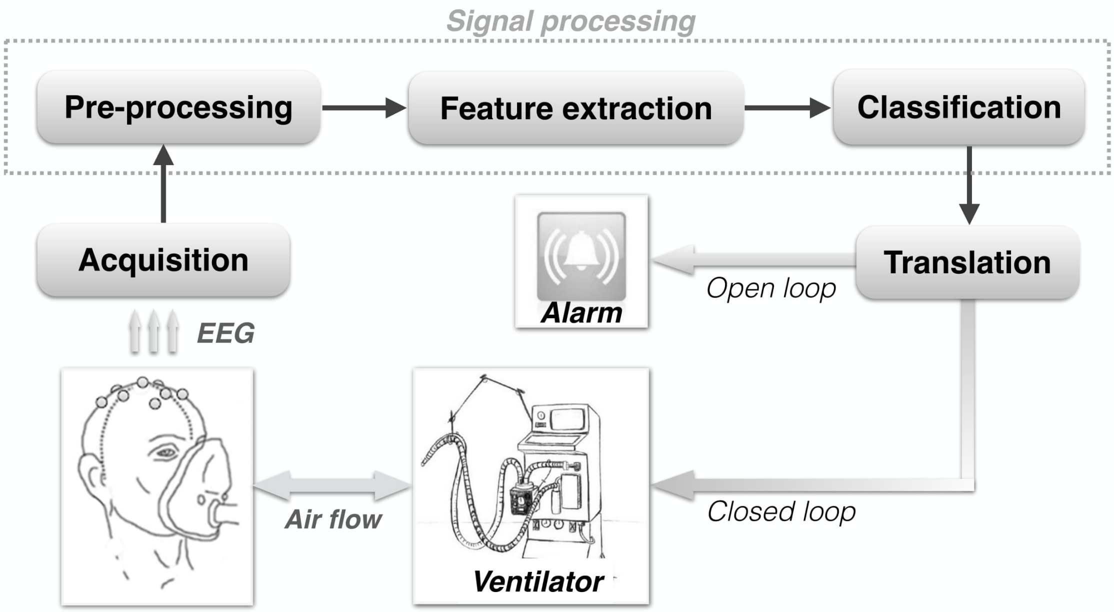

Figure 3. Scheme of a brain-ventilator interface.

More details in [25].

7.5. Interhemispheric Connectivity Characterizes Cortical Reorganization in Motor-Related Networks After Cerebellar Lesions

Participants: Fabrizio de Vico Fallani, Silvia Clausi, Maria Leggio, Mario Chavez, Miguel Valencia, Anton Giulio Maglione, Fabio Babiloni, Febo Cincotti, Donatella Mattia, Marco Molinari [Correspondant].

Although cerebellar-cortical interactions have been studied extensively in animal models and humans using modern neuroimaging techniques, the effects of cerebellar stroke and focal lesions on cerebral cortical processing remain unknown. In the present study, we analyzed the large-scale functional connectivity at the cortical level by combining high-density electroencephalography (EEG) and source imaging techniques to evaluate and quantify the compensatory reorganization of brain networks after cerebellar damage. The experimental protocol comprised a repetitive finger extension task by 10 patients with unilateral focal cerebellar lesions and 10 matched healthy controls. A graph theoretical approach was used to investigate the functional reorganization of cortical networks. Our patients, compared with controls, exhibited significant differences at global and local topological level of their brain networks. An abnormal rise in small-world network efficiency was observed in the gamma band (30-40 Hz) during execution of the task, paralleled by increased long-range connectivity between cortical hemispheres (Figure 4). Our findings show that a pervasive reorganization of the brain network is associated with cerebellar focal damage and support the idea that the cerebellum boosts or refines cortical functions. Clinically, these results suggest that cortical changes after cerebellar damage are achieved through an increase in the interactions between remote cortical areas and that rehabilitation should aim to reshape functional activation patterns. Future studies should determine whether these hypotheses are limited to motor tasks or if they also apply to cerebro-cerebellar dysfunction in general.

More details in [11].

7.6. A topological criterion for filtering information in complex brain networks

Participants: Fabrizio de Vico Fallani [Correspondant], Vito Latora, Mario Chavez.

In many biological systems, the network of interactions between the elements can only be inferred from experimental measurements. In neuroscience, non-invasive imaging tools are extensively used to derive either structural or functional brain networks in-vivo. As a result of the inference process, we obtain a matrix of values corresponding to an unrealistic fully connected and weighted network. To turn this into a useful sparse network, thresholding is typically adopted to cancel a percentage of the weakest connections. The structural properties of the resulting network depend on how much of the inferred connectivity is eventually retained. However, how to fix this threshold is still an open issue. We introduce a criterion, the efficiency cost optimization (ECO), to select a threshold based on the optimization of the trade-off between the efficiency of a network and its wiring cost. We prove analytically and we confirm through numerical simulations that the connection density maximizing this trade-off emphasizes the intrinsic properties of a given network, while preserving its sparsity. Moreover, this density threshold can be determined a-priori, since the number of connections to filter only depends on the network size according to a power-law. We validate this result on several brain networks, from micro- to macro-scales, obtained with different imaging modalities. Finally, we test the potential of ECO in discriminating brain states with respect to alternative filtering methods. ECO advances our ability to analyze and compare biological networks, inferred from experimental data, in a fast and principled way.

More details in [12].

7.7. Robust imaging of hippocampal inner structure at 7T: in vivo acquisition protocol and methodological choices

Participants: Linda Marrakchi-Kacem [Correspondant], Alexandre Vignaud, Julien Sein, Johanne Germain, Thomas Henry, Cyril Poupon, Lucie Hertz-Pannier, Stephane Lehericy, Olivier Colliot, Pierre-François Van de Moortele, Marie Chupin.

Motion is a crucial issue for ultra-high resolution imaging, such as can be achieved with 7T MRI. An acquisition protocol was designed for imaging hippocampal inner structure at 7T. It relies on a compromise between anatomical details visibility and robustness to motion. In order to reduce acquisition time and motion artifacts, the full slab covering the hippocampus was split into separate slabs with lower acquisition time. A robust registration approach was implemented to combine the acquired slabs within a final 3D-consistent high-resolution slab covering the whole hippocampus. Evaluation was performed on 50 subjects overall, made of three groups of subjects acquired using three acquisition settings; it focused on three issues: visibility of hippocampal inner structure, robustness to motion artifacts and registration procedure performance. Overall,

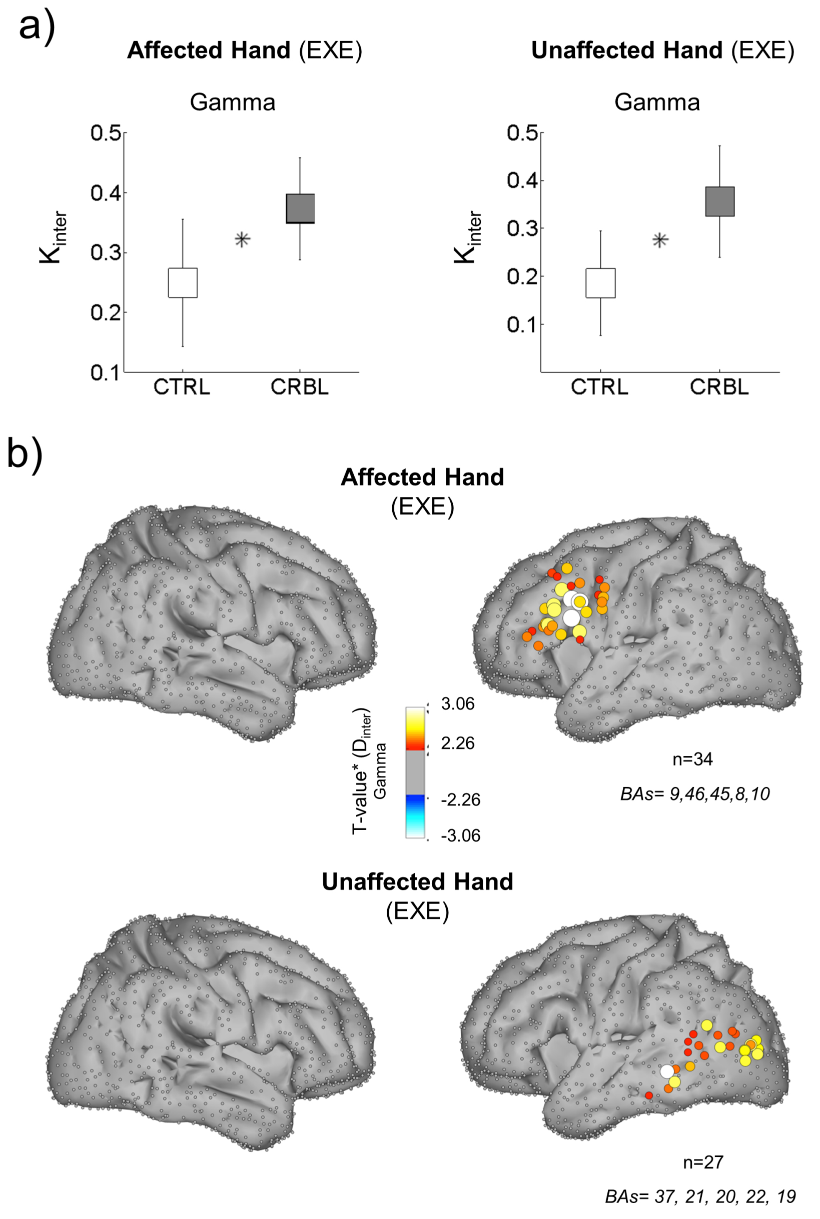

Figure 4. Gamma-band inter-hemispheric density (Kinter) and statistical contrasts of node degrees for brain networks during movement execution Panel a) Averaged Kinter values in the Gamma band for the CTRL and CRBL groups for the affected and unaffected hand conditions in the EXE phase. Panel b) T-value maps of the between-groups contrasts for the node-degree values over lateral views of the MNI cortical model in the Tailarach space in the affected (upper part) and unaffected (bottom part) hand conditions in the EXE phase.

T2-weighted acquisitions with interleaved slabs proved robust. Multi-slab registration yielded high quality datasets in 96% of the subjects, thus compatible with further analyses of hippocampal inner structure. Multi-slab acquisition and registration setting is efficient for reducing acquisition time and consequently motion artifacts for ultra-high resolution imaging of the inner structure of the hippocampus.

More details in [22].

7.8. Improved cerebral microbleeds detection using their magnetic signature on T2*-phase-contrast: a comparison study in a clinical setting

Participants: Takoua Kaaouana [Correspondant], Anne Bertrand, Fatma Ouamer, Bruno Law-Ye, Nadya Pyatigorskaya, Ali Bouyahia, Nathalie Thiery, Carole Dufouil, Christine Delmaire, Didier Dormont, Ludovic de Rochefort, Marie Chupin.

In vivo detection of cerebral microbleeds (CMBs) from T2* gradient recalled echo (GRE) magnitude image suffers from low specificity, modest inter-rater reproducibility and is biased by its sensitivity to acquisition parameters. New methods were proposed for improving this identification, but they mostly rely on 3D acquisitions, not always feasible in clinical practice. A fast 2D phase processing technique for computing internal field maps (IFM) has been shown to make it possible to characterize CMBs through their magnetic signature in routine clinical setting, based on 2D multi-slice acquisitions. However, its clinical interest for CMBs identification with respect to more common images remained to be assessed. To do so, systematic experiments were undertaken to compare the ratings obtained by trained observers with several image types, T2* magnitude, Susceptibility Weighted Imaging reconstructions (SWI) and IFM built from the same T2*-weighted acquisition. 15 participants from the MEMENTO multi-center cohort were selected: six subjects with numerous CMBs (20+/-6 CMBs), five subjects with a few CMBs (2 +/-1 CMBs) and four subjects without CMB. 2D multi-slice T2* GRE sequences were acquired on Philips and Siemens 3T systems. After pilot experiments, T2* magnitude, Susceptibility Weighted Imaging (SWI) minimum intensity projection (mIP) on three slices and IFM were considered for the rating experiments. A graphical user interface (GUI) was designed in order to consistently display images in random order. Six raters of various background and expertise independently selected “definite” or “possible” CMBs. Rating results were compared with respect to a specific consensus reference, on both lesion and subject type points Results: IFM yielded increased sensitivity and decreased false positives rate (FPR) for CMBs identification compared to T2* magnitude and SWI-mIP images. Inter-rater variability was decreased with IFM when identifying subjects with numerous lesions, with only a limited increase in rating time. IFM thus appears as an interesting candidate to improve CMBs identification in clinical setting.

More details in [19].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Air-Liquide Medical Systems

Participants: Mario Chavez [Correspondant], Xavier Navarro.

Project title: Real-time characterisation of respiratory states from EEG

Funded in 2014

Amount: 370 K€

Coordinator: Thomas Similowski

Other partners: UPMC, Inserm UMR 1158

Abstract: The project aims at developing a real-time brain computer interface (BCI) for the monitoring of respiratory states from scalp EEG data of healthy volunteers and patients, recorded at the laboratory, hospital ward, operating room or intensive care units..

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

9.1.1.1. ANR-NIH CRCNS

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, Université Penn, Etats-units

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.

9.1.1.2. 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.

9.1.1.3. 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 (μ- 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 μ-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.

9.1.2. IHU

9.1.2.1. General program

Participants: Olivier Colliot, Mario Chavez, Stanley Durrleman, Marie Chupin, Didier Dormont, Dominique Hasboun, Damien Galanaud, 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.

9.1.2.2. ICM-Internal Research projects

Participants: Anne Bertrand [Correspondant], Takoua Kaaouana, Benoit Delatour, Alexandra Petiet.

Project title: The Histo-MRI project: targeting MR signature of tauopathy from micro- to macroscopy

Founded 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.

9.1.2.3. ICM-Internal Research projects

Participants: Mario Chavez [Correspondant], Fabrizio de Vico Fallani [Correspondant].

Project title: Non-invasive manipulation of brain synchrony to enhance brain function and rehabilitate faulty cognition in humans: A proof of concept

Founded in 2014

Coordinator: Antoni Valero Cabre (ICM-team “Dynamiques Cérébrales, Plasticité et Rééducation”)

Other partners: Service des Urgences Cérébro-Vasculaires de l’Hôpital Pitié-Salpêtrière, Paris.

The long-term goal of this project is to develop the use of non-invasive manipulation of abnormal cerebral oscillations underlying cognitive activity to restore brain function in neurological patients. Cognitive functions emerge from large distributed networks organized in space and time. The short-term goal of this application is to study the causal role played by oscillatory activity in visual awareness and test whether their manipulation by non-invasive brain stimulation has the potential to restore its function in stroke patients.

9.1.2.4. ICM Big Brain Theory Program

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

Founded in 2016-2017

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.

9.1.2.5. IFR49-Internal Research projects

Participants: Mario Chavez [Correspondant], Fabrizio de Vico Fallani [Correspondant].

Project title: Exploring the impact and time frequency signature of rhythmic patterns of Transcranial Magnetic Stimulation (TMS) on network activity by Magneto-Encephalography (MEG)

Founded in 2014

Coordinator: Antoni Valero Cabre (ICM-team "Dynamiques Cérébrales, Plasticité et Rééducation")

Other partners: TMS, EEG and MEG technical platforms of the ICM at the Hopital Pitié-Salptrière; and Service des Urgences Cérébro-Vasculaires de l'Hôpital Pitié-Salpêtrière, Paris.

The long-term goal of this project is to better understand the ability of non invasive neurostimulation to induce lasting local and distributed reorganization effects in the human brain to better plan and document therapies for patients. The short-term goal of this application is to develop a new mapping procedure to be able to capture and characterize in terms of oscillatory activity the lasting impact of repetitive Transcranial Magnetic Stimulation (TMS) on specific brain regions and associated networks.

9.1.3. CATI (Alzheimer Plan)

Participants: Olivier Colliot [Correspondant], Marie Chupin [Correspondant], Stanley Durrleman, Didier Dormont, Chabha Azouani, Ali Bouyahia, Johanne Germain, Kelly Martineau, Sonia Djobeir, Hugo Dary, Ludovic Fillon, Takoua Kaaouana, Alexandre Routier, Mathieu Dubois.

Project acronym: CATI

Project title: Centre d'Acquisition et de Traitement des Images

Funded in 2011

Amount: 9M€

Coordinator: Jean-François Mangin

Other partners: Neurospin, CENIR, Inserm U678, IM2A

Abstract: The CATI project (funded by the National Alzheimer Plan for 9M€, 2.1M€ for ARAMIS) aims at creating a national platform for multicenter neuroimaging studies. CATI aims to be a national resource for the scientific, medical and industrial research community and will provide a wide range of services: access to a national acquisition network, standardization of acquisitions, image quality control, image analysis, databasing/archiving, meta-analyses. Through CATI, our team coordinates a large network composed of over 30 image acquisition centers. CATI already supports over 15 multicenter projects including the national cohort MEMENTO (2300 subjects). CATI is integrated with France Life Imaging (PI: F. Lethimonnier) and the Neugrid for you (N4U, PI: G. Frisoni) network.

9.1.4. National Networks

- GdR Statistics and Medicine - http://gdr.statsante.fr/Accueil.html
- GdR (MaDICS) Masses de Données, Informations et Connaissances en Sciences Big Data - Data ScienceStatistics and Medicine - http://www.madics.fr/reseaux/

9.1.5. Other National Programs

9.1.5.1. Programme Hospitalier de Recherche Clinique (PHRC)

Participants: Olivier Colliot, Marie Chupin, Stanley Durrleman, Didier Dormont, Damien Galanaud.

- PHRC PredictPGRN, co-funding by Alzheimer Plan, *Caractérisation multimodale prospective de la démence frontotemporale dûe à des mutations du gène PGRN à un stade symptomatique et présymptomatique*. (Coordinator : A. Brice)
- PHRC ImaBio3, co-funding by Roche (pharmaceutical industry), *Rôle des réactions cellulaires sanguines, inflammatoires et immunitaires anti-amyloïde centrales et périphériques dans la maladie d'Alzheimer débutante*. (Coordinator : M. Sarazin)
- PHRC CAPP, *Caractérisation linguistique, anatomique/métabolique et biologique des différentes formes d'aphasie primaire progressive : vers le rationnel pour des essais pharmacologiques et des rééducations du langage ciblées*. (Coordinator: M. Teichmann)

9.1.5.2. Institut Universitaire d'Ingénierie pour la Santé (IUIS)

Participants: Mario Chavez, Xavier Navarro.

Project acronym: DYSPEV

Project title: Dépistage de la dyspnée par potentiels évoqués visuels

Funded in 2014

Amount: 38K€

Coordinator: Thomas Similowski

Other partners: UPMC, Inserm UMR 1158

Abstract: Steady state visual evoked potentials (SSVEP) have been widely utilized in brain computer interfacing (BCI) in last years. In this project, we explore the possibilities of SSVEP to manage the communication between patients suffering from respiratory disorders and health care providers. By imposing different breathing constraints, we use a SSVEP-based brain computer interface to help those subjects to communicate their breathing sensations (breathing well/breathing bad).

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. 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 (for this phase): 2016-2018

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.

9.2.1.3. 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

F. De Vico Fallani has a collaboration with the University Penn, Philadelphia, US (Prof. Danielle Bassett).

S. Durrleman has an enduring collaboration with professor Guido Gerig, Tandon School of Engineering, NYU. He is consultant for NIH Grant "4D shape analysis for modeling spatiotemporal change trajectories in Huntington's Disease "predict-HD".

M. Chupin and O. Colliot have an enduring collaboration with the Center for Magnetic Resonance Research, University of Minnesota, USA (P-F Van de Moortele, T. Henry, M. Marjanska, K. Ugurbil) a leading center in 7T MRI.

S. Durrleman and O. Colliot have a collaboration with the Center for Medical Image Computing (CMIC) at University College London (UCL), London, UK (S. Ourselin, D. Alexander, M. Modat).

D. Galanaud has an enduring collaboration with the Massachusetts General Hospital, Harvard University, USA (R. Gupta).

M. Chavez has different collaborations with the Mathematics Departement of the Queen Mary University of London, UK (Prof. V. Latora); and the Physics Department of the Universitat de Barcelona, Spain (Prof. Albert Diaz-Guilera)

F. De Vico Fallani has an enduring collaboration with the University Sapienza, Rome, Italy (Profs. Fabio and Claudio Babiloni) and with the IRCCS Fondazione Santa Lucia, Rome, Italy (M. Molinari and D. Mattia).

A. Bertrand has an enduring collaboration with professor Youssef Z. Wadghiri, head of the Preclinical Imaging Core, Center for Biomedical Imaging, NYU School of Medicine, New York, NY, USA.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

9.4.1.1. Internships

Kuldeep Kumar (Ecole de Technologie Supérieure, Montréal, Canada) is visiting ARAMIS from October 2016 to March 2017 under the MITACS programme.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Selection

10.1.1.1. Member of the Conference Program Committees

O. Colliot was a member of program committee of the Workshop on Patch-based Techniques in Medical Imaging (Patch-MI) held in conjunction with the MICCAI conference.

S. Durrleman was on the advisory panel of MICCAI Workshop on Spectral and Shape Analysis in Medical Imaging (SESAMI)

F. De Vico Fallani was member of the program committee of the Satellite on Brain networks, International Conference on Network Science (NetSci), Seoul, South Korea, 2016

F. De Vico Fallani was member of the program committee of5th International Workshop on Complex Networks and their Applications, Milan, Italy, 2016

10.1.1.2. Reviewer

O. Colliot acted as a reviewer for the annual meeting of the Organization for Human Brain Mapping (OHBM).

S. Durrleman acted as a reviewer for Computer Vision and Pattern Recognition (CVPR), International Conference on Computer Vision (ICCV), and Workshop on Biomedical Image Registration (WBIR).

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

O. Colliot is a member of the Editorial Board of Medical Image Analysis (Elsevier).

S. Durrleman is associate editor of IEEE Transactions on Medical Imaging (TMI)

10.1.2.2. Reviewer - Reviewing Activities

O. Colliot acted as a reviewer for NeuroImage, NeuroImage: Clinical, IEEE Trans Medical Imaging, Medical Image Analysis and Neurobiology of Aging.

S. Durrleman acted as a reviewer for NeuroImage, IEEE Trans Medical Imaging, Medical Image Analysis, Frontiers in Neuroimaging, International Journal of Computer Assisted Radiology and Surgery (IJCARS), Advances in Data Analysis and Classification, among others.

A. Bertrand acted as a reviewer for Neurobiology og Aging, Frontiers in Neuroscience, American Journal of Neuroradiology, Journal of Neuroradiology.

F. De Vico Fallani acted as a reviewer for Brain, Cerebral Cortex, IEEE TBME/TNRSE, Human Brain Mapping, Neuroimage, Plos Computational Biology, J Neurosci Meth, Sci Rep, Brain Connectivity.

10.1.3. Invited Talks

S. Durrleman gave an invited lecture at the International Colloquium "Evolution du cerveau et des capacités cognitives des Hominidés fossiles depuis Sahelanthropus tchadensis, il y a sept millions d'années jusqu'à l'Homme moderne" in Tautavel.

F. De Vico Fallani gave an invited talk at the Workshop on Complex networks, Lipari, Italy, 2016

F. De Vico Fallani gave an invited talk Meeting on Dynamics and synchronization on complex networks, Tarragona, Spain, 2016

F. De Vico Fallani gave an invited talk Workshop on Dynamic networks, Institut Systèmes Complexes, Toulouse, France, 2016

F. De Vico Fallani gave an invited talk Satellite on Brain networks, International Conference on Network Science (NetSci), Seoul, South Korea, 2016

10.1.4. Scientific Expertise

S. Durrleman has served in the “Commission de développement technologique” (CDT) of the Inria Paris center.

S. Durrleman has led a working group on neuroinformatics at the ICM (Brain and Spine Institute).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Olivier Colliot coordinates the module "Méthodes d'imagerie médicale" of the Master 2 in Computer Science of Université Pierre et Marie Curie.

Master: Olivier Colliot, Master in Computer Science, 4.5 hours (eqTD), Université Pierre et Marie Curie

Master: Stanley Durrleman, Master in Computer Science, 9 hours (eqTD), Université Pierre et Marie Curie

Master: Marie Chupin, Master in Computer Science, 3 hours (eqTD), Université Pierre et Marie Curie

Master: Fabrizio De Vico Fallani, Master in "Méthodologies et applications en imagerie médicale", 3 hours (eqTD), Université Pierre et Marie Curie

Master: Damien Galanaud, Master in Medical Physics, 4 hours, Université Paris-Sud

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", Université Pierre et Marie Curie

Medical school: Anne Bertrand gives lectures in Neuroimaging of degenerative diseases and normal aging for residents in Radiology and Neurology, for Radiology technicians, for License students in Orthophony, and in various "University Diploma" medical programs (Neurogeriatrics, Neuroradiology, Alzheimer’s Disease and related disorders, Neurovascular Imaging, Emergency-Stroke, Neuroresuscitation), for a total of 50 hours a year.

Medical school: Didier Dormont, Courses for Medical Students, Université Pierre et Marie Curie

Medical school: Dominique Hasboun, Courses for Medical Students, Université Pierre et Marie Curie

Medical school: Damien Galanaud, Courses for Medical Students, Université Pierre et Marie Curie

Medical school: Anne Bertrand, Courses for Medical Students, Université Pierre et Marie Curie

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: Didier Dormont, Courses to the university diplomas (DU) : "Maladie d’Alzheimer", and "Imagerie Vasculaire non Invasive"

Medical school: Damien Galanaud, courses to the University Diploma (DIU) “Diagnostic and Therapeutic Neuroradiology", Université Pierre et Marie Curie

Paramedical studies: Dominique Hasboun, Psychomotricity, 50 hours, Université Pierre et Marie Curie

10.2.2. Supervision

PhD in progress : Catalina Obando-Forero, “Graph models of cortical plasticity in temporal brain networks”, Inria, started in 2015, advisor: Fabrizio De Vico Fallani

PhD in progress : Jeremy Guillon, “Méthode d’analyse multimodale de connectivités neuronales basée sur la théorie des réseaux complexes multicouches”, EDITE Université Pierre et Marie Curie, started in 2015, advisors: Fabrizio De Vico Fallani and Mario Chavez

PhD Cifre in progress : Fanny Grosselin, “Fouille des données EEG et suivi longitudinal grande échelle pour le diagnostic et la prédiction du niveau de stress chez l’homme”, EDITE Université Pierre et Marie Curie, started in 2016, advisors: Fabrizio De Vico Fallani and Mario Chavez,

PhD in progress : Junhao Wen, “Cortical morphometry for discovering new biomarkers of neurodegenerative diseases”, Université Pierre et Marie Curie, Started in 2015, advisors: Olivier Colliot and Stanley Durrleman

PhD in progress : Jorge Samper-Gonzalez, “Learning from heterogeneous data for prediction of Alzheimer’s disease”, Université Pierre et Marie Curie, Started in 2015, advisors: Olivier Colliot and Theodoros Evgeniou

PhD in progress : Alexandre Routier, “Multimodal neuroimaging for characterization of primary progressive aphasias”, Université Pierre et Marie Curie, Started in 2015, advisors: Marc Teichmann, Olivier Colliot and Marie-Odile Habert

PhD in progress: Jean-Baptiste Schiratti, “Méthodes et algorithmes pour l’analyse statistique de données anatomiques longitudinales – application à la caractérisation des phases pré-symptomatiques des maladies neurodégénératives”, Ecole Polytechnique, Started in 2013, advisors: S. Allassonnière and S. Durrleman

PhD in progress: Barbara Gris, “Approche modulaire des méthodes de grandes déformations pour l’appariement de formes”, Ecole Normale Supérieure de Cachan, Started 2013, advisors: A. Trouvé and S. Durrleman

PhD in progress: Pascal Lu, “Machine learning from multimodal genetic and neuroimaging data for personalized medicine”, Université Pierre et Marie Curie, Started 2016, advisor: O. Colliot

PhD in progress: Wen Wei, “Learning brain alterations in multiple sclerosis from multimodal neuroimaging data”, Université de Nice Sophia-Antipolis, Started 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 2016, advisors: O. Colliot and S. Durrleman

PhD in progress: Manon Ansart, “Automatic recommendation systems built on the statistical exploitation of longitudinal medical data sets”, started 2016, advisors: D. Dormont and S. Durrleman

PhD in progress: Maxime Louis, “Learning spatiotemporal trajectories of iconic-geometric data sets”, started 2016, advisors: S. Durrleman

PhD in progress: Igor Koval, “Construction of disease progression models from multimodal longitudinal data”, started 2016, advisors: S. Allassonnière and S. Durrleman

PhD in progress: Lou Albessard, “Etude de la co-variation morphologique entre le crâne et le cerveau dans le genre Homo”, started 2015, advisors: D. Grimaud-Hervé and S. Durrleman

Master 2: Alexandre Morin, Master in Neuroscience, Université Pierre et Marie Curie, Oct 2015-Aug 2016, advisor: Olivier Colliot

Master 2: Thomas Jacquemont, Master in Neuroscience, Université Pierre et Marie Curie, Oct 2015-Aug 2016, advisor: Olivier Colliot

Master 2: Martina Sundqvist, Master in Cognitive Science, Ecole Normale Supérieure, Oct 2015-Aug 2016, advisors: Olivier Colliot and Marc Teichmann

Master 2: Enrico Valenti, Master in Psychiatry, Université Sapienza, Rome, Italy, Sep 2016-Dec 2016, advisor: Fabrizio De Vico Fallani

Master 2: Carlos Tor Diez, Master in BioMedical Engineering, ParisTech Universite Paris Descartes, Mar-Sept 2016, advisor: Marie Chupin

Internship: Ayoub Louati, Tunisia Polytechnic School, Mar-Sept 2016, advisor: Marie Chupin

10.2.3. Juries

Fabrizio De Vico Fallani participated, as referee, to the PhD committee of Aziz Adebimpe (Université Picardie), 2016 (supervisors: Fabrice Wallois and Ardalan Aarabi).

Mario Chavez participated, as referee, to the PhD committee of Aziz Adebimpe (Université Picardie), 2016 (supervisors: Fabrice Wallois and Ardalan Aarabi).

Olivier Colliot participated, as referee, to the PhD committee of Mehdi Hadj-Hamou (Inria Sophia), 2016 (supervisors: Xavier Pennec and Nicholas Ayache).

Olivier Colliot participated, as examiner, to the PhD committee of Bishesh Kanal (Inria Sophia), 2016 (supervisors: Xavier Pennec and Nicholas Ayache).

Olivier Colliot participated, as examiner, to the PhD committee of Baptiste Morel (Telecom ParisTech), 2016 (supervisors: Isabelle Bloch and Catherine Adamsbaum).

Olivier Colliot participated, as examiner, to the PhD committee of Romain Colle (Université Paris-Sud), 2016 (supervisor: Emmanuelle Corruble).

10.3. Popularization

With the precious help of the communication department of the Inria Paris Center, ARAMIS prepared and presented games on brain data analysis, presented at the "Salon Culture et Jeux Mathématiques" and at the "Fête de la Science".

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] P. GORI.*Statistical models to learn the structural organisation of neural circuits from multimodal brain images - Application to Gilles de la Tourette syndrome*, UPMC - Université Paris 6 Pierre et Marie Curie, January 2016, https://hal.archives-ouvertes.fr/tel-01352477.

Articles in International Peer-Reviewed Journal

[2] A. ALGHAMDI, R. PALICH, R. CALIN, C. HUSSENET, M. PAPO, D. BOUTOLLEAU, A. MATHIAN, D. GALANAUD, P. LE HOANG, E. CAUMES, C. KATLAMA, B. BODAGHI, V. TOUITOU, V. POURCHER-MARTINEZ.*Atypical ocular manifestation of primary varicella zoster virus infection as the first manifestation of AIDS*, in "AIDS", 2016, vol. 30, n^{o} 4, p. 674 - 676 [*DOI :* 10.1097/QAD.0000000000000981], https://hal.inria.fr/hal-01377873.

[3] F. AMEUR, O. COLLIOT, P. CAROPPO, S. STROER, D. DORMONT, A. BRICE, C. AZUAR, B. DUBOIS, I. LE BER, A. BERTRAND.*White matter lesions in FTLD: distinct phenotypes characterize GRN and C9ORF72 mutations*, in "Neurology: Genetics", 2016, vol. 2, n^{o} 1 [*DOI :* 10.1212/NXG.0000000000000047], https://hal.inria.fr/hal-01266596.

[4] J.-L. BACQUET, N. WEISS, C. MEYNIEL, C. ALGRIN, S. CHOQUET, V. LEBLOND, D. GALANAUD, P. LEHOANG, B. BODAGHI, V. TOUITOU.*Neither the patient nor the physician could see anything: Atypical Bing-Neel syndrome*, in "American Journal of Hematology", 2016, vol. 91, n^{o} 8, p. 858 - 859, https://hal.inria.fr/hal-01377870.

[5] A. BEAUDET, J. DUMONCEL, F. DE BEER, B. DUPLOYER, S. DURRLEMAN, E. GILISSEN, J. HOFFMAN, C. TENAILLEAU, J. F. THACKERAY, J. BRAGA.*Morphoarchitectural variation in South African fossil cercopithecoid endocasts*, in "Journal of Human Evolution", September 2016, vol. 101, p. 65-78 [*DOI :* 10.1016/J.JHEVOL.2016.09.003], https://hal.inria.fr/hal-01440065.

[6] A. L. CALDERON-GARCIDUEÑAS, A. IDBAIH, D. GALANAUD, C. DUYCKAERTS, F. BIELLE.*Dural and osteolytic sarcomatoid relapse of a secondary gliosarcoma with tryptase immunoreactivity*, in "Clinical Neuropathology", 2016, vol. 35, n^{o} 05, p. 154 - 158, https://hal.inria.fr/hal-01377874.

[7] E. CAVEDO, B. DUBOIS, O. COLLIOT, S. LISTA, B. CROISILE, G. L. TISSERAND, J. TOUCHON, A. BONAFE, P. J. OUSSET, O. ROUAUD, F. RICOLFI, A. VIGHETTO, F. PASQUIER, S. GALLUZZI, C. DELMAIRE, M. CECCALDI, N. GIRARD, S. LEHERICY, F. DUVEAU, M. CHUPIN, M. SARAZIN, D. DORMONT, H. HAMPEL.*Reduced Regional Cortical Thickness Rate of Change in Donepezil-Treated Subjects With Suspected Prodromal Alzheimer's Disease*, in "Journal of Clinical Psychiatry", October 2016 [*DOI :* 10.4088/JCP.15M10413], https://hal.inria.fr/hal-01388951.

[8] F. COHEN AUBART, J.-F. EMILE, P. MAKSUD, D. GALANAUD, P. CLUZEL, N. BENAMEUR, O. AUMAITRE, Z. AMOURA, J. HAROCHE.*Efficacy of the MEK inhibitor cobimetinib for wild-type BRAF Erdheim-Chester disease*, in "British Journal of Haematology", October 2016, https://hal.inria.fr/hal-01377866.

[9] R. COLLE, C. CURY, M. CHUPIN, E. DEFLESSELLE, P. HARDY, N. GHAIDAA, B. FALISSARD, D. DUCREUX, O. COLLIOT, E. CORRUBLE.*Hippocampal volume predicts antidepressant efficacy*

in depressed patients without incomplete hippocampal inversion, in "Neuroimage-Clinical", 2016 [*DOI :* 10.1016/J.NICL.2016.04.009], http://hal.upmc.fr/hal-01310464.

[10] R. COLLE, I. DUPONG, O. COLLIOT, E. DEFLESSELLE, P. HARDY, B. FALISSARD, D. DUCREUX, M. CHUPIN, E. CORRUBLE.*Smaller hippocampal volumes predict lower antidepressant response/remission rates in depressed patients: A meta-analysis*, in "World Journal of Biological Psychiatry", July 2016, p. 1-23, https://hal.inria.fr/hal-01372398.

[11] F. DE VICO FALLANI, S. CLAUSI, M. LEGGIO, M. CHAVEZ, M. VALENCIA, A. G. MAGLIONE, F. BABILONI, F. CINCOTTI, D. MATTIA, M. MOLINARI.*Interhemispheric Connectivity Characterizes Cortical Reorganization in Motor-Related Networks After Cerebellar Lesions*, in "Cerebellum (London, England)", July 2016, https://hal.inria.fr/hal-01427519.

[12] F. DE VICO FALLANI, V. LATORA, M. CHAVEZ.*A Topological Criterion for Filtering Information in Complex Brain Networks*, in "PLoS Computational Biology", January 2017, vol. 13, n^{o} 1 [*DOI :* 10.1371/JOURNAL.PCBI.1005305], https://hal.inria.fr/hal-01443254.

[13] A. DESBOIS, O. ADDIMANDA, A. BERTRAND, A. DEROUX, L. PÉRARD, R. DEPAZ, E. HACHULLA, M. LAMBERT, D. LAUNAY, B. SUBRAN, F. ACKERMAN, X. MARIE, F. COHEN, I. MARIE, C. SALVARINI, P. CACOUB, D. SAADOUN.*Efficacy of Anti-TNFα in Severe and Refractory Neuro-Behcet Disease*, in "Medicine", 2016, vol. 95, n^{o} 23, https://hal.inria.fr/hal-01377880.

[14] P. GORI, O. COLLIOT, L. MARRAKCHI-KACEM, Y. WORBE, F. DE VICO FALLANI, M. CHAVEZ, C. POUPON, A. HARTMANN, N. AYACHE, S. DURRLEMAN.*Parsimonious Approximation of Streamline Trajectories in White Matter Fiber Bundles*, in "IEEE Transactions on Medical Imaging", 2016, vol. PP, n^{o} 99 [*DOI :* 10.1109/TMI.2016.2591080], https://hal.archives-ouvertes.fr/hal-01346067.

[15] P. GORI, O. COLLIOT, L. MARRAKCHI-KACEM, Y. WORBE, C. POUPON, A. HARTMANN, N. AYACHE, S. DURRLEMAN.*A Bayesian Framework for Joint Morphometry of Surface and Curve meshes in Multi-Object Complexes*, in "Medical Image Analysis", January 2017, vol. 35, p. 458-474 [*DOI :* 10.1016/J.MEDIA.2016.08.011], https://hal.inria.fr/hal-01359423.

[16] L. HAMELIN, J. LAGARDE, G. DOROTHÉE, C. LEROY, M. LABIT, R. A. COMLEY, L. C. DE SOUZA, H. CORNE, L. DAUPHINOT, M. BERTOUX, B. DUBOIS, P. GERVAIS, O. COLLIOT, M. C. POTIER, M. BOTTLAENDER, M. SARAZIN.*Early and protective microglial activation in Alzheimer's disease: a prospective study using 18F-DPA-714 PET imaging*, in "Brain - A Journal of Neurology ", March 2016, vol. 139, n^{o} Pt 4, p. 1252-64, https://hal.inria.fr/hal-01372414.

[17] A. HARTMANN, J. MÜLLNER, N. MEIER, H. HESEKAMP, P. VAN MEERBEECK, M.-O. HABERT, A. KAS, M. TANGUY, M. MAZMANIAN, H. OYA, N. ABUAF, H. GAOUAR, S. SALHI, F. CHARBONNIER-BEAUPEL, M. FIEVET, D. GALANAUD, S. ARGUILLERE, E. ROZE, B. DEGOS, D. GRABLI, L. LACOMBLEZ, C. HUBSCH, M. VIDAILHET, A.-M. BONNET, J.-C. CORVOL, M. SCHÜPBACH, M. SHARMA.*Bee Venom for the Treatment of Parkinson Disease – A Randomized Controlled Clinical Trial*, in "PLoS ONE", July 2016, vol. 11, n^{o} 7, https://hal.inria.fr/hal-01377867.

[18] C. ISABEL, F. COHEN AUBART, P. DODET, D. GALANAUD, Z. AMOURA, E. MAILLART.*Spinal Koebner phenomenon: Medullar sarcoidosis facing a discal hernia*, in "Joint Bone Spine", 2016, https://hal.inria.fr/hal-01377868.

[19] T. KAAOUANA, A. BERTRAND, F. OUAMER, B. LAW-YE, N. PYATIGORSKAYA, A. BOUYAHIA, N. THIERY, C. DUFOUIL, C. DELMAIRE, D. DORMONT, L. DE ROCHEFORT, M. CHUPIN.*Improved cerebral microbleeds detection using their magnetic signature on T2*-phase-contrast: a comparison study in a clinical setting*, in "Neuroimage-Clinical", 2016 [*DOI :* 10.1016/J.NICL.2016.08.005], https://hal.inria.fr/hal-01406920.

[20] B. LAW-YE, B. GEERTS, D. GALANAUD, D. DORMONT, N. PYATIGORSKAYA.*Pseudo-asymmetry of cerebral blood flow in arterial spin labeling caused by unilateral fetal-type circle of Willis: Technical limitation or a way to better understanding physiological variations of cerebral perfusion and improving arterial spin labeling acquisition?*, in "Journal of Cerebral Blood Flow and Metabolism", September 2016, vol. 36, n^{o} 9, p. 1641 - 1643, https://hal.inria.fr/hal-01377864.

[21] E. MAILLART, C. PAPEIX, C. MELLERIO, A. BERTRAND, C. LUBETZKI, C. LOUAPRE.*Extensive and severe CNS demyelination associated with golimumab therapy*, in "Journal of Neurology", 2016, vol. 263, n^{o} 9, p. 1869 - 1871, https://hal.inria.fr/hal-01377879.

[22] L. MARRAKCHI-KACEM, A. VIGNAUD, J. SEIN, J. GERMAIN, T. R. HENRY, C. POUPON, L. HERTZ-PANNIER, S. LEHÉRICY, O. COLLIOT, P.-F. VAN DE MOORTELE, M. CHUPIN.*Robust imaging of hippocampal inner structure at 7T: in vivo acquisition protocol and methodological choices*, in "Magnetic Resonance Materials in Physics, Biology and Medicine", May 2016, https://hal.inria.fr/hal-01321870.

[23] G. MARZOLF, M. SABOU, B. LANNES, F. COTTON, D. MEYRONET, D. GALANAUD, J.-P. COTTIER, S. GRAND, H. DESAL, J. KREUTZ, M. SCHENCK, N. MEYER, F. SCHNEIDER, J.-L. DIETEMANN, M. KOOB, R. HERBRECHT, S. KREMER, H. ZHANG.*Magnetic Resonance Imaging of Cerebral Aspergillosis: Imaging and Pathological Correlations*, in "PLoS ONE", April 2016, vol. 11, n^{o} 4, https://hal.inria.fr/hal-01377871.

[24] A. MOLKA, G. LEROUX, A. BERTRAND, S. JAUREGUIBERRY, D. SAADOUN, D. MAZIER, E. CAUMES, P. CACOUB.*Brain lesions*, in "Revue de Médecine Interne", 2016 [*DOI :* 10.1016/J.REVMED.2016.04.010], https://hal.inria.fr/hal-01406578.

[25] X. NAVARRO-SUNE, A. HUDSON, F. DE VICO FALLANI, J. MARTINERIE, A. WITON, P. POUGET, M. RAUX, T. SIMILOWSKI, M. CHAVEZ.*Riemannian geometry applied to detection of respiratory states from EEG signals: the basis for a brain-ventilator interface*, in "IEEE Transactions on Biomedical Engineering", 2016, 14, 14 pages, 7 figures, https://hal.inria.fr/hal-01427496.

[26] M. OLLIVIER, A. BERTRAND, F. CLARENÇON, S. GERBER, S. DELTOUR, F. DOMONT, S. TRUNET, D. DORMONT, D. LECLERCQ.*Neuroimaging features in Posterior Reversible Encephalopathy Syndrome: A Pictorial Review*, in "Journal of the Neurological Sciences", December 2016 [*DOI :* 10.1016/J.JNS.2016.12.007], https://hal.inria.fr/hal-01427526.

[27] G. OPERTO, M. CHUPIN, B. BATRANCOURT, M.-O. HABERT, O. COLLIOT, H. BENALI, C. POUPON, C. CHAMPSEIX, C. DELMAIRE, S. MARIE, D. RIVIÈRE, M. PÉLÉGRINI-ISSAC, V. PERLBARG, R. TREBOSSEN, M. BOTTLAENDER, V. FROUIN, A. GRIGIS, D. P. ORFANOS, H. DARY, L. FILLON, C. AZOUANI, A. BOUYAHIA, C. FISCHER, L. EDWARD, M. BOUIN, U. THOPRAKARN, J. LI, L. MAKKAOUI, S. PORET, C. DUFOUIL, V. BOUTELOUP, G. CHÉTELAT, B. DUBOIS, S. LEHÉRICY, J.-F. MANGIN, Y. COINTEPAS.*CATI: A Large Distributed Infrastructure for the Neuroimaging of Cohorts*, in "Neuroinformatics", April 2016, p. 1-12, The CATI Consortium [*DOI :* 10.1007/S12021-016-9295-8], http://hal.upmc.fr/hal-01303094.

[28] F. SEDEL, B. CHABROL, B. AUDOIN, E. KAPHAN, C. TRANCHANT, T. BURZYKOWSKI, A. TOURBAH, M. T. VANIER, D. GALANAUD.*Normalisation of brain spectroscopy findings in Niemann–Pick disease type C patients treated with miglustat*, in "Journal of Neurology", 2016, vol. 263, n^{o} 5, p. 927 - 936, https://hal.inria.fr/hal-01377872.

[29] S. SENOVA, M. AGGAD, J.-L. GOLMARD, D. HASBOUN, I. LAMPROGLOU, C. JENNY, P. CORNU, J.-J. MAZERON, C. A. VALÉRY.*Predictors of Trigeminal Neuropathy After Radiosurgery for Vestibular Schwannomas*, in "International Journal of Radiation Oncology Biology Physics", 2016, vol. 95, n^{o} 2, p. 721 - 728, https://hal.inria.fr/hal-01377875.

[30] M. TEICHMANN, C. LESOIL, J. GODARD, M. VERNET, A. BERTRAND, R. LEVY, B. DUBOIS, L. LEMOINE, D. Q. TRUONG, M. BIKSON, A. KAS, A. VALERO-CABRÉ.*Direct current stimulation over the anterior temporal areas boosts semantic processing in primary progressive aphasia*, in "Annals of Neurology", 2016, https://hal.inria.fr/hal-01377876.

[31] S. J. TEIPEL, E. CAVEDO, M. J. GROTHE, S. LISTA, S. GALLUZZI, O. COLLIOT, M. CHUPIN, H. BAKARDJIAN, D. DORMONT, B. DUBOIS, H. HAMPEL.*Predictors of cognitive decline and treatment response in a clinical trial on suspected prodromal Alzheimer's disease*, in "Neuropharmacology", August 2016, vol. 108, p. 128-35, https://hal.inria.fr/hal-01372417.

International Conferences with Proceedings

[32] A. BERTRAND, S. EPELBAUM, M.-O. HABERT, H. HAMPEL, H. BAKARDJIAN, M. CHUPIN, C. AZOUANI, K. MARTINEAU, L. FILLON, C. FISCHER, O. COLLIOT, B. DUBOIS.*THE INSIGHT COHORT: BASELINE ANALYSIS OF STRUCTURAL MR IMAGING IN ASYMPTOMATIC SUBJECTS AT RISK FOR ALZHEIMER'S DISEASE*, in "AAIC", Toronto, Canada, July 2016, vol. 12, n^{o} 7 [*DOI :* 10.1016/J.JALZ.2016.06.612], https://hal.inria.fr/hal-01439571.

[33] S. BOURRELIER, M.-O. HABERT, B. LAW-YE, M. OLLIVIER, N. YENI, P.-M. DAVID, A. KAS, D. DORMONT, A. BERTRAND.*Imagerie hybride TEP-IRM pour l'exploration des troubles cognitifs*, in "Journées Françaises de Radiologie", Paris, France, October 2016, https://hal.inria.fr/hal-01427376.

[34] C. CURY, M. M. LORENZI, D. M. CASH, J. M. NICHOLAS, A. M. ROUTIER, J. ROHRER, S. M. OURSELIN, S. M. DURRLEMAN, M. M. MODAT.*Spatio-Temporal Shape Analysis of Cross-Sectional Data for Detection of Early Changes in Neurodegenerative Disease*, in "Spectral and Shape Analysis in Medical Imaging", Athens, Greece, M. REUTER, C. WACHINGER, H. LOMBAERT (editors), Spectral and Shape Analysis in Medical Imaging, Springer, September 2016, vol. 10126, p. 63 - 75 [*DOI :* 10.1007/978-3-319-51237-2_6], https://hal.inria.fr/hal-01440061.

[35] J. DUMONCEL, G. SUBSOL, S. DURRLEMAN, J. P. JESSEL, A. BEAUDET, J. BRAGA.*How to build an average model when samples are variably incomplete? Application to fossil data*, in "WBIR: Workshop on Biomedical Image Registration", Las Vegas, United States, July 2016, p. 541-548, https://hal-lirmm.ccsd.cnrs.fr/lirmm-01381310.

[36] M. HABERT, M. DAVID, A. BERTRAND, G. BERA, N. YENI, M. BERTAUX, N. PYATIGORSKAYA, D. DORMONT, A. VILLANUEVA, A. KAS.*Imagerie hybride TEP/IRM dans le bilan des pathologies neurodégénératives : retour d'expérience*, in "2es Journées Francophones de Médecine Nucléaire", Grenoble, France, May 2016, vol. 40, n^{o} 3, p. 169 - 170 [*DOI :* 10.1016/J.MEDNUC.2016.03.003], https://hal.inria.fr/hal-01427374.

Other Publications

[37] B. GRIS, S. DURRLEMAN, A. TROUVÉ.*A sub-Riemannian modular framework for diffeomorphism based analysis of shape ensembles*, May 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01321142.

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:

4. - Security and privacy
4.3. - Cryptography
4.3.1. - Public key cryptography
4.3.3. - Cryptographic protocols
4.8. - Privacy-enhancing technologies
7. - Fundamental Algorithmics
7.7. - Number theory

Other Research Topics and Application Domains:

6.4. - Internet of things
9.4.1. - Computer science
9.8. - Privacy

1. Members

Research Scientists

David Pointcheval [Team leader, CNRS, Senior Researcher, HDR]
Michel Ferreira Abdalla [CNRS, Senior Researcher, HDR]
Georg Fuchsbauer [Inria, Researcher]
Hoeteck Wee [CNRS, Researcher, HDR]

Faculty Members

Jacques Stern [ENS Paris, Emeritus Professor, HDR]
Damien Vergnaud [ENS Paris, Associate Professor, HDR]

PhD Students

Fabrice Ben Hamouda–Guichoux [ENS Paris, Fondation CFM, until August 2016]
Raphael Bost [DGA, Associated member]
Florian Bourse [CNRS, ERC CryptoCloud]
Jérémy Chotard [CNRS, ERC CryptoCloud, from October 2016]
Mario Cornejo Ramirez [Inria, CORDI-S, until September 2016]
Geoffroy Couteau [CNRS, ERC CryptoCloud]
Rafael Del Pino [Inria, FUI CryptoComp, until September 2016]
Aurélien Dupin [Thales, Associated member]
Pierre-Alain Dupont [DGA]
Romain Gay [ENS Paris]
Dahmun Gourdazi [CryptoExperts, CIFRE]
Louiza Khati [Oppida]
Pierrick Meaux [Inria, H2020 ICT SAFECrypto]
Thierry Mefenza Nountu [ENS Paris, ANR JCJC ROMAnTIC]
Michele Minelli [ENS Paris, H2020 ITN ECRYPT-NET]
Anca Nitulescu [CNRS, ERC CryptoCloud]
Michele Orrù [CNRS, ERC aSCEND, from November 2016]
Alain Passelegue [ENS Paris, DGA & ANR Prince]

Razvan Rosie [ENS Paris, H2020 ITN ECRYPT-NET]
Quentin Santos [Orange Labs, CIFRE]
Adrian Thillard [ANSSI]

Post-Doctoral Fellows

Pooya Farshim [ENS Paris, ANR EnBiD, from February 2016]
Julia Hesse [CNRS, ERC CryptoCloud, from October 2016]

Administrative Assistants

Nathalie Gaudechoux [Inria]
Joëlle Isnard [CNRS, Administrative Head DI/ENS]

Others

Baptiste Louf [Ecole Polytechnique, Internship, from March to July 2016]
Balthazar Bauer [ENS Lyon, Internship, from September 2016]

2. Overall Objectives

2.1. Presentation

Cryptographic algorithms are the equivalent of locks, seals, security stamps and identification documents on the Internet. They are essential to protect our on-line bank transactions, credit cards, medical and personal information and to support e-commerce and e-government. They come in different flavors. Encryption algorithms are essential 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) replace hand-written signatures in electronic transactions. A similar role can be played by MAC algorithms. Identification protocols allow to securely verify the identity of the party at the other end of the line. Therefore, cryptology is a research area with a high strategic impact for industries, individuals, and for the society as a whole. The research activity of the project-team CASCADE addresses the following topics, which cover almost all the domains that are currently active in the international cryptographic community, but mainly in the public-key area:

1. Implementation of cryptographic and applied cryptography
2. Design and provable security
3. Theoretical and concrete 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 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 into 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 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 to get provable security, without such ideal assumptions, under new and possibly stronger computational assumptions. As a consequence, a cryptographer has to deal with the four following important steps, which are **all** our main goals:

computational assumptions, which are the foundations 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, with additional features, etc.

security proof, which consists in exhibiting a reduction.

3. Research Program

3.1. Randomness in Cryptography

Randomness is a key ingredient for cryptography. Random bits are necessary not only for generating cryptographic keys, but are also often an important part of cryptographic algorithms. In some cases, probabilistic protocols make it possible to perform tasks that are impossible deterministically. In other cases, probabilistic algorithms are faster, more space efficient or simpler than known deterministic algorithms. Cryptographers usually assume that parties have access to perfect randomness but in practice this assumption is often violated and a large body of research is concerned with obtaining such a sequence of random or pseudorandom bits.

One of the project-team research goals is to get a better understanding of the interplay between randomness and cryptography and to study the security of various cryptographic protocols at different levels (information-theoretic and computational security, number-theoretic assumptions, design and provable security of new and existing constructions).

Cryptographic literature usually pays no attention to the fact that in practice randomness is quite difficult to generate and that it should be considered as a resource like space and time. Moreover since the perfect randomness abstraction is not physically realizable, it is interesting to determine whether imperfect randomness is “good enough” for certain cryptographic algorithms and to design algorithms that are robust with respect to deviations of the random sources from true randomness.

The power of randomness in computation is a central problem in complexity theory and in cryptography. Cryptographers should definitely take these considerations into account when proposing new cryptographic schemes: there exist computational tasks that we only know how to perform efficiently using randomness but conversely it is sometimes possible to remove randomness from probabilistic algorithms to obtain efficient deterministic counterparts. Since these constructions may hinder the security of cryptographic schemes, it is of high interest to study the efficiency/security tradeoff provided by randomness in cryptography.

Quite often in practice, the random bits in cryptographic protocols are generated by a pseudorandom number generation process. When this is done, the security of the scheme of course depends in a crucial way on the quality of the random bits produced by the generator. Despite the importance, many protocols used in practice often leave unspecified what pseudorandom number generation to use. It is well-known that pseudorandom generators exist if and only if one-way functions exist and there exist efficient constructions based on various number-theoretic assumptions. Unfortunately, these constructions are too inefficient and many protocols used in practice rely on “ad-hoc” constructions. It is therefore interesting to propose more efficient constructions, to analyze the security of existing ones and of specific cryptographic constructions that use weak pseudorandom number generators.

The project-team undertakes research in these three aspects. The approach adopted is both theoretical and practical, since we provide security results in a mathematical frameworks (information theoretic or computational) with the aim to design protocols among the most efficient known.

3.2. Lattice 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 discrete log. This is somewhat 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 —in particular since they also offer features that other alternative public-key cryptosystems (such as MQ-based, code-based or hash-based schemes) cannot provide.

At its very core, secure communication rests on two foundations: authenticity and secrecy. Authenticity assures the communicating parties that they are indeed communicating with each other and not with some potentially malicious outside party. Secrecy is necessary so that no one except the intended recipient of a message is able to deduce anything about its contents.

Lattice cryptography might find applications towards constructing practical schemes for resolving essential cryptographic problems —in particular, guaranteeing authenticity. On this front, our team is actively involved in pursuing the following two objectives:

1. Construct, implement, and standardize a practical public key digital signature scheme that is secure against quantum adversaries.
2. Construct, implement, and standardize a symmetric key authentication scheme that is secure against side channel attacks and is more efficient than the basic scheme using AES with masking.

Despite the great progress in constructing fairly practical lattice-based encryption and signature schemes, efficiency still remains a very large obstacle for advanced lattice primitives. While constructions of identity-based encryption schemes, group signature schemes, functional encryption schemes, and even fully-homomorphic encryption schemes are known, the implementations of these schemes are extremely inefficient.

Fully Homomorphic Encryption (FHE) is a very active research area. Let us just give one example illustrating the usefulness of computing on encrypted data: Consider an on-line patent database on which firms perform complex novelty queries before filing patents. With current technologies, the database owner might analyze the queries, infer the invention and apply for a patent before the genuine inventor. While such frauds were not reported so far, similar incidents happen during domain name registration. Several websites propose "registration services" preceded by "availability searches". These queries trigger the automated registration of the searched domain names which are then proposed for sale. Algorithms allowing arbitrary computations without disclosing their inputs (and/or their results) are hence of immediate usefulness.

In 2009, IBM announced the discovery of a FHE scheme by Craig Gentry. The security of this algorithm relies on worst-case problems over ideal lattices and on the hardness of the sparse subset sum problem. Gentry's construction is an ingenious combination of two ideas: a somewhat homomorphic scheme (capable of supporting many "logical or" operations but very few "ands") and a procedure that refreshes the homomorphically processed ciphertexts. Gentry's main conceptual achievement is a "bootstrapping" process in which the somewhat homomorphic scheme evaluates its own decryption circuit (self-reference) to refresh (recrypt) ciphertexts.

Unfortunately, it is safe to surmise that if the state of affairs remains as it is in the present, then despite all the theoretical efforts that went into their constructions, these schemes will never be used in practical applications.

Our team is looking at the foundations of these primitives with the hope of achieving a breakthrough that will allow them to be practical in the near future.

3.3. Security amidst Concurrency on the Internet

Cryptographic protocols that are secure when executed in isolation, can be 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 as well as 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,
2. and 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

Electronic cash (e-cash) was first proposed in the 1980s but despite extensive research it has never been deployed on a large scale. Other means of digital payments have instead largely replaced cash and other "analog" payments. Common to all digital payments offered by banks and other payment providers is that they do not respect the citizens' right to privacy, which for legitimate purchases and moderate sums also includes their right of anonymous payments.

Recently the rise of so-called decentralized currencies, such as Bitcoin and the numerous "alt-coins" inspired by it, have established a third way of payments in addition to physical cash, which offers privacy, and card and other electronic payments, which are traceable by its providers. The continuous growth of popularity and usage of this new kind of currencies, also called "cryptocurrencies" as their security and stability crucially relies on the use of cryptography, have triggered a renewed interest in cryptographic e-cash.

Our group investigates "centralized" e-cash, which respects the current economic model where money is issued by (central) banks, as opposed to cryptocurrencies, which use money distribution to incentivize widespread participation in the system, required for stability. Of particular interest among centralized e-cash schemes is transferable e-cash, which allows users to transfer coins between each other without any interaction with a third party. Currently all efficient e-cash schemes require coins to be deposited at the bank once received; they are thus not transferable. Our goal is to propose efficient transferable e-cash schemes.

Another direction concerns cryptocurrencies whose adoption is continuously growing so that now even central banks, like the Swedish *Riksbank*, are considering issuing their own currency as a cryptocurrency. While systems like Bitcoin are perceived as offering anonymous payments, a line of research has shown that this is not the case. One of the major research challenges in this area is thus to devise schemes that offer an anonymity level comparable to that of physical cash. The currently proposed schemes either lack formal security analyses or they are inefficient due to the heavy-duty cryptography used. Our group works towards practical cryptocurrencies with formally analyzed privacy guarantees.

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.

4.2. Hardware Security

Cryptography is only one component of information security, but it is a crucial component. Without cryptography, it would be impossible to establish secure communications between users over insecure networks like the Internet. In particular, public-key cryptography (invented by Diffie and Hellman in 1976) enables to establish secure communications between users who have never met physically before. One can argue that companies like E-Bay or Amazon could not exist without public-key cryptography. Since 30 years the theory of cryptography has developed considerably. However cryptography is not only a theoretical science; namely at some point the cryptographic algorithms must be implemented on physical devices, such as PCs, smart cards or RFIDs. Then problems arise: in general smart cards and RFIDs have limited computing power and leak information through power consumption and electro-magnetic radiations. Similarly a PC can be exposed to various computer viruses which can leak private informations to a remote attacker. Such information leakage can be exploited by an attacker; this is called a **side-channel attack**. It is well known that a cryptographic algorithm which is perfectly secure in theory can be completely insecure in practice if improperly implemented.

In general, countermeasures against side-channel attacks are heuristic and can only make a particular implementation resist particular attacks. Instead of relying on ad-hoc security patches, a better approach consists in working in the framework of **provable security**. The goal is to prove that a cryptosystem does not only resist specific attacks but can resist any possible side-channel attack. As already demonstrated with cryptographic protocols, this approach has the potential to significantly increase the security level of cryptographic products. Recently the cryptography research community has developed new security models to take into account these practical implementation attacks; the most promising such model is called the **leakage-resilient model**.

Therefore, our goal is to define new security models that take into account any possible side-channel attack, and then to design new cryptographic schemes and countermeasures with a proven security guarantee against side-channel attacks.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Conferences

- Hoeteck Wee is one of the invited speakers at Asiacrypt 2016.
- Michel Abdalla is one of the invited speakers at ICISC 2016.

5.1.2. Awards

Romain Gay and Hoeteck Wee, together with Dennis Hofheinz and Eike Kiltz, received the Best Paper Award at Eurocrypt 2016 .
BEST PAPERS AWARDS :
[40] **Advances in Cryptology – EUROCRYPT 2016**. R. GAY, D. HOFHEINZ, E. KILTZ, H. WEE.

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):

- More efficient constructions with lattices
- New e-cash constructions
- Advanced primitives for the privacy in the cloud
- Efficient functional encryption
- Various predicate encryption schemes

7. Partnerships and Cooperations

7.1. National Initiatives with Industrials

7.1.1. SIMPATIC

Title: SIM and PAiring Theory for Information and Communications security

Program: ANR INS

Duration: February 2013 – July 2016

Coordinator: Orange Labs

Partners:

Orange Labs

ENS

INVIA

Oberthur Technologies

STMicroelectronics

Université Bordeaux 1

Université de Caen Basse-Normandie

Université de Paris VIII

Local coordinator: David Pointcheval

We aim at providing the most possible efficient and secure hardware/software implementation of a bilinear pairing in a SIM card.

7.1.2. CryptoComp

Program: FUI

Duration: October 2014 – November 2018

Coordinator: CryptoExperts

Partners:

CEA

CNRS

Kalray

Inria

Dictao

Université de Limoges

VIACESS

Bertin technologies

GEMALTO

Local coordinator: David Pointcheval

We aim at studying delegation of computations to the cloud, in a secure way.

7.2. National Collaborations within Academics

7.2.1. ROMAnTIC

Title: Randomness in Mathematical Cryptography

Program: ANR JCJC

Duration: October 2012 – September 2016

PI: Damien Vergnaud

Partners: ENS Lyon, Université de Limoges

ANSSI

Univ. Paris 7

Univ. Limoges

The goal of this project is to get a better understanding of the interplay between randomness and cryptography and to study the security of various cryptographic protocols at different levels (information-theoretic and computational security, number-theoretic assumptions, design and provable security of new and existing constructions).

7.2.2. EnBiD

Title: Encryption for Big Data

Program: ANR JCJC

Duration: October 2014 – September 2018

PI: Hoeteck Wee

Partners:

Univ. Paris 2

Univ. 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.

7.2.3. EfTrEC

Title: Efficient Transferable E-Cash

Program: ANR JCJC

Duration: October 2016 – September 2020

PI: Georg Fuchsbauer

Partners:

Univ. 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 even resistant to attacks on quantum computers.

7.2.4. ALAMBIC

Title: AppLicAtions of MalleaBIlity in Cryptography

Program: ANR PRC

Duration: October 2016 – September 2020

PI: Damien Vergnaud

Partners:

ENS Lyon

Univ. 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. CryptoAction

Title: Cryptography for Secure Digital Interaction

Program: H2020 ICT COST

Duration: April 2014 – April 2018

Local coordinator: Michel Abdalla

The aim of this COST CryptoAction is to stimulate interaction between the different national efforts in order to develop new cryptographic solutions and to evaluate the security of deployed algorithms with applications to the secure digital interactions between citizens, companies and governments.

7.3.2. CryptoCloud

Title: Cryptography for the Cloud

Program: FP7 ERC Advanced Grant

Duration: June 2014 – May 2019

PI: David Pointcheval

The goal of the CryptoCloud project is to develop new interactive tools to provide privacy to the Cloud.

7.3.3. SAFEcrypto

Title: 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 on 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.

7.3.4. ECRYPT-NET

Title: 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)

École Normale Supérieure (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.

7.3.5. aSCEND

Title: Secure Computation on Encrypted Data

Program: H2020 ERC Starting Grant

Duration: June 2015 – May 2020

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.

7.4. International Research Visitors

- Sanjam Garg (UC Berkeley)
- Yuval Ishai (UCLA/Technion)
- Gregory Neven (IBM Zurich)
- Ryo Nishimaki (NTT)
- Claudio Orlandi (Aarhus)
- Rafael Pass (Cornell)
- Leonid Reyzin (Boston University)
- Alessandra Scafuro (postdoc, BU/NEU)
- Victor Shoup (NY University)
- Vinod Vaikuntanathan (MIT)
- Daniel Wichs (Northeastern University)

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. Events and Activities

- a regular seminar is organized: http://www.di.ens.fr/CryptoSeminaire.html

- quarterly Paris Crypto Days (https://pariscryptoday.github.io) supported by CryptoCloud and aSCEND
- working group on lattices (http://perso.ens-lyon.fr/damien.stehle/LATTICE_MEETINGS.html), joint with ENS Lyon
- 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 PAIRING: Michel Abdalla

8.1.1.3. Other Steering Committees

- steering committee of the Coding and Cryptography working group (GT-C2 - https://crypto.di.ens.fr/c2:main) of the *Groupe de Recherche Informatique Mathématique* (GDR-IM): Damien Vergnaud is the Head of this steering committee

8.1.1.4. Board of International Organisations

- Board of the *International Association for Cryptologic Research* (IACR): Michel Abdalla (2013 – 2018), David Pointcheval (2008–2016)

8.1.2. Scientific Events Selection

8.1.2.1. Program Committee Chair

- Africacrypt '16 – 13-15 April (Fes, Morocco): David Pointcheval

8.1.2.2. Program Committee Member

- Financial Crypto '16 – 22–26 February (Barbados): Damien Vergnaud
- PKC '16 – 6-9 March (Taiwan): David Pointcheval
- Africacrypt '16 – 13-15 April (Fes, Morocco): Georg Fuchsbauer
- Eurocrypt '16 – 8-12 May (Vienna, Austria): Michel Abdalla and Georg Fuchsbauer
- AsiaPKC 2016 – 30 May 30 - 03 June (Xi'an, China): Damien Vergnaud
- Crypto '16 – 14-18 August (Santa Barbara, California, USA): David Pointcheval
- ACM CCS '16 – 24-28 October (Vienna, Austria): Hoeteck Wee
- ProvSec '16 – 10-12 November (Nanjing, China): Georg Fuchsbauer
- CANS '16 – 14-16 November (Milan, Italy): Georg Fuchsbauer
- Asiacrypt '16 – 4-8 December (Hanoi, Vietnam): Georg Fuchsbauer

8.1.3. Editorial Boards of Journals

Editor-in-Chief

- of the *International Journal of Applied Cryptography (IJACT)* – Inderscience Publishers: David Pointcheval

Associate Editor

- of *IET Information Security*: Michel Abdalla
- of *ETRI Journal*: Michel Abdalla
- of *Applicable Algebra in Engineering, Communication and Computing*: David Pointcheval

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

- Master: David Pointcheval, Jacques Stern, Damien Vergnaud, Introduction to Cryptology, M1, ENS
- Master: Michel Abdalla, David Pointcheval, Cryptography, M2, MPRI
- Master: Damien Vergnaud, Advanced Algebra and Applications to Cryptography, Ecole Centrale Paris
- Master: David Pointcheval, Cryptography, M2, ESIEA
- IACR-SEAMS School on "Cryptography: Foundations and New Directions": David Pointcheval

8.2.2. Defenses

- PhD: Adrian Thillard, Counter-measures against side-channel attacks and secure multi-party computation, ENS, December 12th, 2016 (Supervisor: Damien Vergnaud)
- PhD: Alain Passelègue, Algebraic Frameworks for Pseudorandom Functions, ENS, December 9th, 2016 (Supervisor: Michel Abdalla)
- PhD: Mario Cornejo, Security for the cloud, ENS, November 17th, 2016 (Supervisor: Michel Abdalla)
- HdR: Hoeteck Wee, Advances in Functional Encryption, ENS, July 1st, 2016
- PhD: Fabrice Ben Hamouda, Diverse Modules and Zero-Knowledge, ENS, July 1st, 2016 (Supervisors: Michel Abdalla & David Pointcheval)

8.2.3. Supervision

- PhD in progress: Raphael Bost, Symmetric Searchable Encryption, from 2014, David Pointcheval (with Pierre-Alain Fouque, at Rennes)
- PhD in progress: Florian Bourse, Encryption Schemes for the Cloud, from 2014, Michel Abdalla & David Pointcheval
- PhD in progress: Geoffroy Couteau, Efficient secure two-party computation for the Cloud, from 2014, David Pointcheval & Hoeteck Wee
- PhD in progress: Rafael Del Pino, Lattice-Based Cryptography – Complexity and Ideal-Lattices, from 2014, Vadim Lyubashevsky
- PhD in progress: Pierrick Meaux, Lattice-Based Cryptography – Advanced Features, from 2014, Vadim Lyubashevsky
- PhD in progress: Thierry Mefenza Nountu, Number-Theoretic Study of Pseudorandom Cryptographic Primitives, from 2014, Damien Vergnaud
- PhD in progress: Aurélien Dupin, Multi-Party Computations, from 2015, David Pointcheval (with Christophe Bidan, at Rennes)
- PhD in progress: Pierre-Alain Dupont, Secure Communications, from 2015, David Pointcheval
- PhD in progress: Romain Gay, Functional Encryption, from 2015, Michel Abdalla & Hoeteck Wee
- PhD in progress: Dahmun Gourdazi, Secure and Fast Cryptographic Implementation for Embedded Devices, from 2015, Damien Vergnaud
- PhD in progress: Louiza Khati, Disk Encryption Modes, from 2015, Damien Vergnaud
- PhD in progress: Michele Minelli, Increased efficiency and functionality through lattice-based cryptography, from 2015, Michel Abdalla & Hoeteck Wee
- PhD in progress: Anca Nitulescu, Verifiable Outsourced Computations, from 2015, David Pointcheval
- PhD in progress: Razvan Rosie, Practical Functional Encryption Schemes For the Cloud, from 2015, Michel Abdalla & Hoeteck Wee
- PhD in progress: Quentin Santos, Advanced Cryptography from a Blockchain, from 2015, David Pointcheval

- PhD in progress: Jérémy Chotard, Attribute-Based Encryption, from 2016, David Pointcheval (with Duong Hieu Phan, at Limoges)
- PhD in progress: Michele Orrù, Functional Encryption, from 2016, Hoeteck Wee & Georg Fuchsbauer

8.2.4. Juries

- PhD Adrian Thillard. *Countermeasures to side-channel attacks and secure multi-party computation* – ENS – France, December 12th, 2016: Damien Vergnaud (supervisor)
- PhD Alain Passelègue. *Algebraic Frameworks for Pseudorandom Functions* – ENS – France, December 9th, 2016: Michel Abdalla (supervisor)
- PhD Mario Cornejo. *Security for the cloud* – ENS – France, November 17th, 2016: Michel Abdalla (supervisor), David Pointcheval
- PhD Christian Janson. *On the Verification of Computation and Data Retrievability* – Royal Holloway University of London – UK, October 11th, 2016: Michel Abdalla
- PhD Brice Minaud. *Analyse de primitives cryptographiques récentes* – Université Rennes I – France, October 7th, 2016: David Pointcheval
- PhD Houda Ferradi. *Integrity, Authentication and Confidentiality in Public-Key Cryptography* – ENS – France, September 22nd, 2016: Michel Abdalla
- HdR Hoeteck Wee. *Advances in Functional Encryption* – ENS – France, July 1st, 2016: Michel Abdalla, David Pointcheval
- PhD Fabrice Ben Hamouda. *Diverse Modules and Zero-Knowledge* – ENS – France, July 1st, 2016: Michel Abdalla & David Pointcheval (supervisors)
- PhD Alberto Battistello. *On the security of embedded systems against physical attacks* – UVSQ – France, June 29th, 2016: David Pointcheval
- PhD Antoine Delignat-Lavaud. *On the Security of Authentication Protocol for the Web* – ENS – France, March 14th, 2016: David Pointcheval
- PhD Rémy Chrétien. *Automated analysis of equivalence properties for cryptographic protocols* – ENS Cachan – January 11th, 2016: David Pointcheval

9. Bibliography

Major publications by the team in recent years

[1] M. ABDALLA, M. BELLARE, D. CATALANO, E. KILTZ, T. KOHNO, T. LANGE, J. MALONE-LEE, G. NEVEN, P. PAILLIER, H. SHI.*Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous IBE, and Extensions*, in "Journal of Cryptology", July 2008, vol. 21, n^{o} 3, p. 350–391.

[2] 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.

[3] G. BARTHE, D. POINTCHEVAL, S. ZANELLA-BÉGUELIN.*Verified Security of Redundancy-Free Encryption from Rabin and RSA*, in "Proceedings of the 19th ACM Conference on Computer and Communications Security (CCS '12)", Raleigh, NC, USA, T. YU, G. DANEZIS, V. D. GLIGOR (editors), ACM Press, 2012, p. 724–735.

[4] 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.

[5] I. DINUR, O. DUNKELMAN, N. KELLER, A. SHAMIR.*New Attacks on Feistel Structures with Improved Memory Complexities*, in "Advances in Cryptology – Proceedings of CRYPTO '15 (1)", R. GENNARO, M. ROBSHAW (editors), Lecture Notes in Computer Science, Springer, 2015, vol. 9215, p. 433-454.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] V. LYUBASHEVSKY, T. PREST.*Quadratic Time, Linear Space Algorithms for Gram-Schmidt Orthogonalization and Gaussian Sampling in Structured Lattices*, in "Advances in Cryptology – Proceedings of Eurocrypt '15 (1)", E. OSWALD, M. FISCHLIN (editors), Lecture Notes in Computer Science, Springer, 2015, vol. 9056, p. 789-815.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] F. BENHAMOUDA.*Diverse modules and zero-knowledge*, PSL Research University ; ENS Paris, July 2016, https://hal.inria.fr/tel-01399476.

[12] M. CORNEJO-RAMIREZ.*Security for the Cloud*, ENS Paris - Ecole Normale Supérieure de Paris, November 2016, https://hal.inria.fr/tel-01399914.

[13] A. PASSELÈGUE.*Algebraic Frameworks for Pseudorandom Functions*, PSL Research University, December 2016, https://hal.inria.fr/tel-01422093.

[14] A. THILLARD.*Countermeasures to side-channel attacks and secure multi-party computation*, Ecole normale supérieure - ENS PARIS ; PSL Research University, December 2016, https://hal.inria.fr/tel-01415754.

[15] H. WEE.*Advances in Functional Encryption*, ENS Paris - Ecole Normale Supérieure de Paris, July 2016, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01399451.

Articles in International Peer-Reviewed Journal

[16] M. ABDALLA, F. BENHAMOUDA, D. POINTCHEVAL.*Public-key encryption indistinguishable under plaintext-checkable attacks*, in "IET Information Security", November 2016, vol. 10, n^{o} 6, p. 288–303 [*DOI :* 10.1049/IET-IFS.2015.0500], https://hal.inria.fr/hal-01385178.

[17] F. BENHAMOUDA, J. HERRANZ, M. JOYE, B. LIBERT.*Efficient Cryptosystems From 2^k -th Power Residue Symbols*, in "Journal of Cryptology", April 2016 [*DOI :* 10.1007/s00145-016-9229-5], https://hal.inria.fr/hal-01394400.

[18] F. BENHAMOUDA, M. JOYE, B. LIBERT.*A New Framework for Privacy-Preserving Aggregation of Time-Series Data*, in "ACM Transactions on Information and System Security", April 2016, vol. 18, n^o 3, 21 [*DOI :* 10.1145/2873069], https://hal.inria.fr/hal-01181321.

[19] S. CANARD, D. POINTCHEVAL, O. SANDERS, J. TRAORÉ.*Divisible e-cash made practical*, in "IET Information Security", July 2016 [*DOI :* 10.1049/IET-IFS.2015.0485], https://hal.inria.fr/hal-01377998.

[20] C. HAZAY, A. LÓPEZ-ALT, H. WEE, D. WICHS.*Leakage-Resilient Cryptography from Minimal Assumptions*, in "Journal of Cryptology", 2016, vol. 29, n^o 3, p. 514–551 [*DOI :* 10.1007/s00145-015-9200-X], https://hal.inria.fr/hal-01378199.

[21] D. VERGNAUD.*Comment on "A strong provably secure IBE scheme without bilinear map" by M. Zheng, Y. Xiang and H. Zhou [J. Comput. Syst. Sci. 81 (2015) 125–131]*, in "Journal of Computer and System Sciences (JCSS)", August 2016, vol. 82, n^o 5, 2 [*DOI :* 10.1016/J.JCSS.2015.12.003], https://hal.inria.fr/hal-01305462.

International Conferences with Proceedings

[22] M. ABDALLA, M. CORNEJO, A. NITULESCU, D. POINTCHEVAL.*Robust Password-Protected Secret Sharing*, in "ESORICS 2016 - 21st European Symposium on Research in Computer Security", Heraklion, Greece, I. ASKOXYLAKIS, S. IOANNIDIS, S. KATSIKAS, C. MEADOWS (editors), Lecture Notes in Computer Science, Springer, September 2016, vol. 9879, p. 61-79 [*DOI :* 10.1007/978-3-319-45741-3_4], https://hal.archives-ouvertes.fr/hal-01380699.

[23] H. ABUSALAH, G. FUCHSBAUER.*Constrained PRFs for Unbounded Inputs with Short Keys*, in "Applied Cryptography and Network Security - 14th International Conference, ACNS 2016", Guildford, United Kingdom, June 2016 [*DOI :* 10.1007/978-3-319-39555-5_24], https://hal.inria.fr/hal-01384375.

[24] H. ABUSALAH, G. FUCHSBAUER, K. PIETRZAK.*Offline Witness Encryption*, in "Applied Cryptography and Network Security - 14th International Conference, ACNS 2016", Guildford, United Kingdom, June 2016 [*DOI :* 10.1007/978-3-319-39555-5_16], https://hal.inria.fr/hal-01384371.

[25] S. BELAID, F. BENHAMOUDA, A. PASSELÈGUE, E. PROUFF, A. THILLARD, D. VERGNAUD.*Randomness Complexity of Private Circuits for Multiplication*, in "EUROCRYPT 2016", Vienna, Austria, May 2016, p. 616-648 [*DOI :* 10.1007/978-3-662-49896-5_22], https://hal.archives-ouvertes.fr/hal-01324823.

[26] M. BELLARE, G. FUCHSBAUER, A. SCAFURO.*NIZKs with an Untrusted CRS: Security in the Face of Parameter Subversion*, in "Advances in Cryptology - ASIACRYPT 2016 - 22nd International Conference on the Theory and Application of Cryptology and Information Security", Hanoi, Vietnam, December 2016, https://hal.inria.fr/hal-01384384.

[27] F. BENHAMOUDA, C. CHEVALIER, A. THILLARD, D. VERGNAUD.*Easing Coppersmith Methods Using Analytic Combinatorics: Applications to Public-Key Cryptography with Weak Pseudorandomness*, in "Public-Key Cryptography – PKC 2016", Taipei, Taiwan, IACR, March 2016, vol. 9615, 31 [*DOI :* 10.1007/978-3-662-49387-8_3], https://hal.inria.fr/hal-01278460.

[28] N. BITANSKY, R. NISHIMAKI, A. PASSELÈGUE, D. WICHS.*From Cryptomania to Obfustopia through Secret-Key Functional Encryption*, in "TCC 2016-B - Fourteenth IACR Theory of Cryptography Conference", Beijing, China, October 2016, https://hal.inria.fr/hal-01379256.

[29] O. BLAZY, C. CHEVALIER, D. VERGNAUD.*Mitigating Server Breaches in Password-Based Authentication: Secure and Efficient Solutions*, in "CT-RSA 2016", San Francisco, France, K. SAKO (editor), The Cryptographers' Track at the RSA Conference, February 2016, vol. LNCS, n^{o} 9610 [*DOI :* 10.1007/978-3-319-29485-8_1], https://hal.archives-ouvertes.fr/hal-01292699.

[30] F. BOURSE, R. DEL PINO, M. MINELLI, H. WEE.*FHE Circuit Privacy Almost for Free*, in "Crypto 2016 - 36th Annual International Cryptology Conference", Santa Barbara, United States, Crypto 2016, Springer Verlag, August 2016, vol. Lecture Notes in Computer Science, n^{o} 9815 [*DOI :* 10.1007/978-3-662-53008-5_3], https://hal.inria.fr/hal-01360110.

[31] Z. BRAKERSKI, D. CASH, R. TSABARY, H. WEE.*Targeted Homomorphic Attribute-Based Encryption*, in "14th International Conference, TCC 2016-B", Beijing, China, M. HIRT, A. SMITH (editors), Lecture Notes in Computer Science, Springer, October 2016, vol. 9986, p. 330-360, https://hal.inria.fr/hal-01378341.

[32] Z. BRAKERSKI, V. VAIKUNTANATHAN, H. WEE, D. WICHS.*Obfuscating Conjunctions under Entropic Ring LWE*, in "ITCS", Cambridge, United States, 2016 [*DOI :* 10.1145/2840728.2840764], https://hal.inria.fr/hal-01378193.

[33] P. CHAIDOS, V. CORTIER, G. FUCHSBAUER, D. GALINDO.*BeleniosRF: A Non-interactive Receipt-Free Electronic Voting Scheme*, in "23rd ACM Conference on Computer and Communications Security (CCS'16)", Vienna, Austria, October 2016 [*DOI :* 10.1145/2976749.2978337], https://hal.inria.fr/hal-01377917.

[34] C. CHEVALIER, F. LAGUILLAUMIE, D. VERGNAUD.*Privately Outsourcing Exponentiation to a Single Server: Cryptanalysis and Optimal Constructions*, in "Computer Security - ESORICS 2016", Heraklion, Greece, I. G. ASKOXYLAKIS, S. IOANNIDIS, S. K. KATSIKAS, C. A. MEADOWS (editors), Computer Security – ESORICS 2016, Springer, September 2016, vol. 9878, p. 261-278 [*DOI :* 10.1007/978-3-319-45744-4_13], https://hal.inria.fr/hal-01375817.

[35] G. COUTEAU, T. PETERS, T. PETERS, D. POINTCHEVAL.*Encryption Switching Protocols*, in "Crypto 2016 - 36th Annual International Cryptology Conference", Santa Barbara, United States, M. ROBSHAW, J. KATZ (editors), Springer, August 2016 [*DOI :* 10.1007/978-3-662-53018-4_12], https://hal.inria.fr/hal-01407341.

[36] R. DEL PINO, V. LYUBASHEVSKY, D. POINTCHEVAL.*The Whole is Less Than the Sum of Its Parts: Constructing More Efficient Lattice-Based AKEs*, in "SCN 2016 - 10th International Conference Security and Cryptography for Networks", Amalfi, Italy, V. ZIKAS, R. D. PRISCO (editors), Security and Cryptography for Networks, Springer , August 2016, vol. Lecture Notes in Computer Science, n^{o} 9841, p. 273 - 291 [*DOI :* 10.1007/978-3-319-44618-9_15], https://hal.inria.fr/hal-01378005.

[37] H. FERRADI, R. GÉRAUD, D. MAIMUT, D. NACCACHE, D. POINTCHEVAL.*Legally Fair Contract Signing Without Keystones*, in "ACNS 2016 - 14th International Conference Applied Cryptography and Network Security", Guildford, United Kingdom, M. MANULIS, A.-R. SADEGHI, S. SCHNEIDER (editors), Applied Cryptography and Network Security, Springer, June 2016, vol. LNCS, n^{o} 9696, p. 175 - 190 [*DOI :* 10.1007/978-3-319-39555-5_10], https://hal.inria.fr/hal-01377993.

[38] D. FIORE, A. NITULESCU.*On the (In)security of SNARKs in the Presence of Oracles*, in "TCC 2016-B - Fourteenth IACR Theory of Cryptography Conference", Beijing, China, Theory of Cryptography 14th International Conference, TCC 2016-B, Beijing, China, November 1-3, 2016, Proceedings, October 2016, https://hal.inria.fr/hal-01378013.

[39] G. FUCHSBAUER, C. HANSER, C. KAMATH, D. SLAMANIG.*Practical Round-Optimal Blind Signatures in the Standard Model from Weaker Assumptions*, in "Security and Cryptography for Networks - 10th International Conference, SCN 2016", Amalfi, Italy, August 2016 [*DOI :* 10.1007/978-3-319-44618-9_21], https://hal.inria.fr/hal-01384381.

[40] *Best Paper*
R. GAY, D. HOFHEINZ, E. KILTZ, H. WEE.*Tightly CCA-Secure Encryption without Pairings*, in "Advances in Cryptology – EUROCRYPT 2016", Vienna, Austria, Advances in Cryptology – EUROCRYPT 2016, May 2016, vol. Lectures Notes in Computer Science, n^o 9665 [*DOI :* 10.1007/978-3-662-49890-3_1], https://hal.archives-ouvertes.fr/hal-01302516.

[41] D. GOUDARZI, M. RIVAIN.*On the Multiplicative Complexity of Boolean Functions and Bitsliced Higher-Order Masking* , in "CHES", Santa-Barbara, United States, 2016, https://hal.inria.fr/hal-01379296.

[42] D. GOUDARZI, M. RIVAIN, D. VERGNAUD.*Lattice Attacks against Elliptic-Curve Signatures with Blinded Scalar Multiplication*, in "Selected Areas in Cryptography - SAC 2016", St. John's, Canada, R. AVANZI, H. HEYS (editors), Selected Areas in Cryptography - SAC 2016, Springer, August 2016, https://hal.inria.fr/hal-01379249.

[43] L. KHATI, N. MOUHA, D. VERGNAUD.*Full Disk Encryption: Bridging Theory and Practice*, in "CT-RSA 2017 - RSA Conference Cryptographers' Track", San Francisco, United States, Lecture Notes in Computer Science, February 2017, 16, https://hal.inria.fr/hal-01403418.

[44] H. KRAWCZYK, H. WEE.*The OPTLS Protocol and TLS 1.3*, in "EuroS&P", Saarbrücken, Germany, 2016 [*DOI :* 10.1109/EUROSP.2016.18], https://hal.inria.fr/hal-01378195.

[45] T. MEFENZA.*Inferring Sequences Produced by a Linear Congruential Generator on Elliptic Curves Using Coppersmith's Methods*, in "COCOON 2016", Ho Chi Minh City, Vietnam, T. N. DINH, M. T. THAI (editors), 22nd International Computing and Combinatorics Conference, COCOON 2016, Springer Verlag, August 2016, vol. Lecture Notes in Computer Science, n^o 9797, 12 [*DOI :* 10.1007/978-3-319-42634-1_24], https://hal.inria.fr/hal-01381658.

[46] P. MÉAUX, A. JOURNAULT, F.-X. STANDAERT, C. CARLET.*Towards Stream Ciphers for Efficient FHE with Low-Noise Ciphertexts*, in "Advances in Cryptology – EUROCRYPT 2016", WIEN, Austria, Lecture Notes in Computer Science, May 2016, vol. 9665, p. 311-343 [*DOI :* 10.1007/978-3-662-49890-3_13], https://hal.inria.fr/hal-01405859.

[47] D. POINTCHEVAL, O. SANDERS.*Short Randomizable Signatures*, in "The Cryptographers' Track at the RSA Conference 2016", San Francisco, United States, K. SAKO (editor), Springer Verlag, February 2016, vol. LNCS, n^o 9610, p. 111 - 126 [*DOI :* 10.1007/978-3-319-29485-8_7], https://hal.inria.fr/hal-01377997.

[48] H. WEE.*Déjà Q: Encore! Un Petit IBE*, in "TCC 2016 A", Tel Aviv, Israel, 2016 [*DOI :* 10.1007/978-3-662-49099-0_9], https://hal.inria.fr/hal-01378189.

[49] H. WEE.*KDM-Security via Homomorphic Smooth Projective Hashing*, in "Public Key Cryptography", Taipei, Taiwan, 2016 [*DOI :* 10.1007/978-3-662-49387-8_7], https://hal.inria.fr/hal-01378191.

Books or Proceedings Editing

[50] D. POINTCHEVAL, A. NITAJ, T. RACHIDI (editors). *8th International Conference on Cryptology in Africa - Africacrypt 2016*, Africacrypt 2016, Springer Verlag, Fes, Morocco, 2016, vol. LNCS, n^o 9646 [*DOI :* 10.1007/978-3-319-31517-1], https://hal.inria.fr/hal-01377995.

Research Reports

[51] M. ABDALLA, F. BOURSE, A. DE CARO, D. POINTCHEVAL.*Better Security for Functional Encryption for Inner Product Evaluations*, IACR, January 2016, n^o Cryptology ePrint Archive: Report 2016/011, https://hal.archives-ouvertes.fr/hal-01380726.

[52] M. ABDALLA, M. CORNEJO, A. NITULESCU, D. POINTCHEVAL.*Robust Password-Protected Secret Sharing*, IACR, May 2016, n^o Cryptology ePrint Archive: Report 2016/123, https://hal.archives-ouvertes.fr/hal-01380730.

[53] M. ABDALLA, M. RAYKOVA, H. WEE.*Multi-Input Inner-Product Functional Encryption from Pairings*, IACR, April 2016, n^o Cryptology ePrint Archive: Report 2016/425, https://hal.archives-ouvertes.fr/hal-01380735.

[54] J. CHOTARD, D. HIEU PHAN, D. POINTCHEVAL.*Homomorphic-Policy Attribute-Based Key Encapsulation Mechanisms*, IACR Cryptology ePrint Archive, November 2016, n^o Cryptology ePrint Archive: Report 2016/1089, https://hal.inria.fr/hal-01402517.

Other Publications

[55] M. JOYE, A. PASSELÈGUE.*Function-Revealing Encryption*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01379260.

Project-Team CLIME

Coupling environmental data and simulation models for software integration

RESEARCH CENTER
Paris

THEME
Earth, Environmental and Energy Sciences

Table of contents

Project-Team CLIME

Creation of the Project-Team: 2005 September 01, end of the Project-Team: 2016 December 31

Keywords:

Computer Science and Digital Science:

3.3. - Data and knowledge analysis
3.4.1. - Supervised learning
3.4.5. - Bayesian methods
5.3. - Image processing and analysis
5.9. - Signal processing
6.2.1. - Numerical analysis of PDE and ODE
6.2.3. - Probabilistic methods
6.2.7. - High performance computing
6.3. - Computation-data interaction
6.3.1. - Inverse problems
6.3.2. - Data assimilation
6.3.3. - Data processing
6.3.4. - Model reduction
6.3.5. - Uncertainty Quantification

Other Research Topics and Application Domains:

3.3. - Geosciences
3.4. - Risks
4.3.3. - Wind energy
4.3.4. - Solar Energy
8.2. - Connected city

1. Members

Research Scientists

Isabelle Herlin [Team leader, Inria, Senior Researcher, HDR]
Vivien Mallet [Inria, Researcher]

Faculty Members

Julien Brajard [Univ. Paris VI, Associate Professor, until Aug 2016]
Étienne Huot [Univ. Versailles Saint-Quentin-en-Yvelines, Associate Professor]

Technical Staff

Guillaume Chérel [Inria, from Apr 2016]
Nicolas Claude [Inria, until Feb 2016]
Marius Guérard [Inria, from Nov 2016]

PhD Students

Jean Thorey [EDF]
Raphaël Ventura [Inria]

Administrative Assistant

Nathalie Gaudechoux [Inria]

Others

Dominique Béréziat [Univ. Paris VI, Associate Professor, external collaborator, HDR]
Timothée Maziol [Inria, Internship, from Apr 2016 until Aug 2016]
Louis Philippe [Inria, Internship, from Jul 2016]

2. Overall Objectives

2.1. Clime in short

The international politic, economic and scientific contexts are pointing out the role that is played by models and observation systems for forecasting and evaluating environmental risks.

The complexity of environmental phenomena as well as the operational objectives of risk mitigation necessitate an intensive interweaving between geophysical models, data processing, simulation, visualization and database tools.

For illustration purpose, we observe that this situation is met in the domain of atmospheric pollution, whose modeling is gaining an ever-increasing significance and impact, either at local (air quality), regional (transboundary pollution) or global scale (greenhouse effect). In this domain, numerical modeling systems are used for operational forecasts (short or long term), detailed case studies, impact studies for industrial sites, as well as coupled modeling, such as pollution and health or pollution and economy. All these scientific subjects strongly require linking/coupling the models with all available data either of physical origin (e.g., models outputs), coming from raw observations (satellite acquisitions and/or information measured in situ by an observation network) or obtained by processing and analysis of these observations (e.g., chemical concentrations retrieved by inversion of a radiative transfer model).

Clime has been created for studying these questions with researchers specialized in data assimilation and image processing.

Clime carries out research activities in three main areas:

- Data assimilation methods: inverse modeling, network design, ensemble methods, uncertainties estimation, uncertainties propagation.
- Image assimilation: assimilation of structures in environmental forecasting models, study of ill-posed image processing problems with data assimilation technics, definition of dynamic models from images, reduction of models.
- Development of integrated chains for data/models/outputs (system architecture, workflow, database, visualization).

3. Research Program

3.1. Data assimilation and inverse modeling

This activity is one major concern of environmental sciences. It matches up the setting and the use of data assimilation methods, for instance variational methods (such as the 4D-Var method). An emerging issue lies in the propagation of uncertainties by models, notably through ensemble forecasting methods.

Although modeling is not part of the scientific objectives of Clime, the project-team has complete access to air quality models through collaborations with École des Ponts ParisTech and EDF R&D: the models from Polyphemus (pollution forecasting from local to regional scales) and Code_Saturne (urban scale). In regard to other modeling domains, such as oceanography and meteorology, Clime accesses models through co-operation with LOCEAN (Laboratoire d'OCEANographie et du climat, UPMC) and Météo-France.

The research activities of Clime tackle scientific issues such as:

- Within a family of models (differing by their physical formulations and numerical approximations), which is the optimal model for a given set of observations?
- How to reduce dimensionality of problems by Galerkin projection of equations on subspaces? How to define these subspaces in order to keep the main properties of systems?
- How to assess the quality of a forecast and its uncertainty? How do data quality, missing data, data obtained from sub-optimal locations, affect the forecast? How to better include information on uncertainties (of data, of models) within the data assimilation system?
- How to make a forecast (and a better forecast!) by using several models corresponding to different physical formulations? It also raises the question: how should data be assimilated in this context?
- Which observational network should be set up to perform a better forecast, while taking into account additional criteria such as observation cost? What are the optimal location, type and mode of deployment of sensors? How should trajectories of mobile sensors be operated, while the studied phenomenon is evolving in time? This issue is usually referred as “network design”.

3.2. Satellite acquisitions and image assimilation

In geosciences, the issue of coupling data, in particular satellite acquisitions, and models is extensively studied for meteorology, oceanography, chemistry-transport and land surface models. However, satellite images are mostly assimilated on a point-wise basis. Three major approaches arise if taking into account the spatial structures, whose displacement is visualized on image sequences:

- Image approach. Image assimilation allows the extraction of features from image sequences, for instance motion field or structures’ trajectory. A model of the dynamics is considered (obtained by simplification of a geophysical model such as Navier-Stokes equations). An observation operator is defined to express the links between the model state and the pixel values or some image features. In the simplest case, the pixel value corresponds to one coordinate of the model state and the observation operator is reduced to a projection. However, in most cases, this operator is highly complex, implicit and non-linear. Data assimilation techniques are developed to control the initial state or the whole assimilation window. Image assimilation is also applied to learn reduced models from image data and estimate a reliable and small-size reconstruction of the dynamics, which is observed on the sequence.
- Model approach. Image assimilation is used to control an environmental model and obtain improved forecasts. In order to take into account the spatial and temporal coherency of structures, specific image characteristics are considered and dedicated norms and observation error covariances are defined.
- Correcting a model. Another topic, mainly described for meteorology in the literature, concerns the location of structures. How to force the existence and to correct the location of structures in the model state using image information? Most of the operational meteorological forecasting institutes, such as Météo-France (in France), UK-met (in United Kingdom), KNMI (in Netherlands), ZAMG (in Austria) and Met-No (in Norway), study this issue because operational forecasters often modify their forecasts based on visual comparisons between the model outputs and the structures displayed on satellite images.

3.3. Software chains for environmental applications

An objective of Clime is to participate in the design and creation of software chains for impact assessment and environmental crisis management. Such software chains bring together static or dynamic databases, data assimilation systems, forecast models, processing methods for environmental data and images, complex visualization tools, scientific workflows, ...

Clime is currently building, in partnership with École des Ponts ParisTech and EDF R&D, such a system for air pollution modeling: Polyphemus (see the web site http://cerea.enpc.fr/polyphemus/), whose architecture is specified to satisfy data requirements (e.g., various raw data natures and sources, data preprocessing) and to support different uses of an air quality model (e.g., forecasting, data assimilation, ensemble runs).

4. Application Domains

4.1. Introduction

The first application domain of the project-team is atmospheric chemistry. We develop and maintain the air quality modeling system Polyphemus, which includes several numerical models (Gaussian models, Lagrangian model, two 3D Eulerian models including Polair3D) and their adjoints, and different high level methods: ensemble forecast, sequential and variational data assimilation algorithms. Advanced data assimilation methods, network design, inverse modeling, ensemble forecast are studied in the context of air chemistry. Note that addressing these high level issues requires controlling the full software chain (models and data assimilation algorithms).

The activity on assimilation of satellite data is mainly carried out for meteorology and oceanography. This is addressed in cooperation with external partners who provide numerical models. Concerning oceanography, the aim is to assess ocean surface circulation, by assimilating fronts and vortices displayed on image acquisitions. Concerning meteorology, the focus is on correcting the location of structures related to high-impact weather events (cyclones, convective storms, etc.) by assimilating images.

4.2. Air quality

Air quality modeling implies studying the interactions between meteorology and atmospheric chemistry in the various phases of matter, which leads to the development of highly complex models. The different usages of these models comprise operational forecasting, case studies, impact studies, etc., with both societal (e.g., public information on pollution forecast) and economical impacts (e.g., impact studies for dangerous industrial sites). Models lack some appropriate data, for instance better emissions, to perform an accurate forecast and data assimilation techniques are recognized as a major key point for improving forecast's quality.

In this context, Clime is interested in various problems, the following being the crucial ones:

- The development of ensemble forecast methods for estimating the quality of the prediction, in relation with the quality of the model and the observations. The ensemble methods allow sensitivity analysis with respect to the model's parameters so as to identify physical and chemical processes, whose modeling must be improved.
- The development of methodologies for sequential aggregation of ensemble simulations. What ensembles should be generated for that purpose, how spatialized forecasts can be generated with aggregation, how can the different approaches be coupled with data assimilation?
- The definition of second-order data assimilation methods for the design of optimal observation networks. The two main objectives are: management of combinations of sensor types, and deployment modes and dynamic management of mobile sensors' trajectories.
- How to estimate the emission rate of an accidental release of a pollutant, using observations and a dispersion model (from the near-field to the continental scale)? How to optimally predict the evolution of a plume? Hence, how to help people in charge of risk evaluation for the population?
- The definition of non-Gaussian approaches for data assimilation.
- The assimilation of satellite measurements of troposphere chemistry.

The activities of Clime in air quality are supported by the development of the Polyphemus air quality modeling system. This system has a modular design, which makes it easier to manage high level applications such as inverse modeling, data assimilation and ensemble forecast.

4.3. Oceanography

The capacity of performing a high quality forecast of the state of the ocean, from the regional to the global scales, is of major interest. Such a forecast can only be obtained by systematically coupling numerical models and observations (in situ and satellite data). In this context, being able to assimilate image structures becomes a key point. Examples of such image structures are:

- apparent motion field that represents surface velocity;
- trajectories, obtained either from tracking of features or from integration of the velocity field;
- spatial objects, such as fronts, eddies or filaments.

Image models of these structures are developed and take into account the underlying physical processes. Image acquisitions are assimilated into these models to derive pseudo-observations of state variables, which are further assimilated in numerical ocean forecast models.

4.4. Meteorology

Meteorological forecasting constitutes a major applicative challenge for image assimilation. Although satellite data are operationally assimilated within models, this is mainly done on an independent pixel basis: the observed radiance is linked to the state variables via a radiative transfer model, that plays the role of an observation operator. Indeed, because of their limited spatial and temporal resolutions, numerical weather forecast models fail to exploit image structures, such as precursors of high impact weather:

- cyclogenesis related to the intrusion of dry stratospheric air in the troposphere (a precursor of cyclones),
- convective systems (supercells) leading to heavy winter time storms,
- low-level temperature inversion leading to fog and ice formation, etc.

To date, there is no available method for assimilating such data, which are characterized by a strong coherence in space and time. Meteorologists have developed qualitative Conceptual Models (CMs), for describing the high impact weathers and their signature on images, and tools to detect CMs on image data. The result of this detection is used for correcting the numerical models, for instance by modifying the initialization. The aim is therefore to develop a methodological framework allowing to assimilate the detected CMs within numerical forecast models. This is a challenging issue given the considerable impact of the related meteorological events.

4.5. 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 develop. Numerical simulation is well established to study the urban environment, e.g., for road traffic modeling. 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. A number of research issues can be raised:

- How to properly take into account the city geometry that makes the data assimilation problems unique?
- How to make use of the various sensors, sometimes mobile, of low quality but numerous?
- How to couple all the systems that are intricated at urban scale?

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. Awards

Inria and Paris City were awarded a Décibel d'Argent 2016 in research category for the mobile application Ambiciti. The award was attributed by the Conseil National du Bruit, which depends on the Ministry of Ecology, Sustainable Development and Energy, and is a national organization in charge of noise. The selection committee pointed out the Ambiciti articulation between research, citizen involvement, city or government actions and the operational development of a rich and perennial mobile application.

6. New Software and Platforms

6.1. Heimdali

- Participants: Isabelle Herlin, Dominique Bereziat and David Froger
- Contact: Isabelle Herlin

SCIENTIFIC DESCRIPTION

The main components of Heimdali concern:

- the pre/post processing of image sequences,
- the image assimilation with numerical models,
- the visualization of image sequences.

FUNCTIONAL DESCRIPTION

The initial aim of the image processing library Heimdali was to replace an internal Inria library (named Inrimage) by a library based on standard and open source tools, and mostly dedicated to satellite acquisitions.

The leading idea of the library is to allow the following issues:
- making easier the sharing and development of image assimilation softwares. For that purpose, the installation is easily achieved with the package manager Conda.
- developing generic tools for image processing and assimilation based on ITK (Insight Segmentation and Registration Toolkit http://www.itk.org). In reverse provide tools to ITK and contribute to the ITK community. Our software corresponds to issues related to satellite acquisitions but could be of interest for processing medical image sequences.

6.2. Image Forecast

- Authors: Isabelle Herlin and Yann Lepoittevin
- Contact: Isabelle Herlin

SCIENTIFIC DESCRIPTION

From a given number of images, Image Forecast synthetizes the future images at a given and short temporal horizon.

FUNCTIONAL DESCRIPTION

Image forecast includes two components:

- it estimates the dynamics from an image sequence. Various options are available in the software: stationarity, Lagrangian conservation, description of structures. The result is the motion field explaining the temporal evolution of image data.
- the estimated dynamics is applied for forecasting future images at short temporal horizon.

6.3. Polyphemus

- Participants: Sylvain Doré (CEREA, École des Pont ParisTech) and Vivien Mallet
- Contact: Vivien Mallet
- URL: http://cerea.enpc.fr/polyphemus/

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.

Fig. 1 depicts a typical result produced by Polyphemus.

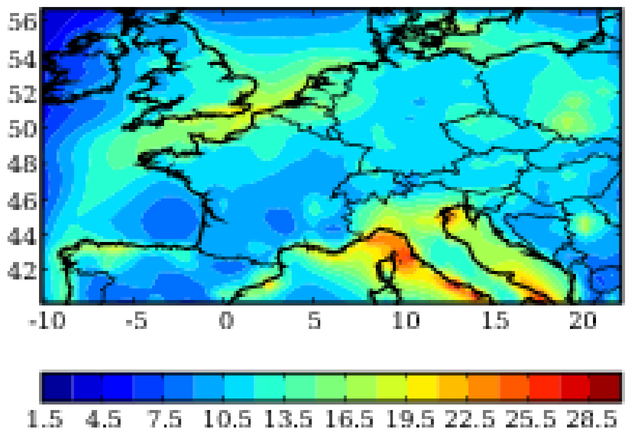

Figure 1. Map of the relative standard deviation (or spread, %) of an ensemble built with Polyphemus (ozone simulations, $\mu g\ m^{-3}$). The standard deviations are averaged over the summer of 2001. They provide an estimation of the simulation uncertainties.

Clime is involved in the overall design of the system and in the development of advanced methods in model coupling, data assimilation and uncertainty quantification (through model drivers and post-processing).

6.4. SoundCity - Ambiciti

- Authors: Pierre-Guillaume Raverdy (SED), Fadwa Rebhi (Mimove), Cong Kinh Nguyen (Mimove), Rajiv Bhatia (TheCivicEngine), Vivien Mallet and Valerie Issarny (Mimove)
- Contact: Valerie Issarny (Mimove)

FUNCTIONAL DESCRIPTION

Ambiciti measures the actual noise levels to which you are exposed. It can monitor noise levels throughout the day and inform you about your instantaneous, hourly and daily exposures.
Ambiciti also computes the air quality index in your region or at the exact location where you stand. You can also access to forecasts.
Ambiciti includes a lot of features:

- Measuring noise level, anytime on demand or automatically during the day,
- Air quality indexes, in the past, present and future hours or days,
- Pollution levels for nitrogen dioxide, fine particulate matter and ozone,
- Statistics on exposure to pollutions, hourly, daily, during daytime and nighttime,
- Maps with your noise measurements,
- Hourly air quality maps, at street resolution in Paris, San Francisco, Oakland, Richmond (California), at present time,
- The recommendation of pedestrian routes which minimize the exposure to noise pollution or to air pollution,
- The ability to take pictures with pollution levels on top.

6.5. Urban noise analysis

- Authors: Vivien Mallet, Raphael Ventura and Guillaume Cherel
- Contact: Vivien Mallet

FUNCTIONAL DESCRIPTION

This software merges noise simulations and mobile observations. It can extract a given region of a noise map and filter out the buildings. It extends a previous software for data assimilation of air pollution observations at city scale. This prior software computes the so-called best linear unbiased estimator (BLUE), with a special background error covariance model that depends on the city geometry. The extension for noise introduces special treatments for the errors in mobile observations, and includes more statistical verifications.

The software also comes with a Python module for the management of a large database of mobile noise measurements, especially with many filters relying the observations metadata.

The software finally includes the automatic generation of a report based on intensive measurements in a city district. This report targets participants of crowdsensing experiments.

6.6. Verdandi

- Participants: Vivien Mallet, Gautier Bureau (Medisim), Dominique Chapelle (Medisim), Sébastien Gilles (Medisim) and Philippe Moireau (Medisim)
- Contact: Vivien Mallet
- URL: http://verdandi.gforge.inria.fr/

FUNCTIONAL DESCRIPTION

Verdandi is a free and open-source (LGPL) library for data assimilation. It includes various 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.

7. New Results

7.1. State estimation: analysis and forecast

One major objective of Clime is the conception of new methods of data assimilation in geophysical sciences. Clime is active on several challenging aspects: non-Gaussian assumptions, multiscale assimilation, minimax filtering, etc.

7.1.1. Assimilation of drifter data in the East Mediterranean Sea

Participants: Julien Brajard, Isabelle Herlin, Leila Issa [Lebanese American University, Lebanon], Laurent Mortier [LOCEAN], Daniel Hayes [Oceanography Centre, Cyprus], Milad Fakhri [CNRS, Lebanon], Pierre-Marie Poulain [Oceanography Institute of Trieste, Italy].

Surface velocity fields of the ocean in the Eastern Levantine Mediterranean are estimated by blending altimetry and surface drifters data. The method is based on a variational assimilation approach for which the velocity is corrected by matching real drifters positions with those predicted by a simple advection model, while taking into account the wind effect. The velocity correction is done in a time-continuous fashion by assimilating at once a whole trajectory of drifters with a temporal sliding window. Except for the wind component, a divergence-free regularization term was added to constrain the velocity field. Results show that, with few drifters, our method improves the estimated velocity in two typical situations: an eddy between the Lebanese coast and Cyprus, and velocities along the Lebanese coast. A description of these results is published in the Ocean Modelling journal.

7.1.2. State estimation for noise pollution

Participants: Raphaël Ventura, Vivien Mallet, Valérie Issarny [Mimove], Pierre-Guillaume Raverdy [SED], Fadwa Rebhi [Mimove], Cong Kinh Nguyen [Mimove].

70 million observations of ambient noise have been collected with the mobile application Ambiciti (previously, SoundCity). An important work was carried out on the calibration of the measurements. Over 100 mobile phones were calibrated against a sound level meter, at various noise intensities and frequencies, in order to test their response and devise a calibration strategy.

A data assimilation procedure has been put in place in order to select and assimilate the most reliable observations. Simulated noise maps have been improved with the observations, by computing the so-called best linear unbiased estimator (BLUE) with error covariance models suitable for noise pollution. The assimilation of mobile observation introduces new errors, like location errors, compared to the assimilation of the more common observations from fixed monitoring stations.

7.2. Image assimilation

Sequences of images, such as satellite acquisitions, display structures evolving in time. This information is recognized of major interest by forecasters (meteorologists, oceanographers, etc.) in order to improve the information provided by numerical models. However, the satellite images are mostly assimilated in geophysical models on a point-wise basis, discarding the space-time coherence visualized by the evolution of structures such as clouds. Assimilating image data in an optimal way is of major interest. This issue is twofold:

- from the model's viewpoint, the location of structures on the observations is used to control the state vector.
- from the image's viewpoint, a model of the dynamics and structures is built from the observations.

7.2.1. Estimation of motion and acceleration from image data

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

Image sequences allow visualizing dynamic systems and understanding their intrinsic characteristics. One first component of this dynamics is obtained by retrieving the velocity of the structures displayed on the sequence. This motion estimation issue has been extensively studied in the literature of image processing and computer vision. In this research, we step beyond the traditional optical flow methods and address the problem of recovering the acceleration from the whole temporal sequence, which has been poorly investigated, even if this is of major importance for some data types, such as fluid flow images. Acceleration is here viewed as the space-time function resulting from the forces applied to the studied system. To solve this issue, we propose a variational approach where a specific energy is designed to model both the motion and the acceleration fields. The contributions are twofold: first, we introduce a unified variational formulation of motion and acceleration under space-time constraints; second, we define the minimization scheme, which allows retrieving the estimations, and provide the full information on the discretization schemes. Experiments are conducted on synthetic and real image sequences, visualizing fluid-like flows, where direct and precise calculation of acceleration is of primary importance.

7.2.2. Rain nowcasting from radar image acquisitions

Participants: Isabelle Herlin, Étienne Huot.

This research concerns the design of an operational method for rainfall nowcasting that aims at prevention of flash floods. The nowcasting method includes two main components:

- a data assimilation method, based on radar images, estimates the state of the atmosphere: this is the estimation phase.
- a forecast method uses this estimation to extrapolate the state of the atmosphere in the future: this is the forecast phase.

Results are analyzed on space-time neighborhoods in order to prevent consequences of flash floods on previously defined zone.

Current research concerns the following issues:

- the use of object components in the state vector. The objective is to improve the description of the image data in order to get a better motion estimation and a more accurate localization of endangered regions.
- the extension of the estimation phase to a multiscale process.
- the merging with measures acquired by a network of pluviometers.

7.2.3. Ensemble Kalman filter based on the image structures

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

One major limitation of the motion estimation methods that are available in the literature concerns the availability of the uncertainty on the result. This is however assessed by a number of filtering methods, such as the ensemble Kalman filter (EnKF). Our research consequently concerns the use of a description of the displayed structures in an ensemble Kalman filter, which is applied for estimating motion on image acquisitions. Compared to the Kalman filter, EnKF does not require propagating in time the error covariance matrix associated to the estimation, resulting in reduced computational requirements. However, EnKF is also known for exhibiting a shrinking effect when taking into account the observations on the studied system at the analysis step. Various methods are available in the literature for correcting this effect, but they do not involve the structures displayed on the image sequence. We defined two alternative solutions to that shrinking effect: a dedicated localization function and an adaptive decomposition domain. These methods are both well suited for fluid flows images and applied on satellite images of the atmosphere.

7.3. Uncertainty quantification and risk assessment

The uncertainty quantification of environmental models raises a number of problems due to:

- the dimension of the inputs, which can easily be 10^5-10^8 at every time step;
- the dimension of the state vector, which is usually 10^5-10^7;
- the high computational cost required when integrating the model in time.

While uncertainty quantification is a very active field in general, its implementation and development for geosciences requires specific approaches that are investigated by Clime. The project-team tries to determine the best strategies for the generation of ensembles of simulations. In particular, this requires addressing the generation of large multimodel ensembles and the issue of dimension reduction and cost reduction. The dimension reduction consists in projecting the inputs and the state vector to low-dimensional subspaces. The cost reduction is carried out by emulation, i.e., the replacement of costly components with fast surrogates.

7.3.1. Sequential aggregation with uncertainty estimation

Participants: Jean Thorey, Vivien Mallet, Christophe Chaussin [EdF R&D].

In the context of ensemble forecasting, one goal is to combine an ensemble of forecasts in order to produce an improved probabilistic forecast. We previously designed a new approach to predict a probability density function or cumulative distribution function, from a weighted ensemble of forecasts. The procedure aims at forecasting the cumulative distribution function of the observation which is simply a Heaviside function centered at the observed value. Our forecast is the weighted empirical cumulative distribution function based on the ensemble of forecasts. Each forecast of the ensemble is attributed a weight which is updated whenever new observations become available. The performance of the forecast is given by the continuous ranked probability score (CRPS), which is the square of the two-norm of the discrepancy between the forecast and the observed cumulative distribution functions. The method guarantees that, in the long run, the forecast cumulative distribution function has a continuous ranked probability score at least as good as the best weighted empirical cumulative function with weights constant in time.

The CRPS computed from an ensemble of forecasts is subject to a bias. We proposed a new way to compute the CRPS in order to mitigate the bias and obtain better aggregation performance.

The work was applied to the forecast of photovoltaics production, both at EDF production sites and for global France production.

7.3.2. Sensitivity analysis of air quality simulations at urban scale

Participants: Vivien Mallet, Louis Philippe, Fabien Brocheton [Numtech], David Poulet [Numtech].

We carried out a sensitivity analysis of the urban air quality model Sirane. We carried out dimension reduction on both inputs and outputs of the air quality model. This designed a reduced-order model, which we then emulated. We sampled the (reduced) inputs to the reduced model, and emulated the response surface of the reduced outputs. A metamodel was derived by the combination of the dimension reduction and the statistical emulation. This metamodel performs as well as the original model, compared to field observations. It is also extremely fast, which allowed us to compute Sobol' indices and carry out a complete sensitivity analysis.

7.3.3. Sensitivity analysis of road traffic simulations and corresponding emissions

Participants: Ruiwei Chen [École des Ponts ParisTech], Vivien Mallet, Vincent Aguiléra [Cerema], Fabien Brocheton [Numtech], David Poulet [Numtech], Florian Cohn [Numtech].

This work deals with the simulation of road traffic at metropolitan scale. We compared state-of-the-art static traffic assignment and dynamic traffic assignment, which better represents congestion. The work was applied in Clermont-Ferrand and its surrounding region, for a time period of two years, and using about 400 traffic loop counters for evaluation. The dynamic model showed similar overall performance as the static model.

We developed an open source software for the computation of the emissions of traffic. It computes the emissions of the main air pollutants, according the vehicle fleet.

For both traffic assignment and pollutant emissions, we carry out sensitivity tests with respect to limit speed, roads capacities or fleet composition. A complete sensitivity analysis is out of reach with the complete, computational intensive, traffic assignment model. Hence further work has been engaged with the metamodeling of the traffic assignment model. Preliminary results are encouraging and tend to show that a very fast metamodel can perform as well as the complete model.

7.3.4. Ensemble variational data assimilation

Participants: Julien Brajard, Isabelle Herlin, Marc Bocquet [CEREA], Jérôme Sirven [LOCEAN], Olivier Talagrand [LMD, ENS], Sylvie Thiria [LOCEAN].

The general objective of ensemble data assimilation is to produce an ensemble of analysis from observations and a numerical model which is representative of the uncertainty of the system. In a bayesian framework, the ensemble represents a sampling of the state vector probability distribution conditioned to the available knowledge of the system, denoted the a-posteriori probability distribution.

Ensemble variational data assimilation (EnsVar) consists in producing such an ensemble by perturbating N times the observations according to their error law, and run a standard variationnal assimilation for each perturbation. An ensemble of N members is then produced. In the case of linear models, there is a theoretical guarantee that this ensemble is a sampling of the a-posteriori probability. But there is no theoretical result in the non-linear case.

The objective of this work is to study the ability of EnsVar to produce "good" ensemble (i.e. that sampled the a posteriori probablility) on a shallow-water model. Statistical properties of the ensemble are evaluated, and the sensitivity to the main features of the assimilation system (number, distribution of observations, size of the assimilation window, ...) are also studied.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

A new Iilab, Rain_Water, has been accepted in 2016. It concerns joint research with the company Weather Measures. Rain_Water aims to define a platform of local meteorology. Users are mainly farmers that will use the platform for monitoring the agricultural practices at the parcel level.

9. Partnerships and Cooperations

9.1. National Initiatives

- The ANR project Estimair aims at quantifying the uncertainties of air quality simulations at urban scale. The propagation of uncertainties requires the use of model reduction and emulation. A key uncertainty source lies in the traffic emissions, which are generated using a dynamic traffic assignment model. Ensembles of traffic assignments are calibrated and used in the uncertainty quantification. Estimair is led by Clime.
- The IPSL project "AVES" (Ensemble Variational Assimilation applied to a shallow-water model) aims at estimating the quality of an ensemble produced by a variational ensemble algorithm applied on a shallow-water numerical model. A focus is made on the bayesian properties of the ensemble, i.e. its capacity to sample the a-posteriori probability law of the model state.
- Two new ANR projects have been accepted in 2016 and will begin in January 2017.
  FireCaster aims at fire forecasting and risk mitigation.
  Cense aims at the estimation of urban noise, using numerical simulation and a dense monitoring network.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. EoCoE

Title: Energy oriented Centre of Excellence for computer applications

Program: H2020

Duration: October 2015 - October 2018

Coordinator: CEA

Partners:

CEA, Commissariat à l'Énergie Atomique et aux Énergies Alternatives (France)

Forschungszentrum Julich (Germany)

Max Planck Gesellschaft (Germany)

ENEA, Agenzia Nazionale Per le Nuove Tecnologie, l'energia E Lo Sviluppo Economico Sostenibile (Italy)

CERFACS, European Centre for Research and Advanced Training in Scientific Computing (France)

Instytut Chemii Bioorganicznej Polskiej Akademii Nauk (Poland)

Universita Degli Studi di Trento (Italy)

Fraunhofer Gesellschaft (Germany)

University of Bath (United Kingdom)

CYL, The Cyprus Institute (Cyprus)

CNR, National Research Council of Italy (Italy)

Université Libre de Bruxelles (Belgium)

BSC, Barcelona Supercomputing Center - Centro Nacional de Supercomputacion (Spain)

Inria contact: Michel Kern (Serena team)

The aim of the projevt is to establish an Energy Oriented Centre of Excellence for computing applications (EoCoE). EoCoE (pronounce "Echo") will use the prodigious potential offered by the ever-growing computing infrastructure to foster and accelerate the European transition to a reliable and low carbon energy supply. To achieve this goal, we believe that the present revolution in hardware technology calls for a similar paradigm change in the way application codes are designed. EoCoE will assist the energy transition via targeted support to four renewable energy pillars: Meteo, Materials, Water and Fusion, each with a heavy reliance on numerical modelling. These four pillars will be anchored within a strong transversal multidisciplinary basis providing high-end expertise in applied mathematics and HPC. EoCoE is structured around a central Franco-German hub coordinating a pan-European network, gathering a total of 8 countries and 23 teams. Its partners are strongly engaged in both the HPC and energy fields; a prerequisite for the long-term sustainability of EoCoE and also ensuring that it is deeply integrated in the overall European strategy for HPC. The primary goal of EoCoE is to create a new, long lasting and sustainable community around computational energy science. At the same time, EoCoE is committed to deliver high-impact results within the first three years. It will resolve current bottlenecks in application codes, leading to new modelling capabilities and scientific advances among the four user communities; it will develop cutting-edge mathematical and numerical methods, and tools to foster the usage of Exascale computing. Dedicated services for laboratories and industries will be established to leverage this expertise and to foster an ecosystem around HPC for energy. EoCoE will give birth to new collaborations and working methods and will encourage widely spread best practices.

9.2.1.2. Env&You 2016

Title: Env&You

Program: EIT Digital

Duration: January 2016 - December 2016

Coordinator: Inria (MiMove)

Partners and subgrantees:

Inria

NUMTECH

Cap Digital

Forum Virium (Finland)

TheCivicEngine (United States)

Ambientic

Inria contact: Valérie Issarny (Mimove project-team)

Env&You aims at delivering the whole picture of urban pollution, from the individual exposure to neighborhood-by-neighborhood and day-to-day variation, to citizens and governments, informing their decisions for healthy urban living. There is a clear, and probably increasing, desire from the citizens to better know their individual exposure to pollution. Partial solutions exist to the exposure data problem but each focuses on one or another domain of information—crowdsourcing exposure, translating governmental open data to usable consumer information, harnessing social media information, harnessing biometrics—what is unique about Env&You is that it will assimilate a multi-dimensional picture of exposure and provide the integrated information to citizen, government, and business use.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

Partner: Marine Hydrophysical Institute, Ukraine.

The collaboration concerns the study of the Black Sea surface circulation and the issue of image assimilation in forecasting models.

Partner: IBM Research, Dublin, Ireland

The collaboration addresses the assimilation of classical observations as well as images, with application to geophysics. New assimilation methods are developed, mainly based on minimax filtering.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Selection

10.1.1.1. Reviewer

- Isabelle Herlin: International Conference on Computer Vision and Pattern Recognition (CVPR), International Conference on Computer Vision (ECCV), Asian Conference on Comuter Vision (ACCV), International Conference on Image Processing (ICIP).

10.1.2. Journal

10.1.2.1. Reviewer - Reviewing Activities

- Julien Brajard: IEEE Transactions on Geoscience and Remote Sensing (TGRS).
- Isabelle Herlin: IEEE Transactions on Geoscience and Remote Sensing (TGRS), IEEE Geoscience and Remote Sensing Letters, Mathematical Methods in Applied Sciences.

10.1.3. Invited Talks

- Vivien Mallet: Processing environmental simulations and observations for smart cities; Perspectives and New Challenges in Data Science, École des Ponts ParisTech; February 2016.
- Vivien Mallet: Assimilation de données et prévision d'ensemble appliquées à la qualité de l'air; CEA seminar; June 2016.

10.1.4. Scientific Expertise

- Isabelle Herlin is a member of the Scientific Council of ANDRA (French national radioactive waste management agency)
- Isabelle Herlin is a member of the Scientific Council of CSFRS (High Council for Strategic Education and Research in France).
- Isabelle Herlin and Vivien Mallet reviewed several research proposals, especially for ANR (France).

10.1.5. Research Administration

- Isabelle Herlin is a member of the program committee of DIGITEO, French research cluster in science and technology of information.
- Isabelle Herlin is a member of the Evaluation Committee at Inria.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master OACOS/WAPE: Marc Bocquet, Vivien Mallet, Jean-Matthieu Haussaire; Introduction to Data Assimilation for Geophysics; 12 hours; M2; UPMC, X, ENS, ENSTA ParisTech, École des Ponts ParisTech; France.

Master SGE and 3rd-year class at École des Ponts ParisTech: Vivien Mallet; Air quality modeling; 4.5 hours; M2; Universities Paris Diderot- Paris 7, Paris 12 and École des Ponts ParisTech, France.

Training: Vivien Mallet; Uncertainty Quantification: Ensembles and Data Assimilation – Application to Climate and Geosciences; 5.25 hours; CERFACS; France.

Training: Vivien Mallet; Introduction to data assimilation: Kalman filters and ensembles; 4.5 hours; CEMRACS (summer school).

10.2.2. Supervision

- PhD in progress : Pacôme Eberhart, "Génération automatique de codes performants et fiables pour l'assimilation de données", September 2013, Fabienne Jezequel, Pierre Fortin and Julien Brajard
- PhD in progress : Jean Thorey, "Prévision d'ensemble du rayonnement solaire pour la production photovoltaïque du parc EDF", November 2013, Vivien Mallet.
- PhD in progress: Ruiwei Chen, "Quantification d'incertitude en simulation des émissions du trafic routier", November 2014, Vivien Mallet.
- PhD in progress: Raphaël Ventura, "Simulation numérique de la ville par couplage entre la modélisation et l'observation", September 2014, Vivien Mallet.
- PhD in progress: Ngoc Bao Tran Le, "Quantification d'incertitude par réduction de modèle de dispersion atmosphérique", November 2016, Vivien Mallet.

10.2.3. Juries

- Vivien Mallet for the PhD defense of Michaël Zamo, “Statistical Post-processing of Deterministic and Ensemble Windspeed Forecasts on a Grid”, December 2016, Météo France.

10.3. Popularization

Vivien Mallet and Raphaël Ventura, together with Cong Kinh Nguyen (MiMove), Pierre-Guillaume Raverdy (SED), Fadwa Rebhi (MiMove), Fabienne Giboudeaux (Paris city), Awa Ndiaye (Paris city), Gilles Plattner (Particitae) and Laure Turcati (Particitae), took part to collaborative measuring of noise pollution with volunteers from the general public. Using the mobile application Ambiciti, the volunteers carried out a journey in a city district in order to measure noise pollution with their smartphones, and therefore help to map noise pollution in the area. At the end of the journeys, the observations were merged with an existing noise map so as to produce an improved map. Discussion about the strengths and limitations of mobile observation were engaged, and the activity helped to inform about noise pollution. This was carried out in Paris at the opening of the “Canopée des Halles”, at Futur en Seine festival and during the “journée sans voiture”.

11. Bibliography

Major publications by the team in recent years

[1] M. BOCQUET, P. SAKOV.*An iterative ensemble Kalman smoother*, in "Quarterly Journal of the Royal Meteorological Society", October 2013 [*DOI :* 10.1002/QJ.2236], http://hal.inria.fr/hal-00918488.

[2] D. BÉRÉZIAT, I. HERLIN.*Solving ill-posed Image Processing problems using Data Assimilation*, in "Numerical Algorithms", February 2011, vol. 56, n^{o} 2, p. 219-252 [*DOI :* 10.1007/S11075-010-9383-Z], http://hal.inria.fr/inria-00538510.

[3] D. GARAUD, V. MALLET.*Automatic calibration of an ensemble for uncertainty estimation and probabilistic forecast: Application to air quality*, in "Journal of Geophysical Research", October 2011, vol. 116 [*DOI :* 10.1029/2011JD015780], http://hal.inria.fr/hal-00655771.

[4] I. HERLIN, D. BÉRÉZIAT, N. MERCIER, S. ZHUK.*Divergence-Free Motion Estimation*, in "ECCV 2012 - European Conference on Computer Vision", Florence, Italie, A. FITZGIBBON, S. LAZEBNIK, P. PERONA, Y. SATO, C. SCHMID (editors), Lecture Notes in Computer Science, Springer, October 2012, vol. 7575, p. 15-27 [*DOI :* 10.1007/978-3-642-33765-9_2], http://hal.inria.fr/hal-00742021.

[5] M. R. KOOHKAN, M. BOCQUET.*Accounting for representativeness errors in the inversion of atmospheric constituent emissions: application to the retrieval of regional carbon monoxide fluxes*, in "Tellus B", July 2012, vol. 64, n^{o} 19047 [*DOI :* 10.3402/TELLUSB.V64I0.19047], http://hal.inria.fr/hal-00741930.

[6] G. K. KOROTAEV, E. HUOT, F.-X. LE DIMET, I. HERLIN, S. STANICHNY, D. SOLOVYEV, L. WU.*Retrieving ocean surface current by 4-D variational assimilation of sea surface temperature images*, in "Remote Sensing of Environment", April 2008, vol. 112, n^{o} 4, p. 1464-1475, Remote Sensing Data Assimilation Special Issue [*DOI :* 10.1016/J.RSE.2007.04.020], http://hal.inria.fr/hal-00283896.

[7] V. MALLET.*Ensemble forecast of analyses: Coupling data assimilation and sequential aggregation*, in "Journal of Geophysical Research", December 2010, vol. 115 [*DOI :* 10.1029/2010JD014259], http://hal.inria.fr/inria-00547903.

[8] B. SPORTISSE.*Pollution atmosphérique. Des processus à la modélisation*, Ingénierie et développement durable, Springer-Verlag France, 2008, 350, http://hal.inria.fr/inria-00581172.

[9] V. WINIAREK, M. BOCQUET, O. SAUNIER, A. MATHIEU.*Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the cesium-137 and iodine-131 source terms from the Fukushima Daiichi power plant*, in "Journal of Geophysical Research Atmospheres", March 2012, vol. 117 [*DOI :* 10.1029/2011JD016932], http://hal.inria.fr/hal-00704999.

[10] L. WU, V. MALLET, M. BOCQUET, B. SPORTISSE.*A comparison study of data assimilation algorithms for ozone forecasts*, in "Journal of Geophysical Research", October 2008, vol. 113 [*DOI :* 10.1029/2008JD009991], http://hal.inria.fr/inria-00582376.

Publications of the year

Articles in International Peer-Reviewed Journal

[11] L. ISSA, J. BRAJARD, M. FAKHRI, D. HAYES, L. MORTIER, P.-M. POULAIN.*Modelling surface currents in the Eastern Levantine Mediterranean using surface drifters and satellite altimetry*, in "Ocean Modelling", 2016, vol. 104, p. 1-14 [*DOI :* 10.1016/J.OCEMOD.2016.05.006], https://hal.inria.fr/hal-01357669.

International Conferences with Proceedings

[12] D. BÉRÉZIAT, I. HERLIN, Y. LEPOITTEVIN.*Ensemble Kalman Filter based on the image structures*, in "International Conference on Computer Vision Theory and Applications", Porto, Portugal, February 2017, http://hal.upmc.fr/hal-01414786.

[13] P. EBERHART, J. BRAJARD, P. FORTIN, F. JÉZÉQUEL.*Estimation of Round-off Errors in OpenMP Codes*, in "IWOMP 2016 - 12th International Workshop on OpenMP", Nara, Japan, Lecture Notes in Computer Science, Springer International Publishing, October 2016, vol. 9903, p. 3-16 [*DOI :* 10.1007/978-3-319-45550-1_1], https://hal.inria.fr/hal-01380131.

[14] V. ISSARNY, V. MALLET, K. NGUYEN, P.-G. RAVERDY, F. REBHI, R. VENTURA.*Dos and Don'ts in Mobile Phone Sensing Middleware: Learning from a Large-Scale Experiment*, in "ACM/IFIP/USENIX Middleware 2016", Trento, Italy, Proceedings of the 2016 International Middleware Conference, December 2016 [*DOI :* 10.1145/2988336.2988353], https://hal.inria.fr/hal-01366610.

[15] Y. LEPOITTEVIN, I. HERLIN.*Regularization terms for motion estimation. Links with spatial correlations*, in "VISAPP - International Conference on Computer Vision Theory and Applications", Rome, Italy, Springer Verlag, February 2016, vol. 3, p. 458-466 [*DOI :* 10.5220/0005712104560464], https://hal.inria.fr/hal-01235718.

[16] V. MALLET, I. HERLIN.*Quantification of uncertainties from ensembles of simulations*, in "International Meeting "Foreknowledge Assessment Series"", Paris, France, December 2016, https://hal.inria.fr/hal-01399175.

Other Publications

[17] I. HERLIN, E. HUOT.*Image-Based Forecast of the Surface Circulation of Black Sea*, September 2016, Black Sea from Space Workshop, Poster, https://hal.inria.fr/hal-01349391.

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:

1.2.4. - QoS, performance evaluation
6.1.4. - Multiscale modeling
6.2.3. - Probabilistic methods
7.2. - Discrete mathematics, combinatorics
7.3. - Optimization
7.5. - Geometry, Topology
7.8. - Information theory
7.9. - Graph theory
7.10. - Network science
7.11. - Performance evaluation

Other Research Topics and Application Domains:

4.3. - Renewable energy production
6.2.2. - Radio technology
6.3.4. - Social Networks

1. Members

Research Scientists

Marc Lelarge [Team leader, Inria, Senior Researcher, HDR]
François Baccelli [Inria, Senior Researcher, part time, HDR]
Bartlomiej Blaszczyszyn [Inria, Senior Researcher, HDR]
Ana Busic [Inria, Researcher]
Francesco Caltagirone [Inria, Starting Research Position]

Faculty Members

Anne Bouillard [ENS Paris, Associate Professor, HDR]
Jocelyne Elias [Univ. Paris V, Associate Professor, by delegation]

Technical Staff

Virag Shah [Inria]

PhD Students

Lennart Gulikers [Inria]
Md Umar Hashmi [PSL]
Dalia-Georgiana Herculea [Nokia Bell Labs, granted by CIFRE]
Alexandre Hollocou [Min. de la Défense]
Leo Miolane [Ecole Polytechnique, from Sep 2016]
Christelle Rovetta [Univ. Paris VI, granted by ANR DYOGEN-MARMOTE- project]
Sébastien Samain [Inria, from Nov 2016]
Rémi Varloot [Inria]

Post-Doctoral Fellows

Arpan Chattopadhyay [Inria, until Oct 2016]
Arpan Mukhopadhyay [Inria, from Mar 2016]

Visiting Scientists

Adithya Munegowda Devraj [University of Florida, from May 2016 until Jul 2016]
Sean Meyn [University of Florida, Jun 2016]

Administrative Assistant

Helene Milome [Inria]

Others

Pierre Bremaud [Professor emeritus]
Eric Tramel [Inria, Engineer from Nov 2016]

2. Overall Objectives

2.1. Overall Objectives

A large number of real-world structures and phenomena can be described by networks: separable elements with connections between certain pairs of them. Among such networks, the best known and the most studied in computer science is the Internet. Moreover, the Internet (as the physical underlying network) gives itself rise to many new networks, like the networks of hyperlinks, Internet based social networks, distributed data bases, codes on graphs, local interactions of wireless devices. These huge networks pose exciting challenges for the mathematician and the mathematical theory of networks faces novel, unconventional problems. For example, very large networks cannot be completely known, and data about them can be collected only by indirect means like random local sampling or by monitoring the behavior of various aggregated quantities.

The scientific focus of DYOGENE is 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.

3. Research Program

3.1. Network Calculus

Network calculus [53] is a theory for obtaining deterministic upper bounds in networks that has been developed by R. Cruz [41], [42]. From the modelling point of view, it is an algebra for computing and propagating constraints given in terms of envelopes. A flow is represented by its cumulative function $R(t)$ (that is, the amount of data sent by the flow up to time t). A constraint on a flow is expressed by an arrival curve $\alpha(t)$ that gives an upper bound for the amount of data that can be sent during any interval of length t. Flows cross service elements that offer guarantees on the service. A constraint on a service is a service curve $\beta(t)$ that is used to compute the amount of data that can be served during an interval of length t. It is also possible to define in the same way minimal arrival curves and maximum service curves. Then such constraints envelop the processes and the services. Network calculus enables the following operations:

- computing the exact output cumulative function or at least bounding functions;
- computing output constraints for a flow (like an output arrival curve);

• computing the remaining service curve (that is, the service that of not used by the flows crossing a server);
• composing several servers in tandem;
• giving upper bounds on the worst-case delay and backlog (bounds are tight for a single server or a single flow).
The operations used for this are an adaptation of filtering theory to $(\min, +)$: $(\min, +)$ convolution and deconvolution, sub-additive closure.

We investigate the complexity of computing exact worst-case performance bounds in network calculus and to develop algorithms that present a good trade off between algorithmic efficiency and accuracy of the bounds.

3.2. Perfect Simulation

Simulation approaches can be used to efficiently estimate the stationary behavior of Markov chains by providing independent samples distributed according to their stationary distribution, even when it is impossible to compute this distribution numerically.

The classical Markov Chain Monte Carlo simulation techniques suffer from two main problems:
• The convergence to the stationary distribution can be very slow, and it is in general difficult to estimate;
• Even if one has an effective convergence criterion, the sample obtained after any finite number of iterations is biased.

To overcome these issues, Propp and Wilson [56] have introduced a perfect sampling algorithm (PSA) that has later been extended and applied in various contexts, including statistical physics [47], stochastic geometry [52], theoretical computer science [33], and communications networks [30], [46] (see also the bibliography at http://dimacs.rutgers.edu/~dbwilson/exact.html/ annotated by David B. Wilson.

Perfect sampling uses coupling arguments to give an unbiased sample from the stationary distribution of an ergodic Markov chain on a finite state space $\mathcal{X}$. Assume the chain is given by an update function Φ and an i.i.d. sequence of innovations $(U_n)_{n\in\mathbb{Z}}$, so that

$$X_{n+1} = \Phi(X_n, U_{n+1}). \tag{1}$$

The algorithm is based on a backward coupling scheme: it computes the trajectories from all $x \in \mathcal{X}$ at some time in the past $t = -T$ until time $t = 0$, using the same innovations. If the final state is the same for all trajectories (i.e. $|\{\Phi(x, U_{-T+1}, ..., U_0) : x \in \mathcal{X}\}| = 1$, where $\Phi(x, U_{-T+1}, ..., U_0) := \Phi(\Phi(x, U_{-T+1}), U_{-T+2}, ..., U_0)$ is defined by induction on T), then we say that the chain has globally coupled and the final state has the stationary distribution of the Markov chain. Otherwise, the simulations are started further in the past.

Any ergodic Markov chain on a finite state space has a representation of type (1) that couples in finite time with probability 1, so Propp and Wilson's PSA gives a "perfect" algorithm in the sense that it provides an *unbiased* sample in *finite time*. Furthermore, the stopping criterion is given by the coupling from the past scheme, and knowing the explicit bounds on the coupling time is not needed for the validity of the algorithm.

However, from the computational side, PSA is efficient only under some monotonicity assumptions that allow reducing the number of trajectories considered in the coupling from the past procedure only to extremal initial conditions. Our goal is to propose new algorithms solving this issue by exploiting semantic and geometric properties of the event space and the state space.

3.3. Stochastic Geometry

Stochastic geometry [40] is a rich branch of applied probability which allows one to quantify random phenomena on the plane or in higher dimension. It is intrinsically related to the theory of point processes. Initially its development was stimulated by applications to biology, astronomy and material sciences. Nowadays it is also widely used in image analysis. It provides a way of estimating and computing “spatial averages”. A typical example, with obvious communication implications, is the so called Boolean model, which is defined as the union of discs with random radii (communication ranges) centered at the points of a Poisson point process (user locations) of the Euclidean plane (e.g., a city). A first typical question is that of the prediction of the fraction of the plane which is covered by this union (statistics of coverage). A second one is whether this union has an infinite component or not (connectivity). Further classical models include shot noise processes and random tessellations. Our research consists of analyzing these models with the aim of better understanding wireless communication networks in order to predict and control various network performance metrics. The models require using techniques from stochastic geometry and related fields including point processes, spatial statistics, geometric probability, percolation theory.

F. Baccelli, B. Blaszczyszyn in collaboration with M. Karray (Orange Labs) are preparing a new book focusing on the mathematical tools at the basis of stochastic geometry. The book will cover the main mathematical foundations of the field, namely the theory of point processes and random measures as well as the theory of random closed sets. The basis will be the graduate classes and the research courses taught by the authors at a variety of places worldwide.

The collaboration of F. Baccelli with V. Anantharam (UC Berkeley) continues in new directions on high dimensional stochastic geometry, primarily in relation with Information Theory, cf. Section 7.23.

The collaboration of B. Blaszczyszyn with D. Yogeshwaran (Indian Statistical Institute) and Y. Yukich (Lehigh University) led to the development of the limit theory for geometric statistics on general input processes, cf. Section 7.22.

3.4. Information Theory and Wireless Networks

Classical models of stochastic geometry (SG) are not sufficient for analyzing wireless networks as they ignore the specific nature of radio channels.

Consider a wireless communication network made of a collection of nodes which in turn can be transmitters or receivers. At a given time, some subset of this collection of nodes simultaneously transmit, each toward its own receiver. Each transmitter–receiver pair in this snapshot requires its own wireless link. For each such wireless link, the power of the signal received from the link transmitter is jammed by the powers of the signals received from the other transmitters. Even in the simplest model where the power radiated from a point decays in some isotropic way with Euclidean distance, the geometry of the location of nodes plays a key role within this setting since it determines the signal to interference and noise ratio (SINR) at the receiver of each such link and hence the possibility of establishing simultaneously this collection of links at a given bit rate, as shown by information theory (IT). In this definition, the interference seen by some receiver is the sum of the powers of the signals received from all transmitters excepting its own. The SINR field, which is of an essentially geometric nature, hence determines the connectivity and the capacity of the network in a broad sense. The essential point here is that the characteristics and even the feasibilities of the radio links that are simultaneously active are strongly interdependent and determined by the geometry. Our work is centered on the development of an IT-aware stochastic geometry addressing this interdependence. Dyogene members published in 2009 a two-volume book [1], [2] on Stochastic Geometry and Wireless Networks that became a reference publication in this domain.

In collaboration with Martin Haenggi (University of Notre Dame Notre Dame, IN, USA), Paul Keeler (Weierstrass Institute for Applied Analysis and Stochastics Berlin, Germany) and Sayandev Mukherjee (DOCOMO Innovations, Inc. Palo Alto, CA, USA), B. Blaszczyszyn is currently working on a book project that is intended to bridge a gap between academic and industrial approach to the design of next-generation cellular networks. In fact, simulation-only approach adopted by a majority of industry practitioners does not scale up with the increasing network complexity and analytical treatment is still yet not widely accepted in various bodies working out future standards specifications. The monograph is intended to bridge that gap, and make the methods, tools, approaches, and results of stochastic geometry available to a wide group of researchers (both in academia and in industry), systems engineers, and network designers. We expect that academic researchers and graduate students will appreciate that the book collects and organizes the most recent research results in a convenient way.

3.5. The Cavity Method for Network Algorithms

The cavity method combined with geometric networks concepts has recently led to spectacular progresses in digital communications through error-correcting codes. More than fifty years after Shannon's theorems, some coding schemes like turbo codes and low-density parity-check codes (LDPC) now approach the limits predicted by information theory. One of the main ingredients of these schemes is message-passing decoding strategies originally conceived by Gallager, which can be seen as direct applications of the cavity method on a random bipartite graph (with two types of nodes representing information symbols and parity check symbols, see [57]).

Modern coding theory is only one example of application of the cavity method. The concepts and techniques developed for its understanding have applications in theoretical computer science and a rich class of *complex systems*, in the field of networking, economics and social sciences. The cavity method can be used both for the analysis of randomized algorithms and for the study of random ensembles of computational problems representative real-world situations. In order to analyze the performance of algorithms, one generally defines a family of instances and endows it with a probability measure, in the same way as one defines a family of samples in the case of spin glasses or LDPC codes. The discovery that the hardest-to-solve instances, with all existing algorithms, lie close to a *phase transition* boundary has spurred a lot of interest. Theoretical physicists suggest that the reason is a structural one, namely a change in the geometry of the set of solutions related to the *replica symmetry breaking* in the cavity method. Phase transitions, which lie at the core of statistical physics, also play a key role in computer science [60], signal processing [44] and social sciences [49]. Their analysis is a major challenge, that may have a strong impact on the design of related algorithms.

We develop mathematical tools in the theory of discrete probabilities and theoretical computer science in order to contribute to a rigorous formalization of the cavity method, with applications to network algorithms, statistical inference, and at the interface between computer science and economics (EconCS).

3.6. Statistical Learning

Sparse graph structures are useful in a number of information processing tasks where the computational problem can be described as follows: infer the values of a large collection of random variables, given a set of constraints or observations, that induce relations among them. Similar design ideas have been proposed in sensing and signal processing and have applications in coding [38], network measurements, group testing or multi-user detection. While the computational problem is generally hard, sparse graphical structures lead to low-complexity algorithms that are very effective in practice. We develop tools in order to contribute to a precise analysis of these algorithms and of their gap to optimal inference which remains a largely open problem.

A second line of activities concerns the design of protocols and algorithms enabling a transmitter to learn its environment (the statistical properties of the channel quality to the corresponding receiver, as well as their interfering neighbouring transmitters) so as to optimise their transmission strategies and to fairly and efficiently share radio resources. This second objective calls for the development and use of machine learning techniques (e.g. bandit optimisation).

4. Application Domains

4.1. Wireless Networks

Wireless networks can be efficiently modelled as dynamic stochastic geometric networks. Their analysis requires taking into account, in addition to their geometric structure, the specific nature of radio channels and their statistical properties which are often unknown a priori, as well as the interaction through interference of the various individual point-to-point links. Established results contribute in particular to the development of network dimensioning methods and some of them are currently used in Orange internal tools for network capacity calculations.

4.2. Embedded Networks

Critical real-time embedded systems (cars, aircrafts, spacecrafts) are nowadays made up of multiple computers communicating with each other. The real-time constraints typically associated with operating systems now extend to the networks of communication between sensors/actuators and computers, and between the computers themselves. Once a media is shared, the time between sending and receiving a message depends not only on technological constraints, but also, and mainly from the interactions between the different streams of data sharing the media. It is therefore necessary to have techniques to guarantee maximum network delays, in addition to local scheduling constraints, to ensure a correct global real-time behaviour to distributed applications/functions.

Moreover, pessimistic estimate may lead to an overdimensioning of the network, which involves extra weight and power consumption. In addition, these techniques must be scalable. In a modern aircraft, thousands of data streams share the network backbone. Therefore algorithm complexity should be at most polynomial.

4.3. Distributed Content Delivery Networks

A content distribution network (CDN) is a globally distributed network of proxy servers deployed in multiple data centers. The goal of a CDN is to serve content to end-users with high availability and high performance. CDNs serve a large fraction of the Internet content today, including web objects (text, graphics and scripts), downloadable objects (media files, software, documents), applications (e-commerce, portals), live streaming media, on-demand streaming media, and social networks.

A. Bouillard and F. Baccelli started a collaboration with Virag Shah (Postdoc at the Inria-Microsoft Saclay center) on the analysis of delays in data clusters. Their focus is on the way delays scale with the size of a request and on the way delays compare under different policies for coding, data dissemination, and delivery. A paper on the matter is submitted.

4.4. Probabilistic Algorithms for Renewable Integration in Smart Grids

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 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. 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.

A challenge in residential communities is that many loads are either on or off. How can an on/off load track the continuously varying regulation signal broadcast by a grid operator? The answer proposed in our recent work is based on probabilistic algorithms: A single load cannot track a regulation signal such as the balancing reserves. A collection of loads can, provided they are equipped with local control. The value of probabilistic algorithms is that a) they can be designed with minimal communication, b) they avoid synchronization of load responses, and c) it is shown in our recent work that they can be designed to simplify control at the grid level (see the survey [31] and [54], [39]).

This research is developed within the Inria Associate Team PARIS.

4.5. Algorithms for Finding Communities

In the study of complex networks, a network is said to have community structure if the nodes of the network can be easily grouped into (potentially overlapping) sets of nodes such that each set of nodes is densely connected internally. Community structures are quite common in real networks. Social networks include community groups (the origin of the term, in fact) based on common location, interests, occupation, etc. Metabolic networks have communities based on functional groupings. Citation networks form communities by research topic. Being able to identify these sub-structures within a network can provide insight into how network function and topology affect each other. We propose several algorithms for this problem and extensions [50], [58], [32], [59]

4.6. Mean-Field Limits for Queuing Networks with Node Motion

The work with S. Rybko, S. Vladimorov (IPIT, Moscow) and S. Shlosman (CNRS Marseille) which started through some funding from CNRS and which led to several visits of S. Rybko and S. Vladimorov in Paris led to a series of research projects on queuing theory. The first one, on mean-fields for networks with node motion [5] was published in 2016; cf. Section 7.3.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

F. Baccelli received a Honorary Doctorate of Heriot-Watt University. The graduation took place on November 17, 2016, in Edinburgh, United Kingdom.

6. New Software and Platforms

6.1. CloNES

CLOsed queueing Networks Exact Sampling
FUNCTIONAL DESCRIPTION

Clones is a Matlab toolbox for exact sampling of closed queueing networks.

- Participant: Christelle Rovetta
- Contact: Christelle Rovetta
- URL: http://www.di.ens.fr/~rovetta/Clones/index.html

7. New Results

7.1. Fast Weak KAM Integrators for Separable Hamiltonian Systems

In [7], we consider a numerical scheme for Hamilton–Jacobi equations based on a direct discretization of the Lax–Oleinik semi-group. We prove that this method is convergent with respect to the time and space stepsizes provided the solution is Lipschitz, and give an error estimate. Moreover, we prove that the numerical scheme is a *geometric integrator* satisfying a discrete weak–KAM theorem which allows to control its long time behavior. Taking advantage of a fast algorithm for computing min–plus convolutions based on the decomposition of the function into concave and convex parts, we show that the numerical scheme can be implemented in a very efficient way.

7.2. Low Complexity State Space Representation and Algorithms for Closed Queueing Networks Exact Sampling

In [6] we consider exact sampling from the stationary distribution of a closed queueing network with finite capacities. In a recent work a compact representation of sets of states was proposed that enables exact sampling from the stationary distribution without considering all initial conditions in the coupling from the past (CFTP) scheme.This representation reduces the complexity of the one-step transition in the CFTP algorithm to O(KM^2), where K is the number of queues and M the total number of customers; while the cardinality of the state space is exponential in the number of queues. In this paper, we extend these previous results to the multiserver case. The main focus and the contribution of this work is on the algorithmic and the implementation issues. We propose a new representation, that leads to one-step transition complexity of the CFTP algorithm that is in $O(KM)$. We provide a detailed description of our matrix-based implementation. Matlab toolbox Clones (CLOsed queueing Networks Exact Sampling) can be downloaded at http://www.di.ens.fr/~rovetta/Clones

7.3. Queueing Networks with Mobile Servers: The Mean-Field Approach

In [5] we consider queueing networks which are made from servers exchanging their positions on a graph. When two servers exchange their positions, they take their customers with them. Each customer has a fixed destination. Customers use the network to reach their destinations, which is complicated by movements of the servers. We develop the general theory of such networks and establish the convergence of the symmetrized version of such a network to some nonlinear Markov process.

7.4. Distributed Randomized Control for Demand Dispatch

This work, reported in [14], concerns design of control systems for Demand Dispatch to obtain ancillary services to the power grid by harnessing inherent flexibility in many loads. The role of "local intelligence" at the load has been advocated in prior work, randomized local controllers that manifest this intelligence are convenient for loads with a finite number of states. The present work introduces two new design techniques for these randomized controllers: (i) The Individual Perspective Design (IPD) is based on the solution to a one-dimensional family of Markov Decision Processes, whose objective function is formulated from the point of view of a single load. The family of dynamic programming equation appears complex, but it is shown that it is obtained through the solution of a single ordinary differential equation. (ii) The System Perspective Design (SPD) is motivated by a single objective of the grid operator: Passivity of any linearization of the aggregate input-output model. A solution is obtained that can again be computed through the solution of a single ordinary differential equation. Numerical results complement these theoretical results.

7.5. Smart Fridge / Dumb Grid? Demand Dispatch for the Power Grid of 2020

In our previous research [31], it was argued that loads can provide most of the ancillary services required today and in the future. Through load-level and grid-level control design, high-quality ancillary service for the grid is obtained without impacting quality of service delivered to the consumer. This approach to grid regulation is called demand dispatch: loads are providing service continuously and automatically, without consumer interference. In [19] work we investigate what intelligence is required at the grid-level. In particular, does the grid-operator require more than one-way communication to the loads? Our main conclusion: risk is not great in lower frequency ranges, e.g., PJM's RegA or BPA's balancing reserves. In particular, ancillary services from refrigerators and pool-pumps can be obtained successfully with only one-way communication. This requires intelligence at the loads, and much less intelligence at the grid level.

7.6. Efficient Orchestration Mechanisms for Congestion Mitigation in Network Functions Virtualization: Models and Algorithms

Nowadays, telecommunication infrastructures are composed of property hardware operated by a single entity to offer communication services to their final users. While this architecture simplifies the design and optimization of the network equipment for specific tasks, its low degree of flexibility represents the main limitation for the evolution of the network infrastructure. For this reason, network operators and equipment manufacturers have started the standardization process of a plethora of virtualization solutions that have been individually developed in recent years for enabling the sharing of general-purpose resources and increasing the flexibility of their network architectures. Such a process has led to the specification of the Network Functions Virtualization (NFV) technology, which promises to bring about several benefits, such as reduced CAPEX and OPEX (CAPital and OPerational EXpenditure), low time-tomarket for new network services, higher flexibility to scale up and down the services according to users' demand, simple and cheap testing of new services. Nevertheless, the consolidation of the virtualization technology represents one of the main challenging problems for its success and widespread utilization in telecommunication infrastructures, which still consist of a huge set of property hardware appliances and software systems. Indeed, the sharing of the physical infrastructure among multiple virtual operators as well as the simple configuration of network services require the design of complex management mechanisms for the orchestration of the network equipment, with the final goal of dynamically adapting the infrastructure to the resource utilization.

In particular, spatio-temporal correlation of traffic demands and computational loads can result in high congestion and low network performance for virtual operators, thus leading to service level agreement breaches. In [10], we propose novel orchestration mechanisms to optimally control and mitigate the resource congestion of a physical infrastructure based on the NFV paradigm. More specifically, we analyze the congestion resulting from the sharing of the physical infrastructure and propose innovative orchestration mechanisms based on both centralized and distributed approaches, aimed at unleashing the potential of the NFV technology. In particular, we first formulate the network functions composition problem as a non-linear optimization model to accurately capture the congestion of physical resources. To further simplify the network management, we also propose a dynamic pricing strategy of network resources, proving that the resulting system achieves a stable equilibrium in a completely distributed fashion, even when all virtual operators independently select their best network configuration. Numerical results show that the proposed approaches consistently reduce resource congestion. Furthermore, the distributed solution well approaches the performance that can be achieved using a centralized network orchestration system.

7.7. Optimal Planning of Virtual Content Delivery Networks under Uncertain Traffic Demands

Content Delivery Networks (CDNs) have been identified as one of the relevant use cases where the emerging paradigm of Network Functions Virtualization (NFV) will likely be beneficial. In fact, virtualization fosters flexibility, since on-demand resource allocation of virtual CDN nodes can accommodate sudden traffic demand changes. However, there are cases where physical appliances should still be preferred, therefore we envision a mixed architecture in between these two solutions, capable to exploit the advantages of both of them. Motivated by these reasons, in [13] we formulate a two-stage stochastic planning model that can be used by CDN operators to compute the optimal long-term network planning decision, deploying physical CDN appliances in the network and/or leasing resources for virtual CDN nodes in data centers. Key findings demonstrate that for a large range of pricing options and traffic profiles, NFV can significantly save network costs spent by the operator to provide the content distribution service.

7.8. Distributed Spectrum Management in TV White Space Networks

The radio frequency (RF) spectrum is a scarce resource that has recently become particularly critical with the increased wireless demand. For this reason, the Federal Communications Commission (FCC) has recently allowed for opportunistic access to the unused spectrum in the TV bands (also called "white space"). With opportunistic access, however, there is a need to deploy enhanced channel allocation and power control techniques to mitigate interference, including Adjacent-Channel Interference (ACI). TV White Space (TVWS) spectrum access is often investigated without taking into account ACI between the transmissions of TV Bands Devices (TVBDs) and licensed TV stations. Guard Bands (GBs) can be used to protect data transmissions and mitigate the ACI problem. Therefore, in [9] we consider a spectrum database that is administrated by a database operator, and an opportunistic secondary system, in which every TVBD is equipped with a single antenna that can be tuned to a subset of licensed channels. This can be done, for example, through adaptive channel aggregation or bonding techniques.

We investigate the distributed spectrum management problem in opportunistic TVWS systems using a game theoretical approach that accounts for adjacent channel interference and spatial reuse. TVBDs compete to access idle TV channels and select channel "blocks" that optimize an objective function. This function provides a tradeoff between the achieved rate and a cost factor that depends on the interference between TVBDs. We consider practical cases where contiguous or non-contiguous channels can be accessed by TVBDs, imposing realistic constraints on the maximum frequency span between the aggregated/bonded channels. We show that under general conditions, the proposed TVWS management games admit a potential function. Accordingly, a "best response" strategy allows us to determine the spectrum assignment of all players. This algorithm is shown to converge in a few iterations to a Nash Equilibrium (NE). Furthermore, we propose an effective algorithm based on Imitation dynamics, where a TVBD probabilistically imitates successful selection strategies of other TVBDs in order to improve its objective function. Numerical results show that our game theoretical framework provides a very effective tradeoff (close to optimal, centralized spectrum allocations) between efficient TV spectrum use and reduction of interference between TVBDs.

7.9. Straight: Stochastic Geometry and User History Based Mobility Estimation

5G is envisioned to support scalable networks and improved user experience with virtually zero latency and ultra broad-band service. Supporting unlimited seamless mobility is one of the key issues and also for network resource utilization efficiency. In [16], we focus on mobility management and user equipment (UE) speed class estimation, also known as mobility state estimation (MSE). We propose a method for estimating the UE mobility which is compliant with UE history information specifications by 3GPP (3rd Generation Partnership Project). We also exploit the impact of the environment on the UE trajectory and speed when determining UE mobility state. We evaluate the effectiveness of our algorithm using realistic mobility traces and network topology of the city of Cologne in Germany provided by the Kolntrace project. Results show that the speed classification of UEs can be achieved with much higher accuracy compared to existing legacy 3GPP LTE MSE procedures.

7.10. Mobility State Estimation in LTE

Estimating mobile user speed is a problematic issue which has significant impacts to radio resource management and also to the mobility management of Long Term Evolution (LTE) networks. In [15] introduces two algorithms that can estimate the speed of mobile user equipments (UE), with low computational requirement, and without modification of neither current user equipment nor 3GPP standard protocol. The proposed methods rely on uplink (UL) sounding reference signal (SRS) power measurements performed at the eNodeB (eNB) and remain efficient with large sampling period (e.g., 40 ms or beyond). We evaluate the effectiveness of our algorithms using realistic LTE system data provided by the eNB Layer1 team of Alcatel-Lucent. Results show that the classification of UE's speed required by LTE can be achieved with high accuracy. In addition, they have minimal impact to the central processing unit (CPU) and the memory of eNB modem. We see that they are very practical to today's LTE networks and would allow a continuous and real-time UE speed estimation

7.11. Cell Planning for Mobility Management in Heterogeneous Cellular Networks

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, in [25] 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 (cf [43]), 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 (see [36]).

7.12. Location Aware Opportunistic Bandwidth Sharing between Static and Mobile Users with Stochastic Learning in Cellular Networks

In [26] 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.

7.13. Gibbsian On-Line Distributed Content Caching Strategy for Cellular Networks

In [27] we develop Gibbs sampling based techniques for learning the optimal content placement in a cellular network. A collection of base stations are scattered on the space, each having a cell (possibly overlapping with other cells). Mobile users request for downloads from a finite set of contents according to some popularity distribution. Each base station can store only a strict subset of the contents at a time; if a requested content is not available at any serving base station, it has to be downloaded from the backhaul. Thus, there arises the problem of optimal content placement which can minimize the download rate from the backhaul, or equivalently maximize the cache hit rate. Using similar ideas as Gibbs sampling, we propose simple sequential content update rules that decide whether to store a content at a base station based on the knowledge of contents in neighbouring base stations. The update rule is shown to be asymptotically converging to the optimal content placement for all nodes. Next, we extend the algorithm to address the situation where content popularities and cell topology are initially unknown, but are estimated as new requests arrive to the base stations. Finally, improvement in cache hit rate is demonstrated numerically.

7.14. Spatial Disparity of QoS Metrics Between Base Stations in Wireless Cellular Networks

This work contributes to the line of research on the development of analytic tools for the QoS evaluation and dimensioning of operator cellular networks which is the subject of long-term collaboration between TREC/DYOGENE and Orange Labs (cf Section 8.1.1). Our focus in [8] is to explicitly characterize the disparity of quality of service (QoS) metrics between base stations in large heterogeneous wireless cellular networks. The considered QoS metrics are cell load, users' number, and user throughput. The spatial disparity of these metrics is due to the irregularity of the cells' geometry. In order to consider these irregularities, we assume a Poisson point process of base station locations, random transmission powers, and log-normal shadowing. The interdependency between the performances of the base stations is characterized by a system of load equations. The typical cell simulation model consists in resolving this system in order to find the loads and then deduce the remaining characteristics for each cell of the network. Using stochastic geometric and queueing theoretic techniques, we define the QoS averages, variances, and distributions. Inspired by the analysis of the typical cell model, several investigations lead us to propose a fully analytic approach, called mean cell model, that approximates the averages, variances, and distributions of these QoS metrics. Numerical experiments show a good agreement between the proposed approximations, simulation results, and real-life network measurements.

7.15. Stronger Wireless Signals Appear More Poisson

This work contributes to the line of research on Poisson convergence in wireless networks with strong shadowing initiated in [37], [35]. More recently, Keeler, Ross and Xia derived in [51] approximation and convergence results, which imply that the point process formed from the signal strengths received by an observer in a wireless network under a general statistical propagation model can be modeled by an inhomogeneous Poisson point process on the positive real line. The basic requirement for the results to apply is that there must be a large number of transmitters with a small proportion having a strong signal. The aim of [12] is to apply some of the main results of [51] in a less general but more easily applicable form, to illustrate how the results can apply to functions of the point process of signal strengths, and to gain intuition on when the Poisson model for transmitter locations is appropriate. A new and useful observation is that it is the stronger signals that behave more Poisson, which supports recent experimental work.

7.16. On Some Diffusion and Spanning Problems in Configuration Model

A number of real-world systems consisting of interacting agents can be usefully modelled by graphs, where the agents are represented by the vertices of the graph and the interactions by the edges. Such systems can be as diverse and complex as social networks (traditional or online), protein-protein interaction networks, internet, transport network and inter-bank loan networks. One important question that arises in the study of these networks is: to what extent, the local statistics of a network determine its global topology. This problem can be approached by constructing a random graph constrained to have some of the same local statistics as those observed in the graph of interest. One such random graph model is configuration model, which is constructed in such a way that a uniformly chosen vertex has a given degree distribution. This is the random graph which provides the underlying framework for the problems considered in the PhD thesis [3]. As our first problem, we consider propagation of influence on configuration model, where each vertex can be influenced by any of its neighbours but in its turn, it can only influence a random subset of its neighbours. Our (enhanced) model is described by the total degree of the typical vertex and the number of neighbours it is able to influence. We give a tight condition, involving the joint distribution of these two degrees, which allows with high probability the influence to reach an essentially unique non-negligible set of the vertices, called a big influenced component, provided that the source vertex is chosen from a set of good pioneers. We explicitly evaluate the asymptotic relative size of the influenced component as well as of the set of good pioneers, provided it is non-negligible. Our proof uses the joint exploration of the configuration model and the propagation of the influence up to the time when a big influenced component is completed, a technique introduced in Janson and Luczak [48]. Our model can be seen as a generalization of the classical Bond and Node percolation on configuration model, with the difference stemming from the oriented conductivity of edges in our model. We illustrate these results using a few examples which are interesting from either theoretical or real-world perspective. The examples are, in particular, motivated by the viral marketing phenomenon in the context of social networks. Next, we consider the isolated vertices and the longest edge of the minimum spanning tree of a weighted configuration model. Using Stein-Chen method, we compute the asymptotic distribution of the number of vertices which are separated from the rest of the graph by some critical distance, say alpha. This distribution gives the scaling of the length of the longest edge of the nearest neighbour graph with the size of the graph. We then use the results of Fountoulakis [45] on percolation to prove that after removing all the edges of length greater than alpha, the subgraph obtained is connected but for the isolated vertices. This leads us to conclude that the longest edge of the minimal spanning tree and that of the nearest neighbour graph coincide with high probability. Finally, we investigate a more general question, that is, whether some ordering based on local statistics of the graph would lead to an ordering of the global topological properties, so that the bounds for more complex graphs could be obtained from their simplified versions. To this end, we introduce a convex order on random graphs and discuss some implications, particularly how it can lead to the ordering of percolation probabilities in certain situations.

7.17. Inferring Sparsity: Compressed Sensing Using Generalized Restricted Boltzmann Machines

In [23] we consider compressed sensing reconstruction from M measurements of K-sparse structured signals which do not possess a writable correlation model. Assuming that a generative statistical model, such as a Boltzmann machine, can be trained in an unsupervised manner on example signals, we demonstrate how this signal model can be used within a Bayesian framework of signal reconstruction. By deriving a message-passing inference for general distribution restricted Boltzmann machines, we are able to integrate these inferred signal models into approximate message passing for compressed sensing reconstruction. Finally, we show for the MNIST dataset that this approach can be very effective, even for $M < K$.

7.18. Recovering Asymmetric Communities in the Stochastic Block Model

In [22], we consider the sparse stochastic block model in the case where the degrees are uninformative. The case where the two communities have approximately the same size has been extensively studied and we concentrate here on the community detection problem in the case of unbalanced communities. In this setting, spectral algorithms based on the non-backtracking matrix are known to solve the community detection problem (i.e. do strictly better than a random guess) when the signal is sufficiently large namely above the so-called Kesten Stigum threshold. In this regime and when the average degree tends to infinity, we show that if the community of a vanishing fraction of the vertices is revealed, then a local algorithm (belief propagation) is optimal down to Kesten Stigum threshold and we quantify explicitly its performance. Below the Kesten Stigum threshold, we show that, in the large degree limit, there is a second threshold called the spinodal curve below which, the community detection problem is not solvable. The spinodal curve is equal to the Kesten Stigum threshold when the fraction of vertices in the smallest community is above $p^* = \frac{1}{2} - \frac{1}{2\sqrt{3}}$, so that the Kesten Stigum threshold is the threshold for solvability of the community detection in this case. However when the smallest community is smaller than p^*, the spinodal curve only provides a lower bound on the threshold for solvability. In the regime below the Kesten Stigum bound and above the spinodal curve, we also characterize the performance of best local algorithms as a function of the fraction of revealed vertices. Our proof relies on a careful analysis of the associated reconstruction problem on trees which might be of independent interest. In particular, we show that the spinodal curve corresponds to the reconstruction threshold on the tree.

7.19. A Spectral Algorithm with Additive Clustering for the Recovery of Overlapping Communities in Networks

[17] presents a novel spectral algorithm with additive clustering, designed to identify overlapping communities in networks. The algorithm is based on geometric properties of the spectrum of the expected adjacency matrix in a random graph model that we call stochastic blockmodel withoverlap (SBMO). An adaptive version of the algorithm, that does not require the knowledge of the number of hidden communities, is proved to be

consistent under the SBMO when the degrees in the graph are (slightly more than) logarithmic. The algorithm is shown to perform well on simulated data and on real-world graphs with known overlapping communities.

7.20. Impact of Community Structure on Cascades

The threshold model is widely used to study the propagation of opinions and technologies in social networks. In this model individuals adopt the new behavior based on how many neighbors have already chosen it. In [20] we study cascades under the threshold model on sparse random graphs with community structure to see whether the existence of communities affects the number of individuals who finally adopt the new behavior. Specifically, we consider the permanent adoption model where nodes that have adopted the new behavior cannot change their state. When seeding a small number of agents with the new behavior, the community structure has little effect on the final proportion of people that adopt it, i.e., the contagion threshold is the same as if there were just one community. On the other hand, seeding a fraction of population with the new behavior has a significant impact on the cascade with the optimal seeding strategy depending on how strongly the communities are connected. In particular, when the communities are strongly connected, seeding in one community outperforms the symmetric seeding strategy that seeds equally in all communities.

7.21. Clustering from Sparse Pairwise Measurements

In [21] We consider the problem of grouping items into clusters based on few random pairwise comparisons between the items. We introduce three closely related algorithms for this task: a belief propagation algorithm approximating the Bayes optimal solution, and two spectral algorithms based on the non-backtracking and Bethe Hessian operators. For the case of two symmetric clusters, we conjecture that these algorithms are asymptotically optimal in that they detect the clusters as soon as it is information theoretically possible to do so. We substantiate this claim for one of the spectral approaches we introduce.

7.22. Limit Theory for Geometric Statistics of Clustering Point Processes

Let P be a simple, stationary, clustering point process on the d-dimensional Euclidean space, in the sense that its correlation functions factorize up to an additive error decaying exponentially fast with the separation distance. Let P_n be its restriction to a hypercube windows of volume n. We consider statistics of P_n admitting the representation as sums of spatially dependent terms $H_n = \sum_{x \in P_n} \xi(x, P_n)$, where $\xi(x, P_n)$ is a real valued (score) function, representing the interaction of x with P_n. When the score function depends locally on P_n in the sense that its radius of stabilization has an exponential tail, we establish expectation asymptotics, variance asymptotics, and central limit theorems for H_n as the volume n of the window goes to infinity.

This gives the limit theory for 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) of determinantal point processes with fast decreasing kernels, including the α-Ginibre ensembles. It also gives the limit theory for geometric U-statistics of permanental point processes as well as the zero set of Gaussian entire functions. This extends the existing literature treating the limit theory of sums of stabilizing scores of Poisson and binomial input. In the setting of clustering point processes, it also extends the results of Soshnikov [61] as well as work of Nazarov and Sodin [55].

The proof of the central limit theorem relies on a factorial moment expansion originating in Blaszczyszyn [34] to show clustering of mixed moments of the score function. Clustering extends the cumulant method to the setting of purely atomic random measures, yielding the asymptotic normality of H_n.

7.23. The Boolean Model in the Shannon Regime: Three Thresholds and Related Asymptotics

In [4] we consider a family of Boolean models, indexed by integers $n \geq 1$, where the n-th model features a Poisson point process in $\mathbb{R}^n$ of intensity $e^{n\rho_n}$ with $\rho_n \to \rho$ as $n \to \infty$, and balls of independent and identically distributed radii distributed like $\overline{X}_n\sqrt{n}$, with $\overline{X}_n$ satisfying a large deviations principle. It is shown that there exist three deterministic thresholds: τ_d the degree threshold; τ_p the percolation threshold; and τ_v the volume fraction threshold; such that asymptotically as n tends to infinity, in a sense made precise in the paper: (i) for $\rho < \tau_d$, almost every point is isolated, namely its ball intersects no other ball; (ii) for $\tau_d < \rho < \tau_p$, almost every ball intersects an infinite number of balls and nevertheless there is no percolation; (iii) for $\tau_p < \rho < \tau_v$, the volume fraction is 0 and nevertheless percolation occurs; (iv) for $\tau_d < \rho < \tau_v$, almost every ball intersects an infinite number of balls and nevertheless the volume fraction is 0; (v) for $\rho > \tau_v$, the whole space covered. The analysis of this asymptotic regime is motivated by related problems in information theory, and may be of interest in other applications of stochastic geometry.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. CRE with Orange

One year CRE contract titled “Mise au point d’une méthode d’évaluation de la qualité de service pour le sens montant d’un réseau cellulaire LTE validée avec les mesures terrain” between Inria and Orange Labs have been signed in 2015 end realized in 2016. It is a part of the long-term collaboration between TREC/DYOGENE and Orange Labs, represented by M. K. Karray, for the development of analytic tools for the QoS evaluation and dimensioning of operator cellular networks. Arpan Chattopadhyay was hired by Inria as a post-doctoral fellow thanks to this contract.

8.1.2. Joint Research Lab with Nokia Bell Labs

Arpan Mukhopadhyay was hired by Inria as a post-doctoral fellow within this lab dedicated to the research on communication networks of the future; https://www.inria.fr/en/institute/partnerships/industrial-partnerships2/alcatel-lucent-bell-labs-france.

8.2. Bilateral Grants with Industry

8.2.1. CIFRE Nokia

PhD: Dalia-Georgiana Herculea, co-advised by B. Blaszczyszyn, E. Altman and Ph. Jacquet

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. V. Anantharam [UC Berkeley] was invited professor by LINCS in June-July 2016. He was a speaker at the LINCS Shannon Day organized by M. Lelarge and F. Baccelli in June 2016.

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 Pierre Calka [Université de Rouen]. 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 IM

Members of Dyogene participate in GdR-IM (Informatique-Mathématiques), https://www.gdr-im.fr/, working groups ALEA and SDA2 (Systèmes dynamiques, Automates et Algorithmique).

9.2.3. 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

9.2.4. PGMO

Gaspard Monge Program for Optimization and Operations Research project Decentralized control for renewable integration in smart-grids (2015-17). PI: A. Busic.

9.2.5. ANR MARMOTE

Markovian Modeling Tools and Environments - coordinator: Alain Jean-Marie (Inria Maestro); local coordinator (for partner Inria Paris-Rocquencourt): A. Bušić; Started: January 2013; Duration: 48 months; partners: Inria Paris-Rocquencourt (EPI DYOGENE), Inria Sophia Antipolis Méditerranée (EPI MAESTRO), Inria Grenoble Rhône-Alpes (EPI MESCAL), Université Versaillese-St Quentin, Telecom SudParis, Université Paris-Est Creteil, Université Pierre et Marie Curie.

The aim of the project is to realize a modeling environment dedicated to Markov models. One part will develop the Perfect Simulation techniques, which allow to sample from the stationary distribution of the process. A second one will develop parallelization techniques for Monte Carlo simulation. A third one will develop numerical computation techniques for a wide class of Markov models. All these developments will be integrated into a programming environment allowing the specification of models and their solution strategy. Several applications will be studied in various scientific disciplines: physics, biology, economics, network engineering.

9.3. International Initiatives

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

9.3.1.1. PARIS

Title: Probabilistic Algorithms for Renewable Integration in Smart Grid

International Partner (Institution - Laboratory - Researcher):

University of Florida (United States) — Sean Meyn

Start year: 2015

See also: http://www.di.ens.fr/~busic/PARIS/

The importance of statistical modeling and probabilistic control techniques in the power systems area is now evident to practitioners in both the U.S. and Europe. Renewable generation has brought unforeseen volatility to the grid that require new techniques in distributed and probabilistic control. In a series of recent papers the two PIs have brought together their complementary skills in optimization, Markov modeling, simulation, and stochastic networks that may help to solve some pressing open problems in this area. This new research also opens many exciting new scientific questions.

9.3.2. Inria International Partners

9.3.2.1. Informal International Partners

- B. Blaszczyszyn is collaborationg with T. Rolski, R. Szekli, (University of Wroclaw), D. Yogeshwaran (Indian Statistical Institute) and Y. Yukich (Lehigh University)

- A. Busic is participating to the ARPA-E Powernet project led by Ram Rajagopal (Stanford); https://web.stanford.edu/~ramr/powernet.htm

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Sean Meyn [Professor, University of Florida, Jun 2016]
- Adithya Munegowda Devraj [PhD student, University of Florida, May – Jul 2016]
- Sebastien Ziesche [PdD student, Karlsruhe Institute of Technology, March 2016]

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

- Bartlomiej Blaszczyszyn: 5th Stochastic Models Conference in Będlewo, Poland; http://www.math.uni.wroc.pl/~lorek/bedlewo2016/index.php
- Ana Busic: Workshop on the stochastic optimization and games with applications to energy and networks; http://gdrro.lip6.fr/?q=node/147

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

- Bartlomiej Blaszczyszyn: WiOpt/Spaswin 2016.
- Anne Bouillard: Valuetools 2016 conference.
- Ana Busic: QEST 2016, IEEE SmartGridComm 2016,
- Jocelyne Elias: GSNC 2016, Globecom 2016 SAC CN, IEEE WCNC 2016, WD 2016, IEEE ICCVE 2016.
- Marc Lelarge: AISTATS 2017, NIPS 2016, Workshop on Algorithms and Models for the Web Graph 2016, ACM SIGMETRICS 2017, 2016, ACM Mobihoc 2017, 2016, IEEE INFOCOM 2017, 2016, ICALP 2016.

10.1.2.2. Reviewer

- Dalia-Georgiana Herculea: IEEE VTC , IEEE ICT.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Marc Lelarge: IEEE Transactions on Network Science and Engineering, Bernoulli Journal, Queueing Systems.

10.1.3.2. Reviewer - Reviewing Activities

- Bartlomiej Blaszczyszyn: Ann. Appl. Probab., IEEE TNSE, TWC, WCL.
- Dalia-Georgiana Herculea: IEEE Access, Wiley Transactions on Emerging Telecommunications Technologies.

10.1.4. Invited Talks

- Bartlomiej Blaszczyszyn: Workshop on Continuum Percolation in Lille http://math.univ-lille1.fr/~heinrich/Contperc2016/ , 1st Symposium on Spatial Networks in Oxford http://www.eng.ox.ac.uk/sen/events.html, Workshop on Probabilistic Methods in Telecommunication WIAS Berlin https://www.wias-berlin.de/workshops/PMT16/.
- Anne Bouillard: ASMTA conference.
- Ana Busic: Institute for Mathematics and its Applications, Univ. of Minnesota, Indo-UK workshop on Energy Management, ICMS Edinburgh,Workshop EDF Lab', Simons Institute Berkeley, Faculty of Electrical Engineering and Computing, University of Zagreb,
- Dalia-Georgiana Herculea: Institute of Stochastics, University of Ulm, LINCS Internal Workshop.

10.1.5. Leadership within the Scientific Community

- Ana Busic is co-responsable of the research group COSMOS (Stochastic optimization and control, modeling and simulation) of the GDR-RO; http://gdrro.lip6.fr/?q=node/78.
- Dalia-Georgiana Herculea: Representation of Phd students in the Board of Laboratory of Information, Networking and Communication Sciences (LINCS).

10.1.6. Scientific Expertise

- Bartlomiej Blaszczyszyn: ANR France, ISF Israel, NSC Poland, ERC Europe.

10.1.7. Research Administration

- A. Busic: Co-president of CES (Commission des Emplois Scientifiques) of Inria Paris, Member of the hiring committee for CR2 research positions at Inria Paris, Member of CDT (Commission de développement technologique) of Inria Paris.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence: Anne Bouillard (Cours) et Rémi Varloot (TD) **Structures et algorithmes aléatoires** 80heqTD, L3, ENS, France.

Licence: Anne Bouillard (Cours) **Théorie de l'information et du codage** 24 heqTD, L3, ENS, France.

Licence: Anne Bouillard (Cours) **Algorithmique et programmation** 21 heqTD, L3, ENS, France.

Master: Bartlomiej Blaszczyszyn (Cours) **Processus ponctuels, graphes aléatoires et géeométrie stochastique** 39heqTD, M2 Probabilités et Modèles Aléatoires, UPMC, France

Master: Anne Bouillard (Cours + TD) **Fondements de la modélisation des réseaux** 18heqTD, M2, MPRI, France

Master: Ana Busic and Marc Lelarge (Cours) et Rémi Varloot (TD) **Modéles et algorithmes de réseaux**, 50 heqTD, M1, ENS, Paris, France.

Master: Ana Busic (Cours) **Simulation**, 13.5 heqTD, M2 AMIS UVSQ, France.

Master: Marc Lelarge (TD) **Networks: distributed control and emerging phenomena** (cours given by Laurent Massoulié) M2, Ecole Polytechnique, France.

Master: Marc Lelarge (TD) **Aléatoire** (cours given by Sylvie Mél'eard) M2, Ecole Polytechnique, France.

10.2.2. Supervision

PhD: Kumar Gaurav, On some diffusion and spanning problems in configuration model [3], defended in November 2016, supervised by B. Blaszczyszyn

PhD in progress: Léeo Miolane, supervised by Marc Lelarge

PhD in progress: Dalia-Georgiana Herculea, co-advised by B. Blaszczyszyn, E. Altman and Ph. Jacquet

PhD in progress: Lennart Gulikers, supervised by Marc Lelarge with Laurent Massoulié

PhD in progress: Md Umar Hashmi, Decentralized control for renewable integration in smartgrids, from December 2015, co-advised by A. Busic and M. Lelarge

PhD in progress: Alexandre Hollocou: supervised by Marc Lelarge with Thomas Bonald

PhD in progress: Christelle Rovetta, Applications of perfect sampling to queuing networks and random generation of combinatorial objects, from December 2013, co-advised by Anne Bouillard and Ana Busic

PhD in progress: Sébastien Samain, Monte Carlo methods for performance evaluation and reinforcement learning, from November 2016, supervised by A. Busic

PhD in progress: Rémi Varloot, supervised by Marc Lelarge with Laurent Massoulié

PostDoc: Arpan Mukhopadhyay, supervised by Marc Lelarge with Nidhi Hegde

PostDoc: Virag Shah: supervised by Marc Lelarge with Laurent Massoulié and Milan Vojnovic

10.2.3. Juries

- Anne Bouillard: reviewer of the PhD thesis of Mickael Back (University of Kaiserslautern).
- Ana Busic: PhD jury of Alexandra Ugolnikova, LIPN (Université Paris Nord), Rim Kaddah (Télécom ParisTech).
- Bartlomiej Blaszczyszyn: reviewer of the PhD thesis of Jihong YU LRI (Université Paris-Sud).
- Marc Lelarge: reviewer for the PhD thesis of Kevin Scaman (ENS Cachan) and Van Hao Can (Université Aix-Marseille).

10.3. Popularization

- Anne Bouillard: invited speaker at “Girls can code” 2016 session (Training session for female junior and senior highschool students).

11. Bibliography

Major publications by the team in recent years

[1] F. Baccelli, B. Błaszczyszyn.*Stochastic Geometry and Wireless Networks, Volume I — Theory*, Foundations and Trends in Networking, NoW Publishers, 2009, vol. 3, No 3–4, p. 249–449.

[2] F. Baccelli, B. Błaszczyszyn.*Stochastic Geometry and Wireless Networks, Volume II — Applications*, Foundations and Trends in Networking, Now Publishers, 2009, vol. 4, No 1–2, p. 1–312.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[3] K. Gaurav.*On some diffusion and spanning problems in configuration model*, UPMC - Université Paris 6 Pierre et Marie Curie, November 2016, https://hal.inria.fr/tel-01400999.

Articles in International Peer-Reviewed Journal

[4] V. Anantharam, F. Baccelli.*The Boolean Model in the Shannon Regime: Three Thresholds and Related Asymptotics*, in "Journal of Applied Probability", 2016, vol. 53, n^{o} 4, p. 1001 - 1018, https://hal.inria.fr/hal-01259177.

[5] F. BACCELLI, A. N. RYBKO, S. B. SHLOSMAN.*Queueing networks with mobile servers: The mean-field approach*, in "Problems of Information Transmission", October 2016, vol. 52, n^{o} 2, p. 178 - 199 [*DOI :* 10.1134/S0032946016020071], https://hal.inria.fr/hal-01394044.

[6] A. BOUILLARD, A. BUSIC, C. ROVETTA.*Low complexity state space representation and algorithms for closed queueing networks exact sampling*, in "Performance Evaluation", September 2016, vol. 103, p. 2-22 [*DOI :* 10.1016/J.PEVA.2016.06.006], https://hal.inria.fr/hal-01396074.

[7] A. BOUILLARD, E. FAOU, M. ZAVIDOVIQUE.*Fast Weak-Kam Integrators for separable Hamiltonian systems*, in "Mathematics of Computation", 2016, vol. 85, n^{o} 297, p. 85-117, https://hal.archives-ouvertes.fr/hal-00743462.

[8] B. BŁASZCZYSZYN, R. IBRAHIM, M. K. KARRAY.*Spatial disparity of QoS metrics between base stations in wireless cellular networks*, in "IEEE Transactions on Communications", October 2016, vol. 64, n^{o} 10, p. 4381 - 4393 [*DOI :* 10.1109/TCOMM.2016.2600668], https://hal.inria.fr/hal-01427698.

[9] J. ELIAS, F. MARTIGNON, L. CHEN, M. KRUNZ.*Distributed Spectrum Management in TV White Space Networks*, in "IEEE Transactions on Vehicular Technology", July 2016, https://hal.inria.fr/hal-01350583.

[10] J. ELIAS, F. MARTIGNON, S. PARIS, J. WANG.*Efficient Orchestration Mechanisms for Congestion Mitigation in NFV: Models and Algorithms*, in "IEEE Transactions on Services Computing", January 2016 [*DOI :* 10.1109/TSC.2015.2498176], https://hal.archives-ouvertes.fr/hal-01256392.

[11] M. FÜGGER, T. NOWAK, U. SCHMID.*Unfaithful Glitch Propagation in Existing Binary Circuit Models*, in "IEEE Transactions on Computers", March 2016, vol. 65, n^{o} 3, p. 964-978 [*DOI :* 10.1109/TC.2015.2435791], https://hal.archives-ouvertes.fr/hal-01231501.

[12] H. P. KEELER, N. ROSS, A. XIA, B. BŁASZCZYSZYN.*Stronger wireless signals appear more Poisson*, in "IEEE wireless communications letters", December 2016, vol. 5, n^{o} 6, p. 572 - 575, 9 pages with 1.5 line spacing [*DOI :* 10.1109/LWC.2016.2601913], https://hal.inria.fr/hal-01331897.

[13] M. MANGILI, J. ELIAS, F. MARTIGNON, A. CAPONE.*Optimal Planning of Virtual Content Delivery Networks under Uncertain Traffic Demands*, in "Computer Networks", June 2016, https://hal.inria.fr/hal-01338680.

International Conferences with Proceedings

[14] A. BUSIC, S. MEYN.*Distributed Randomized Control for Demand Dispatch*, in "55th IEEE Conference on Decision and Control (CDC)", Las Vegas, United States, Proceedings of 55th IEEE Conference on Decision and Control, December 2016, https://hal.archives-ouvertes.fr/hal-01423479.

[15] M. HADDAD, D.-G. HERCULEA, E. ALTMAN, N. BEN RACHED, V. CAPDEVIELLE, C. SHUE CHEN, F. RATOVELOMANANA.*Mobility State Estimation in LTE*, in "IEEE Wireless Communications and Networking Conference", Doha, Qatar, 10.1109/WCNC.2016.7564917, IEEE, April 2016 [*DOI :* 10.1109/WCNC.2016.7564917], https://hal.inria.fr/hal-01291728.

[16] D.-G. HERCULEA, C. SHUE CHEN, M. HADDAD, V. CAPDEVIELLE.*Straight: stochastic geometry and user history based mobility estimation*, in "HotPOST '16 Proceedings of the 8th ACM International Workshop on Hot Topics in Planet-scale mObile computing and online Social neTworking", Paderborn, Germany,

Proceedings of the 8th ACM International Workshop on Hot Topics in Planet-scale mObile Computing and Online Social neTworking, July 2016, 6 [*DOI :* 10.1145/2944789.2944790], https://hal.inria.fr/hal-01414185.

[17] E. KAUFMANN, T. BONALD, M. LELARGE.*A Spectral Algorithm with Additive Clustering for the Recovery of Overlapping Communities in Networks*, in "ALT 2016 - Algorithmic Learning Theory", Bari, Italy, R. ORTNER, H. U. SIMON, S. ZILLES (editors), Lecture Notes in Computer Science, Springer, October 2016, vol. 9925, p. 355-370 [*DOI :* 10.1007/978-3-319-46379-7_24], https://hal.archives-ouvertes.fr/hal-01163147.

[18] J. MATHIAS, A. BUSIC, S. MEYN.*Demand Dispatch with Heterogeneous Intelligent Loads*, in "50th Annual Hawaii International Conference on System Sciences (HICSS)", Waikoloa, HI, United States, Proc. of 55th Hawaii International Conference on System Sciences (HICSS), January 2017, Extended version of paper to appear in Proc. 50th Annual Hawaii International Conference on System Sciences (HICSS), 2017, https://hal.archives-ouvertes.fr/hal-01423485.

[19] J. MATHIAS, R. KADDAH, A. BUSIC, S. MEYN.*Smart Fridge / Dumb Grid? Demand Dispatch for the Power Grid of 2020*, in "49th Hawaii International Conference on System Sciences (HICSS)", Koloa, HI, United States, Proc. of 49th Hawaii International Conference on System Sciences (HICSS), January 2016, p. 2498-2507 [*DOI :* 10.1109/HICSS.2016.312], https://hal.archives-ouvertes.fr/hal-01423483.

[20] M. MOHARRAMI, V. SUBRAMANIAN, M. LIU, M. LELARGE.*Impact of Community Structure on Cascades*, in "ACM EC 2016", Maastricht, Netherlands, EC '16 Proceedings of the 2016 ACM Conference on Economics and Computation, July 2016, p. 635 - 636 [*DOI :* 10.1145/2940716.2953924], https://hal.archives-ouvertes.fr/hal-01391590.

[21] A. SAADE, F. KRZAKALA, M. LELARGE, L. ZDEBOROVÁ.*Clustering from sparse pairwise measurements*, in "2016 IEEE International Symposium on Information Theory (ISIT 2016)", Barcelone, Spain, July 2016, p. 780 - 784 [*DOI :* 10.1109/ISIT.2016.7541405], https://hal.archives-ouvertes.fr/hal-01391585.

Conferences without Proceedings

[22] F. CALTAGIRONE, M. LELARGE, L. MIOLANE.*Recovering asymmetric communities in the stochastic block model*, in "allerton 2016 54th Annual Allerton Conference on Communication, Control, and Computing", Monticello, United States, September 2016, https://hal.archives-ouvertes.fr/hal-01391609.

[23] E. W. TRAMEL, A. MANOEL, F. CALTAGIRONE, M. GABRIÉ, F. KRZAKALA.*Inferring sparsity: Compressed sensing using generalized restricted Boltzmann machines*, in "Information Theory Workshop (ITW), 2016 IEEE", Cambridge, United Kingdom, September 2016, p. 265 - 269 [*DOI :* 10.1109/ITW.2016.7606837], https://hal.inria.fr/hal-01416262.

Other Publications

[24] B. BŁASZCZYSZYN, D. YOGESHWARAN, J. E. YUKICH.*Limit theory for geometric statistics of clustering point processes*, June 2016, 69 pages, https://hal.inria.fr/hal-01331939.

[25] A. CHATTOPADHYAY, B. BŁASZCZYSZYN, E. ALTMAN.*Cell planning for mobility management in heterogeneous cellular networks*, June 2016, 13 pages, 5 diagrams, 2 plots, 3 images, https://hal.inria.fr/hal-01331936.

[26] A. CHATTOPADHYAY, B. BŁASZCZYSZYN, E. ALTMAN.*Location Aware Opportunistic Bandwidth Sharing between Static and Mobile Users with Stochastic Learning in Cellular Networks*, 2016, 16 Pages, 1 Figure, 1 Table, https://hal.inria.fr/hal-01401007.

[27] A. CHATTOPADHYAY, B. BŁASZCZYSZYN.*Gibbsian On-Line Distributed Content Caching Strategy for Cellular Networks*, 2016, 9 pages, 2 figures, https://hal.inria.fr/hal-01401010.

[28] L. GULIKERS, M. LELARGE, L. MASSOULIÉ.*A spectral method for community detection in moderately-sparse degree-corrected stochastic block models*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01258191.

[29] L. GULIKERS, M. LELARGE, L. MASSOULIÉ.*An Impossibility Result for Reconstruction in a Degree-Corrected Planted-Partition Model*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01258194.

References in notes

[30] F. BACCELLI, B. BŁASZCZYSZYN, F. TOURNOIS.*Spatial averages of coverage characteristics in large CDMA networks*, in "Wirel. Netw.", 2002, vol. 8, p. 569–586, http://dx.doi.org/10.1023/A:1020321501945.

[31] P. BAROOAH, A. BUSIC, S. MEYN.*Spectral decomposition of demand-side flexibility for reliable ancillary services in a smart grid*, in "System Sciences (HICSS), 2015 48th Hawaii International Conference on", IEEE, 2015, p. 2700–2709.

[32] C. BORDENAVE, M. LELARGE, L. MASSOULIÉ.*Non-backtracking spectrum of random graphs: community detection and non-regular Ramanujan graphs*, in "2015 IEEE 56th Annual Symposium on Foundations of Computer Science", Berkeley, United States, 2015 IEEE 56th Annual Symposium on Foundations of Computer Science, October 2015 [*DOI :* 10.1109/FOCS.2015.86], https://hal.archives-ouvertes.fr/hal-01226796.

[33] A. BUSIC, J. MAIRESSE, I. MARCOVICI.*Probabilistic cellular automata, invariant measures, and perfect sampling*, in "Proc. of 28th International Symposium on Theoretical Aspects of Computer Science, (STACS)", 2011, p. 296-307.

[34] B. BŁASZCZYSZYN.*Factorial-moment expansion for stochastic systems*, in "Stoch. Proc. Appl.", 1995, vol. 56, p. 321–335.

[35] B. BŁASZCZYSZYN, M. K. KARRAY, H. P. KEELER.*Wireless networks appear Poissonian due to strong shadowing*, in "IEEE Trans. Wireless Commun.", 2015, vol. 14, n^{o} 8, p. 4379–4390, Publised online 7 April 2015.

[36] B. BŁASZCZYSZYN, H. P. KEELER.*Equivalence and comparison of heterogeneous cellular networks*, in "Personal, Indoor and Mobile Radio Communications (PIMRC Workshops), 2013 IEEE 24th International Symposium on", IEEE, 2013, p. 153–157.

[37] B. BŁASZCZYSZYN, K. K. MOHAMED, H. P. KEELER.*Using Poisson processes to model lattice cellular networks*, in "Proc. of IEEE INFOCOM", 2013.

[38] E. J. CANDÈS, T. TAO.*Decoding by linear programming*, in "IEEE Trans. Inform. Theory", 2005, vol. 51, n^o 12, p. 4203–4215.

[39] Y. CHEN, A. BUSIC, S. MEYN.*State Estimation for the Individual and the Population in Mean Field Control with Application to Demand Dispatch*, in "54th IEEE Conference on Decision and Control", 2015.

[40] S. N. CHIU, D. STOYAN, W. S. KENDALL, J. MECKE.*Stochastic geometry and its applications*, John Wiley & Sons, 2013.

[41] R. L. CRUZ.*A calculus for network delay, Part I: Network elements in isolation*, in "IEEE Transactions on Information Theory", 1991, vol. 37, n^o 1, p. 114-131.

[42] R. L. CRUZ.*A calculus for network delay, Part II: Network analysis*, in "IEEE Transactions on Information Theory", 1991, vol. 37, n^o 1, p. 132-141.

[43] H. S. DHILLON, R. K. GANTI, F. BACCELLI, J. G. ANDREWS.*Modeling and analysis of K-tier downlink heterogeneous cellular networks*, in "IEEE Journal on Selected Areas in Communications", 2012, vol. 30, n^o 3, p. 550–560.

[44] D. L. DONOHO, J. TANNER.*Observed Universality of Phase Transitions in High-Dimensional Geometry, with Implications for Modern Data Analysis and Signal Processing*, in "Phil. Trans. R. Soc. A", 2011, p. 4273-4293.

[45] N. FOUNTOULAKIS.*Percolation on Sparse Random Graphs with Given Degree Sequence*, in "Internet Mathematics", January 2007, vol. 4, n^o 4, p. 329–356, http://dx.doi.org/10.1080/15427951.2007.10129148.

[46] B. GAUJAL, F. PERRONNIN, R. BERTIN.*Perfect simulation of a class of stochastic hybrid systems with an application to peer to peer systems*, in "Journal of Discrete Event Dynamic Systems", 2007, Special Issue on Hybrid Systems.

[47] O. HÄGGSTRÖM, K. NELANDER.*Exact sampling from anti-monotone systems*, in "Statist. Neerlandica", 1998, vol. 52, n^o 3, p. 360–380, http://dx.doi.org/10.1111/1467-9574.00090.

[48] S. JANSON, M. J. LUCZAK.*A new approach to the giant component problem*, in "Random Structures and Algorithms", March 2009, vol. 34, n^o 2, p. 197–216, http://doi.wiley.com/10.1002/rsa.20231.

[49] G. KALAI, S. SAFRA.*Threshold phenomena and influence: perspectives from mathematics, computer science, and economics*, in "Computational complexity and statistical physics", New York, St. Fe Inst. Stud. Sci. Complex., Oxford Univ. Press, 2006, p. 25–60.

[50] E. KAUFMANN, T. BONALD, M. LELARGE.*An Adaptive Spectral Algorithm for the Recovery of Overlapping Communities in Networks*, June 2015, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01163147.

[51] H. P. KEELER, N. ROSS, A. XIA.*When do wireless network signals appear Poisson?*, in "arXiv preprint arXiv:1411.3757", 2014, to appear in Bernoulli.

[52] W. S. KENDALL.*Perfect simulation for the area-interaction point process*, in "Probability Towards 2000", New York, L. ACCARDI, C. C. HEYDE (editors), Springer-Verlag, 1998, p. 218–234.

[53] J.-Y. LE BOUDEC, P. THIRAN.*Network Calculus: A Theory of Deterministic Queuing Systems for the Internet*, revised version 4, May 10, 2004, Springer-Verlag, 2001, vol. LNCS 2050.

[54] S. MEYN, P. BAROOAH, A. BUSIC, Y. CHEN, J. EHREN.*Ancillary service to the grid using intelligent deferrable loads*, in "IEEE Transactions on Automatic Control", 2015, vol. 60, n^{o} 11, p. 2847–2862.

[55] F. NAZAROV, M. SODIN.*Correlation functions for random complex zeroes: strong clustering and local universality*, in "Communications in Mathematical Physics", 2012, vol. 310, n^{o} 1, p. 75–98.

[56] J. G. PROPP, D. B. WILSON.*Exact sampling with coupled Markov chains and applications to statistical mechanics*, in "Random Structures and Algorithms", 1996, vol. 9, n^{o} 1-2, p. 223-252, http://dbwilson.com/eus/.

[57] T. RICHARDSON, R. URBANKE.*Modern coding theory*, Cambridge University Press, Cambridge, 2008, xvi+572.

[58] W. RUI, J. XU, S. RAYADURGAM, M. LELARGE, L. MASSOULIÉ, B. HAJEK.*Clustering and Inference From Pairwise Comparisons*, in "SIGMETRICS '15 Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems", Portland, United States, SIGMETRICS '15 Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, 2015, vol. 43, n^{o} 1, 2 [*DOI :* 10.1145/2796314.2745887], https://hal.archives-ouvertes.fr/hal-01226785.

[59] A. SAADE, F. KRZAKALA, M. LELARGE, L. ZDEBOROVÁ.*Spectral Detection in the Censored Block Model*, in "ISIT 2015", Hong Kong, China, June 2015, https://hal.archives-ouvertes.fr/hal-01137955.

[60] A. SLY.*Computational Transition at the Uniqueness Threshold*, in "FOCS", IEEE Computer Society, 2010, p. 287-296.

[61] A. SOSHNIKOV.*Gaussian limit for determinantal random point fields*, in "Annals of probability", 2002, p. 171–187.

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:

1. - Architectures, systems and networks
1.2.1. - Dynamic reconfiguration
1.2.3. - Routing
1.2.4. - QoS, performance evaluation
1.2.5. - Internet of things
1.2.6. - Sensor networks
1.2.8. - Network security
1.2.9. - Social Networks
1.4. - Ubiquitous Systems
1.6. - Green Computing
2.3. - Embedded and cyber-physical systems
3.4. - Machine learning and statistics
3.5. - Social networks
4.1. - Threat analysis
4.4. - Security of equipment and software
4.6. - Authentication
4.7. - Access control
6.1. - Mathematical Modeling
6.1.2. - Stochastic Modeling (SPDE, SDE)
7.3. - Optimization
7.10. - Network science
7.11. - Performance evaluation
7.14. - Game Theory
8.2. - Machine learning
8.6. - Decision support

Other Research Topics and Application Domains:

4.2. - Nuclear Energy
4.3. - Renewable energy production
5.1. - Factory of the future
5.9. - Industrial maintenance
6.3.2. - Network protocols
6.3.3. - Network Management
6.4. - Internet of things
7.2. - Smart travel
7.2.1. - Smart vehicles
7.2.2. - Smart road
8.1.2. - Sensor networks for smart buildings

1. Members

Research Scientists

Pascale Minet [Inria, Research Scientist, HDR]
Paul Muhlethaler [Inria, Team leader, Research Scientist, HDR]
Thomas Watteyne [Inria, Research Scientist]

Technical Staff

Tengfei Chang [Inria, Postdoctoral Research Engineer]
Ines Khoufi [Inria, Postdoctoral Research Engineer]
Remy Leone [Inria, Postdoctoral Research Engineer, started Feb 2016]
Erwan Livolant [Inria, until Nov 2016]
Malisa Vucinic [Inria, Postdoctoral Research Lead, started Sep 2016]

PhD Students

Nesrine Ben Hassine [Inria]
Younes Bouchaala [Vedecom]
Keoma Brun-Laguna [Inria]
Mohamed Elhadad Or Hadded [ATER Paris V]
Fatma Marzouk [Ensi]
Jonathan Muñoz [Gridbe, CIFRE]

Post-Doctoral Fellow

Ehsan Ebrahimi Khaleghi [Inria, started May 2016]

Visiting Scientists

Felipe Lalanne [Reseacher, Inria Chile, Chile, visit 19–26 Octobre 2016]
Mario Gerla [Professor, UCLA, USA, visit 31 August - 23 September 10-20 December 2016]
Travis Massey [PhD Student, UC Berkeley, USA, visit 22 July 2016]
Diego Dujovne [Professor, Universidad Diego Portales, Chile, visit 22-31 July 2016]
Carlos Oroza [PhD Student, UC Berkeley, USA, visit 23-29 July 2016]
David Burnett [PhD Student, UC Berkeley, USA, visit 13-15 June 2016]
Branko Kerkez [Professor, U. Michigan, USA, visit 17-22 June 2016]
Steven Glaser [Professor, UC Berkeley, USA, visit 21-25 March 2016]
Filip Barac [PhD Student, Mid Sweden University, Sweden, visit 8-19 February 2016]
Xavi Vilajosana [Professor, UOC/OpenMote, Spain, visit 2-4 February 2016]

Administrative Assistant

Christine Anocq [Inria]

Others

Nadjib Achir [Univ. Paris XIII, Associate Professor, HDR]
Selma Boumerdassi [CNAM, Associate Professor]
Samia Bouzefrane [CNAM, Associate Professor]
Philippe Jacquet [Bell Labs (Alcatel), Division Leader, HDR]
Anis Laouiti [France Telecom, Associate Professor, HDR]
Dana Marinca [Univ. Versailles,Associate Professor]
Nadjib Ait Saadi [Univ. Paris Est, Associate Professor]
Jiangnan Yang [Inria, Internship, Apr-Aug 2016]

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 focus 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 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. Generalities

EVA inherits its expertise in designing algorithms and protocols from HiPERCOM2 (e.g. OLSR). EVA also inherit know-how in modeling, simulation, experimentation and standardization. Through this know-how and experience, the results obtained are both far-reaching and useful.

3.2. Physical Layer

We plan to study how advanced physical layers can be used in low-power wireless networks. For instance, collaborative techniques such as multiple antennas (e.g. the Massive MIMO technology) can improve communication efficiency. The idea is to use a 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 will pursue 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 will focus particularly on scheduled medium access in the context of deterministic industrial networks; this will involve optimizing the joint time slot and channel assignment. Distributed approaches will be considered, and the EVA team will determine their limits in terms of reliability, latency and throughput. Furthermore, adaptivity to application or environment changes will be 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 will study 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 will study 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. Nevertheless, 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 802.15.4e 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 around 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 will investigate self-learning networks. Machine learning approaches, based on experts and forecasters, will be 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 the popularity, expressed by the number of solicitations per day, of 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. Security Trade-off in Constrained Wireless Networks

Ensuring security is a sine qua non condition for the widespread acceptance and adoption of the IoT, in particular in industrial and military applications. While the Public-Key Infrastructure (PKI) approach is ubiquitous on the traditional Internet, constraints in terms of embedded memory, communication bandwidth and computational power make translating PKI to constrained networks non-trivial.

Two related standardization working groups were created in 2013 to address this issue. DICE (DTLS In Constrained Environments) is defining a DTLS (Datagram Transport Layer Security) profile that is suitable for IoT applications, using the (Constrained Application Protocol) CoAP protocol. ACE is standardizing authentication and authorization mechanisms for constrained environments.

The issue is to find the best trade-off between a communication and computation overhead compatible with the limited capacity of sensor nodes and the level of protection required by the application.

4. Application Domains

4.1. Generalities

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.

4.2. Industrial Process Automation

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.

[0] Highway Addressable Remote Transducer, http://en.hartcomm.org/.

Through the involvement of team members in standardization activities, the protocols and techniques will be 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.

4.3. 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 will work on such applications in an associate team “REALMS” comprising members from EVA, the university of Berkeley and the university of Michigan.

4.4. 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.

The Internet of Things (IoT) has received continuous attention since 2013, and has been a marketing tool for industry giants such as IBM and Cisco, and the focal point of major events such the Consumer Electronics Show and the IETF. The danger of such exposure is that any under-performance may ultimately disappoint early adopters.

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.

With the general public becoming increasingly familiar with the term “Internet of Things”, its definition is broadening to include all devices which can be interacted with from a network, and which do not fall under the generic term of “computer”.

The EVA team is dedicated to understanding and contributing to the IoT. In particular, the team will maintain 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 will help establish which technology best fits which application.

4.5. 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.

4.6. Smart Cities

It has been estimated that by 2030, 60% of the world’s population will live in cities. On the one hand, smart cities aim at making everyday life more attractive and pleasant for citizens; on the other hand, they facilitate how those citizens can participate in the life of the city.

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.

4.7. 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.

5. Highlights of the Year

5.1. Highlights of the Year

Awards

- Prof. Steven Glaser (UC Berkeley) and **Thomas Watteyne** recipients of the France-Berkeley Fund award for the project "SHRIMP: Smart Harbor Implementation", August 2016.
- Keoma Brun-Laguna and **Thomas Watteyne**, together with Ana Laura Diedrichs, Javier Emilio Chaar, Diego Dujovne, Juan Carlos Taffernaberry, Gustavo Mercado. Runner up IEEE SECON 2016 Best Demo Award with "A Demo of the PEACH IoT-based Frost Event Prediction System for Precision Agriculture", London, UK, 28 June 2016.
- Remy Leone and **Thomas Watteyne**. Recipient Google IoT Technology Research Award on "6TiSCH and WiFi coexistence with OpenWSN", March 2016.
- Tengfei Chang and **Thomas Watteyne**, together with Pedro Henrique Gomes, Pradipta Gosh, Bhaskar Krishnamachari. EWSN dependability competition 4th place with project "Reliability through Time-Slotted Channel Hopping and Flooding-based Routing", 16 February 2016.

Meeting & Seminars

Organization of Workshops and Conferences

- ***PEMWN 2016*** international conference on Performance Evaluation and modeling in Wired and wireless Networks, co-chaired by Leila Saidane and **Pascale Minet** and Farouk Kamoun , held in Paris, France, November 2016. Pascale Minet was general co-chair with Leila Saidane from ENSI (Tunisia) of the PEMWN 2016 conference, the 5th IFIP international conference on Performance Evaluation and Modeling of Wired and Wireless Networks, technically co-sponsored by IFIP WG6.2 and IEEE ComSoc (see https://sites.google.com/site/pemwn2016/). This conference was held at CNAM in Paris, 22-24 November 2016. It was sponsored by Inria, CNAM and ENSI. The organization co-chairs were Samia Bouzefrane and Selma Boumerdassi. Three tutorials were given:

   - *Internet of Vehicles: From Intelligent Grid to Autonomous Cars and Vehicular Clouds* by Mario Gerla, Professor, University of California, Los Angeles.
   - *5G: Can we make it by 2020?* by Merouane Debbah, Mathematical and Algorithmic Sciences Lab, Huawei, France.
   - *Internet of Things, hyper-massive wireless networks, where are the theoretical limits?* by Philippe Jacquet, NOKIA, France.

  Sixteen papers have been selected by the technical program committee and presented during the three days of the PEMWN 2016 conference.

- ***InterIoT 2016*** The 2nd EAI International Conference on Interoperability in IoT was co-organized by Nathalie Mitton, Thomas Noël (general co-chairs) and Thomas Watteyne (TPC chair). It took place 26-27 October 2016 in Paris, France.

Tutorials

- Standards for the Industrial IoT: a Hands-on Tutorial with OpenWSN and OpenMote. Xavier Vilajosana, Pere Tuset-Peiro, Tengfei Chang, **Thomas Watteyne**. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Valencia, Spain, 4-8 September 2016.
- Introduction to the IETF 6TiSCH stack with OpenWSN & OpenMote. **Thomas Watteyne**, Xavier Vilajosana, Pere Tuset-Peiro, Tengfei Chang. International Conference on Telecommunications (ICT), Thessaloniki, Greece, 16-18 May 2016.

Standardization Activities

- ***Standardization*** *meeting co-chaired by Inria-EVA*
  6TiSCH working group meeting at IETF 97, 17 November 2016, Seoul, South Korea.
- ***Standardization*** *meeting co-chaired by Inria-EVA*
  6TiSCH working group meeting at IETF 96, 18 July 2016, Berlin, Germany.
- ***Interop event*** *organized by ETSI and Inria-EVA*
  ETSI 6TiSCH 3 plugtests, 15-16 July 2016, Berlin, Germany.
- ***Standardization*** *meeting co-chaired by Inria-EVA*
  6TiSCH working group meeting at IETF 95, 4 April 2016, Buenos Aires, Argentina.
- ***Standardization*** *meeting co-chaired by Inria-EVA*
  ETSI 6TiSCH 2 plugtests, 2-4 February 2016, Paris, France.

Real-World Deployments

The networking technology developed at Inria-EVA has reached the level of maturity for it to be used in real-world deployment. We have worked on 3 main sets of deployments in 2016:

- **Save the Peaches** (http://www.savethepeaches.com/), a 23-node network in Western Argentina which monitors temperature and humidity to be predict frost events in peach orchards.
- **SnowHow** (http://www.snowhow.io/), a set of 18 low-power wireless networks (945 sensors total) deployed throughout the Californian Sierra Nevada to monitor the snowpack.
- *(current work)* A Smart Building deployment in the Inria-Paris research center.

From a networking point of view, these deployments SolSystem (see Section 6.8) as a back-end solution. Sensor data and network statistics are available at our Inria-Paris servers (https://sol.paris.inria.fr/) seconds after they were measured in the field.

Distinguished Visitors

- ***Invited Professor Mario Gerla***, from UCLA, USA. He stayed in the EVA team during 2 1-week stays (31 August-23 September, 10-20 December) to work with the EVA team on shock-wave mitigation using vehicular ad hoc networks.
- ***Invited Professor Leila Saidane***, from ENSI, Tunisia. She stayed in the EVA team from 28 November to 2 December 2016 to prepare common publications and identify further research directions.
- ***Invited Professor Diego Dujovne***, from Universidad Diego Portales, Chile. He stayed in the EVA team for a 1-week visit (22-31 July 2016) to integrate sensors in the low-power wireless platforms, to be deployed in Argantina as part of the PEACH project.
- ***Invited Professor Steven Glaser***, from UC Berkeley, USA. He stayed in the EVA team for a 1-week visit (21-25 June 2016) to explore funding opportunities beyond the REALMS associate team.
- ***Invited Professor Branko Kerkez***, from U. Michigan, USA. He stayed in the EVA team for a 1-week visit (17-22 June 2016) to work on the Internet of Water (2 papers submitted). This visit was part of the REALMS associate team.

6. New Software and Platforms

6.1. OpenWSN (Software)

Participants: Tengfei Chang, Jonathan Muñoz, Malisa Vucinic, Thomas Watteyne.

OpenWSN (http://www.openwsn.org/) 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.

The OpenWSN ADT started in 2015, with Research Engineer Tengfei Chang who joined the EVA team.

Highlights for 2016:

- Development:
  - better (continuous) testing of the existing OpenWSN code.
  - Build a image for a RaspberryPi which contains the OpenVisualizer preinstalled.
  - Port OpenWSN to the First prototype of the "Single-Chip uMote" (SCuM), developed in Prof. Pister's lab at UC Berkeley.
  - Create a Virtual Machine image with all OpenWSN development tools preinstalled.
  - Implementation of 6TiSCH standards as they appear, including revisions.
  - Maintenance of the "Golden Image" used as a reference during interoperability testing
- Recognition:
  - OpenWSN remains ETSI's reference implementation for IETF 6TiSCH-related standards. It is therefore the base for the ETSI's Golden Device for 6TiSCH standards, including IEEE802.15.4e TSCH, 6LoWPAN and RPL.
- Events:
  - **Tutorial**
    Standards for the Industrial IoT: a Hands-on Tutorial with OpenWSN and OpenMote. Xavier Vilajosana, Pere Tuset-Peiro, Tengfei Chang, **Thomas Watteyne**. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Valencia, Spain, 4-8 September 2016.
  - **Tutorial**

Introduction to the IETF 6TiSCH stack with OpenWSN & OpenMote. **Thomas Watteyne**, Xavier Vilajosana, Pere Tuset-Peiro, Tengfei Chang. International Conference on Telecommunications (ICT), Thessaloniki, Greece, 16-18 May 2016.

– **Interop Event**
  ETSI 6TiSCH 3 plugtests, 15-16 Juny 2016, Berlin, Germany
– **Interop Event**
  ETSI 6TiSCH 2 plugtests, 2-4 Februay 2016, Paris, France

This software appears in https://bil.inria.fr/.

6.2. OPERA and OCARI (Software)

Participants: Erwan Livolant, Pascale Minet.

The OPERA software was developed by the Hipercom2 team in the OCARI project It includes EOLSR, an energy efficient routing protocol and OSERENA, a coloring algorithm optimized for dense wireless networks. It was registered by the APP. In 2013, OPERA has been made available for download as an open software from the InriaGForge site: https://gforge.inria.fr/scm/?group_id=4665

In 2014, OPERA has been ported on a more powerful platform based on the Atmel transceiver AT86RF233 and on a 32 bits microcontroler Cortex M3. More details and documentation about this software are available in the website made by the Eva team: http://opera.gforge.inria.fr/index.html

In 2016, Erwan Livolant developed extensions to allow the remote management of the OCARI network and the transmission of commands to sensors and actuators.

6.3. F-Interop (Software)

Participants: Remy Leone, Thomas Watteyne.

F-Interop is revolutionizing the way interoperability events are conducted. We are building a cloud-based system which allows implementors to meet online to test their implementations against one another, verify compliance in a automated way, and verify the performance of their implementations on large scale testbeds. This significantly cuts down time-to-market for standards-based solutions, and eventually leads to more standards-based products on the market. The F-Interop platform starts by targeting 6TiSCH, 6LoWPAN and CoAP standards, but our ambition is for F-Interop to become the standard way of doing interoperability, at least at the IETF IoT level.

This implementation is done as part of the H2020 F-Interop project. More information at http://www.f-interop.eu/.

This software appears in https://bil.inria.fr/.

6.4. 6TiSCH Simulator (Software)

Participants: Malisa Vucinic, Thomas Watteyne.

The 6TiSCH simulator allows one to conduct high-level simulator of an IETF 6TiSCH network and answer questions such as How long does it take the nodes to join the network? What is the average power consumption? What does the latency distribution look like?

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.

This implementation is done as part of the H2020 ARMOUR project and the standardization activities of the Inria-EVA team at the IETF. It is published under an open-source BSD license and maintained at https://bitbucket.org/6tisch/simulator/.

This software appears in https://bil.inria.fr/.

6.5. 6TiSCH Wireshark Dissector (Software)

Participants: Jonathan Muñoz, Thomas Watteyne.

The goal of this project is to maintain Wireshark dissectors for 6TiSCH (and 6TiSCH-related) standards up-to-date.

Implementation on the dissectors is done through an open-source repository, stable code is regularly contributed back to the main Wireshark code base.

This implementation is done as part of the collaboration with Gridbee and the standardization activities of the Inria-EVA team at the IETF. It is published under an open-source BSD license and maintained at https://github.com/openwsn-berkeley/dissectors.

This software appears in https://bil.inria.fr/.

6.6. Mercator (Software)

Participants: Keoma Brun-Laguna, Thomas Watteyne.

Mercator allows one to evaluate the connectivity in a low-power wireless network. It is a collection of tools, including the firmware to load on the devices, the scripts that automate the measurements of the connectivity and the tools to structure and display results.

The firmware is written as part of the OpenWSN project. Scripts and analysis tools are written in Python.

It is published under an open-source BSD license and maintained at https://github.com/openwsn-berkeley/mercator.

This software appears in https://bil.inria.fr/.

6.7. CONNEXION (Software)

Participants: Pascale Minet, Erwan Livolant.

In the CONNEXION project, the integration of the OCARI wireless sensor network in a Service-Oriented Architecture using the OPC-UA/ROSA middleware went on with Telecom ParisTech. More precisely, we developed the remote managment of the OCARI network as well as the possibility to generate commands to the sensors and actuators.

Erwan Livolant developed an OCARI frame dissector plugin for Wireshark (https://www.wireshark.org) available from the Git repository at OCARI website. This tool displays the contents of the packets sniffed for the MAC, the NWK and the Application layers, taking into account the specificities of OCARI.

6.8. SolSystem (Software)

Participants: Keoma Brun-Laguna, Thomas Watteyne.

SolSystem is a back-end solution for a low-power wireless mesh network based on SmartMesh IP. It defines how low-power wireless devices must format and transfer sensor data, and the tools to gather, store and display sensor data.

This system is used in several deployments, including http://savethepeaches.com/ and http://snowhow.io/.

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).

It is published under an open-source BSD license. Information and overview at http://www.solsystem.io/.

This software appears in https://bil.inria.fr/.

6.9. Argus (Software)

Participants: Remy Leone, Thomas Watteyne.

Share your low-power wireless sniffer through the cloud!

Imagine you are a team of low-power wireless enthusiasts developing the next generation of products in standards. One essential tool in your toolkit is a low-power wireless sniffer, which shows you the frames which fly through the air.

Rather than requiring each person in your team to have a sniffer, Argus allows you to put in a share a sniffer through the cloud.

There are three piece to 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.

It is published under an open-source BSD license, maintained at https://github.com/openwsn-berkeley/argus.

This software appears in https://bil.inria.fr/.

6.10. SAHARA (Software)

Participants: Ines Khoufi, Erwan Livolant, Pascale Minet.

Ines Khoufi developed modules for the simulation of the Time Slotted Channel Hopping (TSCH) on the ns3 simulation tool. These modules include multi-interface management and transmission management according to a given schedule.

Erwan Livolant developed a SAHARA frame dissector plugin for Wireshark (https://www.wireshark.org). This tool displays the contents of the packets sniffed for the MAC and the NWK layers, taking into account the specificities of the SAHARA project.

6.11. FIT IoT-LAB (Platform)

Participants: Thomas Watteyne, Tengfei Chang.

Note well: IoT-lab is *not* a project of Inria-EVA. It is a large project which runs from 2011 to 2021 and which involves the following other partners Inria (Lille, Sophia-Antipolis, Grenoble), INSA, UPMC, Institut Télécom Paris, Institut Télécom Evry, LSIIT Strasbourg. This section highlights Inria-EVA activity and contribution to the IoT-lab testbed in 2016.

The Inria-EVA team has been using the platform extensively throughout 2016. During the process, we have been interacting closely with the IoT-lab team of engineers.

IoT-lab-related activities include:

- **Running OpenWSN networks at scale**. The IoT-lab has been an incredibly powerful tool to verify the scalability of the protocols and implementations in OpenWSN (Section 6.1). *Lead: Tengfei Chang*.
- **Assessing connectivity with Mercator**. The Mercator project (Section 6.6) is targetted at measuring the connectivity of the IoT-lab platform in time, space and frequency. *Lead: Keoma Brun-Laguna*.
- **16-channel Sniffer with Argus**. We are currently working extending the Argus project (Section 6.9) with support for IoT-lab motes. That is, rather than using a 16-channel sniffer, turn 16 IoT-lab nodes into a distributed multi-frequency sniffer. *Lead: Remy Leone*.

- **IoT-lab for Interop Testing with F-Interop**. Through the H2020 F-Interop project, we are developing the tools (see Section 6.3) to run an F-Interop user's code on the IoT-lab to verify conformance and interoperability. *Lead: Remy Leone*.
- **Scalable Security Solution with H2020 ARMOUR**. Through the H2020 ARMOUR project, Inria-EVA is testing whether security solutions standardized at the IETF are scalable. *Lead: Malisa Vucinic*.

7. New Results

7.1. Wireless Sensor Networks

7.1.1. Deployment of Wireless Sensor Networks

Participants: Ines Khoufi, Pascale Minet, Anis Laouiti.

In 2016, we studied two types of deployment for wireless sensor networks:

- those ensuring full area coverage and network connectivity;
- those covering some given Points of Interest (PoI) and ensuring network connectivity.

Deployment of sensor nodes to fully cover an area has caught the interest of many researchers. However, some simplifying assumptions are adopted such as the knowledge of obstacles, a centralized algorithm... To cope with these drawbacks, we propose OA-DVFA (Obstacles Avoidance Distributed Virtual Forces Algorithm), a self-deployment algorithm to ensure full area coverage and network connectivity. This fully distributed algorithm is based on virtual forces to move sensor nodes. We show how to avoid the problem of node oscillations and to detect the end of the deployment in a distributed way. We evaluate the impact of the number, shape and position of obstacles on the coverage rate, the distance traveled by all nodes and the number of active nodes. Simulation results show the very good behavior of OA-DVFA. This work done in collaboration with Anis Laouiti has been presented at the CCNC 2016 conference [35].

We also focus on wireless sensor networks deployed to cover some given Points of Interest (PoIs), achieve connectivity with the sink and be robust against link and node failures. The Relay Node Placement problem (RNP) consists in minimizing the number of relays needed and the maximum length of the paths connecting each PoI with the sink. We propose a solution that determines the positions of relay nodes based on the virtual grid computed by the optimal deployment for full area coverage. We compare our solution with two different solutions based respectively on (1) the straight line that builds the shortest path between each PoI and the sink, (2) the Steiner point that connects PoIs together. We then extend these algorithms to achieve k-connectivity. Our solution outperforms the Steiner points solution in terms of maximum path length on the one hand, and the straight line solution in terms of total number of relay nodes deployed on the other hand. We also apply our solution in an area containing obstacles and show that it provides very good performances. This study has been presented at the MASS 2016 conference [34].

7.1.2. Path Planning of Mobile Wireless Nodes Gathering Data

Participants: Ines Khoufi, Pascale Minet, Nadjib Achir.

Mobile wireless nodes in charge of collecting data from static wireless sensor nodes constitute a very attractive solution for wireless sensor networks, WSNs, where the application requirements in terms of node autonomy are very strong unlike the requirement in terms of latency. Mobile nodes allow wireless sensor nodes to save energy.

In 2016 we focused on the path planning problem of mobile wireless nodes gathering data according two different objectives:

- to ensure the monitoring of a given area;
- to visit some given Points of Interest (PoI) in a delay less than a given latency.

For the first objective, we are interested in area monitoring using Unmanned Aerial Vehicles (UAVs). Basically, we propose a path planning approach for area monitoring where UAVs are considered as mobile collectors. The area to be monitored is divided into cells. The goal is to determine the path of each UAV such that each cell is covered by exactly one UAV, fairness is ensured in terms of the number of cells visited by each UAV and the path of each UAV is minimized. To meet our goal, we proceed in two steps. In the first step, we assign to each UAV the cells to visit. In the second step, we optimize the path of each UAV visiting its cells. For the first step, we propose two solutions. The first solution is based on cluster formation, each cluster is made up of the set of cells monitored by a same UAV. The second solution is based on game theory and uses coalition formation to determine the cells to be monitored by each UAV. In the second step and for both solutions, we propose to apply optimization techniques to minimize the path of each UAV that visits all its cells. This study done in collaboration with Nadjib Achir was presented at the PEMWN 2016 conference [32].

For the second objective, we use game theory to model the problem. Game theory is often used to find equilibria where no player can unilaterally increase its own payoff by changing its strategy without changing the strategies of other players. In this paper, we propose to use coalition formation to compute the optimized tours of mobile sinks in charge of collecting data from static wireless sensor nodes. The associated coalition formation problem has a stable solution given by the final partition obtained. However, the order in which the players play has a major impact on the final result. We determine the best order to minimize the number of mobile sinks needed. We evaluate the complexity of this coalition game in terms of the number of rounds and the processing time needed to get convergence, as well as the impact of the number of collect points on the number of mobile sinks needed and on the maximum tour duration of these mobile sinks. In addition, we show how to extend the coalition game to support different latencies for different types of data. Finally, we formalize our problem as a multi-objective optimization problem. We compare the coalition game with a genetic algorithm: for 20 nodes to visit, the coalition game requires a processing time 327 times less than the genetic algorithm. The coalition game provides a scalable solution. These results have been presented at the IPCCC 2016 conference. This work was done in cooperation Mohamed-Amine Koulali and Abdellatif Kobbane [33].

7.1.3. Centralized Scheduling in TSCH-based Wireless Sensor Networks

Participants: Erwan Livolant, Pascale Minet, Thomas Watteyne.

Scheduling in an IEEE802.15.4e TSCH(Time Slotted Channel Hopping 6TiSCH) low-power wireless network can be done in a centralized or distributed way. When using centralized scheduling, a scheduler installs a communication schedule into the network. This can be done in a standards-based way using CoAP. In this study, we compute the number of packets and the latency this takes, on real-world examples. The result is that the cost is very high using today's standards, much higher than when using an ad-hoc solution such as OCARI. We conclude by making recommendations to drastically reduce the number of messages and improve the efficiency of the standardized approach.

7.1.4. Using an IEEE 802.15.4e TSCH network

Participants: Ines Khoufi, Pascale Minet, Erwan Livolant, Thomas Watteyne.

Most wireless sensor networks that are currently deployed use a technology based on the IEEE 802.15.4 standard. However, this standard does not meet all requirements of industrial applications in terms of latency, throughput and robustness. That is why the IEEE 802.15.4e amendment has been designed, including the Time Slotted Channel Hopping (TSCH) mode.

In 2016, we evaluated the time needed for a joining node to detect beacons advertising the TSCH network. This time may be long due to channel hopping in the TSCH network. The beacon advertising policy is left unspecified by the standard. We propose DBA, a Deterministic Beacon Advertising algorithm. DBA ensures a collision-free and regular transmission of beacons on all the frequencies used by the TSCH network. DBA outperforms two solutions already published that are Random Horizontal and Random Vertical. Some results have been presented as a poster at the IPCCC 2016 conference [48].

The medium access in a TSCH network is ruled by a schedule that determines for each pair (slot offset, channel offset) the transmitting node(s) and the receiving node(s). Each node in the TSCH network must have this schedule. The question is how to install it on all nodes. We proposed and evaluated different ways of installing a schedule in a TSCH network, comparing them in terms of the number of messages required. This study has been presented at the AdHocNow 2016 conference [36].

7.1.5. The OCARI Wireless Sensor Network

Participants: Erwan Livolant, Pascale Minet, Mohammed Tahar Hammi.

Wireless Sensor Networks and Industrial Internet of Things use smart, autonomous and usually limited capacity devices in order to sense and monitor industrial environments. The devices in a wireless sensor network are managed by a controller, which should authenticate them before they join the network. OCARI is a wireless sensor network technology providing optimized protocols in order to reduce the energy consumption.

To enhance OCARI security and ensure a robust authentication of devices, we propose a strong authentication method based on the One Time Password algorithm and deployed at the MAC layer. This method is specially designed to be implemented on devices with low storage and computing capacities. This work has been done in collaboration with Mohammed Tahar Hammi from Telecom ParisTech and presented at the PEMWN 2016 conference [30].

We also evaluated the performances of the building of an OCARI network. The goal was to identify the most time consuming steps among node association, neighborhood discovery, routing tree building, stabilization of the routing tree and node coloring.

7.1.6. Security in Wireless Sensor Networks

Participants: Selma Boumerdassi, Paul Muhlethaler.

Sensor networks are often used to collect data from the environment where they are located. These data can then be transmitted regularly to a special node called a *sink*, which can be fixed or mobile. For critical data (like military or medical data), it is important that sinks and simple sensors can mutually authenticate so as to avoid data to be collected and/or accessed by fake nodes. For some applications, the collection frequency can be very high. As a result, the authentication mechanism used between a node and a sink must be fast and efficient both in terms of calculation time and energy consumption. This is especially important for nodes which computing capabilities and battery lifetime are very low. Moreover, an extra effort has been done to develop alternative solutions to secure, authenticate, and ensure the confidentiality of sensors, and the distribution of keys in the sensor network. Specific researches have also been conducted for large-scale sensors. At present, we work on an exchange protocol between sensors and sinks based on low-cost shifts and xor operations. This study was published in [21]. After this publication, we have been working on the performance evaluation of the solution to determine the memory overhead together with both computing and communication latencies.

7.1.7. Massive MIMO Cooperative Communications for Wireless Sensor Networks

Participants: Nadjib Achir, Paul Muhlethaler.

This work is a collaboration with Mérouane Debbah (Supelec, France).

The objective of this work is to propose a framework for massive MIMO cooperative communications for Wireless Sensor Networks. Our main objective is to analyze the performances of the deployment of a large number of sensors. This deployment should cope with a high demand for real time monitoring and should also take into account energy consumption. We have assumed a communication protocol with two phases: an initial training period followed by a second transmit period. The first period allows the sensors to estimate the channel state and the objective of the second period is to transmit the data sensed. We start analyzing the impact of the time devoted to each period. We study the throughput obtained with respect to the number of sensors when there is one sink. We also compute the optimal number of sinks with respect to the energy spent for different values of sensors. This work is a first step to establish a complete framework to study energy efficient Wireless Sensor Networks where the sensors collaborate to send information to a sink. Currently, we are exploring the multi-hop case.

7.2. Machine Learning for an efficient and dynamic management of network resources and services

7.2.1. Machine Learning in Networks

Participants: Dana Marinca, Nesrine Ben Hassine, Pascale Minet.

Machine learning techniques can be used to improve the quality of experience for the end users of Content Delivery Networks (CDNs). In a CDN, the most popular video contents are cached near the end-users in order to minimize the contents delivery latency. The idea developed hereafter consists in using prediction techniques to evaluate the future popularity of video contents in order to decide which ones should be cached. The popularity of a video content is evaluated by the number of daily requests for this content.

We consider various prediction methods, called experts, coming from different fields (e.g. statistics, control theory). To evaluate the accuracy of the experts' popularity predictions, we assess these experts according to three criteria: cumulated loss, maximum instantaneous loss and best ranking. The loss function expresses the discrepancy between the prediction value and the real number of requests. We use real traces extracted from YouTube to compare different prediction methods and determine the best tuning of their parameters. The goal is to find the best trade-off between complexity and accuracy of the prediction methods used.

We also show the importance of a decision maker, called forecaster, that predicts the popularity based on the predictions of selection of several experts. The forecaster based on the best K experts outperforms in terms of cumulated loss the individual experts' predictions and those of the forecaster based on only one expert, even if this expert varies over time. This study has been presented at the IWCMC 2016 conference [18].

In the paper presented at the WiMob 2016 conference [18], we apply these prediction methods to caching. We first selected the best experts in charge of predicting the popularity of video contents in real traces of YouTube. We tuned the parameters of the DES expert. We proved that the well-known LFU caching strategy can also be considered as a prediction based strategy on the Basic expert. Simulation results show that the DES prediction-based caching strategy provides similar Hit Ratio to the well-known LFU caching strategy. These results are usually close to the optimal ones that can be achieved only when knowing in advance the popularity of each video content for the next day, which is an unrealistic assumption. The exception occurs when a content whose popularity was very poor becomes suddenly very popular with millions of solicitations. In such a case, the accuracy of prediction methods becomes poor. This opens a research direction where the knowledge of societal events is taken into account to improve the prediction.

7.3. Protocols and Models for Wireless Networks - Application to VANETs

7.3.1. Protocols for VANETs

7.3.1.1. An Infrastructure-Free Slot Assignment Algorithm for Reliable Broadcast of Periodic Messages in VANETs

Participants: Mohamed Elhadad, Paul Muhlethaler, Anis Laouiti.

We have designed a novel **D**istributed **T**DMA based **MAC** protocol, named DTMAC, developed specifically for a highway scenario. DTMAC is designed to provide the efficient delivery of both periodic and event-driven safety messages. The protocol uses the vehicles' location and a new slot reuse concept to ensure that vehicles in adjacent areas have a collision-free schedule. Simulation results and analysis in a highway scenario have been carried out to evaluate the performance of DTMAC and compare it with the VeMAC protocol.

We propose a completely distributed and infrastructure free TDMA scheduling scheme which exploits the linear feature of VANET topologies. The vehicles' movements in a highway environment are linear due to the fact that their movements are constrained by the road topology. Our scheduling mechanism is also based on the assumption that each road is divided into N small fixed areas, denoted by $x_i, i = 1, ..., N$ (see Figure 1). Area IDs can be easily derived using map and GPS Information.

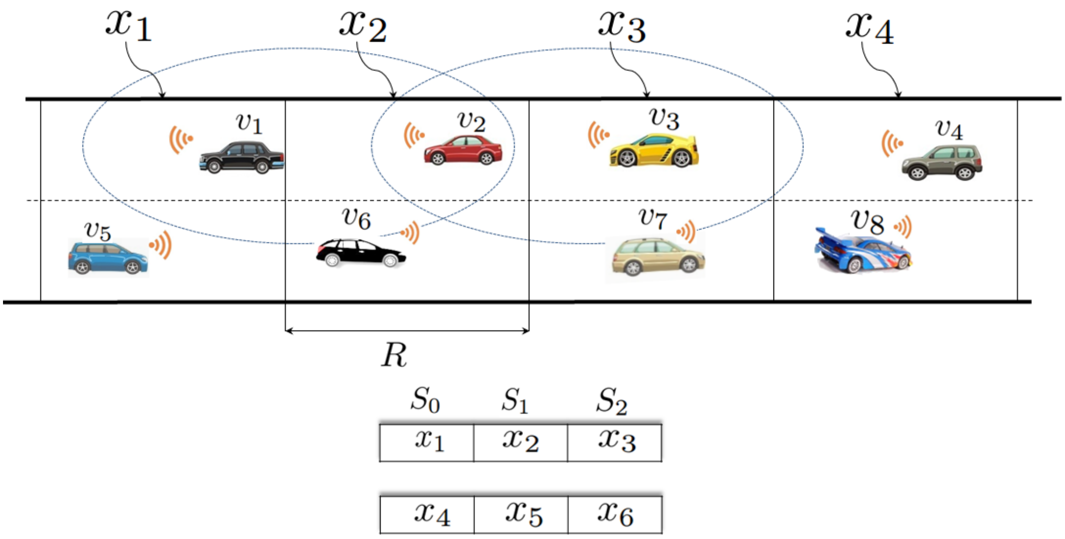

Figure 1. TDMA slots scheduling principle.

The time slots in each TDMA frame are partitioned into three sets S_0, S_1 and S_2 associated with vehicles in three contiguous areas: x_i, x_{i+1} and x_{i+2}, respectively (see Figure 1). Each frame consists of a constant number of time slots, denoted by τ and each time slot is of a fixed time duration, denoted by s. Each vehicle can detect the start time of each frame as well as the start time of a time slot. In the VANET studied, all the vehicles are equipped with a GPS and thus the one-Pulse-Per-Second (1PPS) signal that a GPS receiver gets from GPS satellites can be used for slot synchronization.

To prevent collisions on the transmission channel, our TDMA scheduling mechanism requires that every packet transmitted by any vehicle must contain additional information, called Frame Information (FI). For the frame, this field gives the status of the slot (Idle, Busy, Collision) and the ID of the vehicles accessing each slot with the characteristic of the data sent: periodic or event-driven safety messages.

The simulation results show that, compared to VeMAC which is the reference in terms of TDMA protocols for VANETs, DTMAC provides a lower rate of access and merging collisions, which results in significantly improved broadcast coverage. For further details see [27].

7.3.1.2. TRPM: a TDMA-aware routing protocol for multi-hop communications in VANETs

Participants: Mohamed Elhadad Or Hadded, Paul Muhlethaler, Anis Laouiti.

The main idea of TRPM is to select the next hop using the vehicle position and the time slot information from the TDMA scheduling. Like the GPSR protocol, we assume that each transmitting vehicle knows the position of the packet's destination. In TRPM, the TDMA scheduling information and the position of a packet's destination are sufficient to make correct forwarding decisions at each transmitting vehicle. Specifically, if a source vehicle is moving in area x_i, the locally optimal choice of next hop is the neighbor geographically located in area x_{i+1} or x_{i-1} according to the position of the packet's destination. As a result, the TDMA slot scheduling obtained by DTMAC can be used to determine the set of next hops that are geographically closer to the destination. In fact, each vehicle that is moving in the area x_i can know the locally optimal set of next hops that are located in adjacent areas x_{i+1} or x_{i-1} by observing the set of time slots $S_{(i+3)\%3}$ or $S_{(i+1)\%3}$, respectively. We consider the same example presented above when vehicle G as the destination vehicle that will broadcast a message received from vehicle A. As shown in Figure 2, only two relay vehicles are needed to ensure a multi-hop path between vehicle A and G (one relay node in the area x_2 and another one in the area x_3).

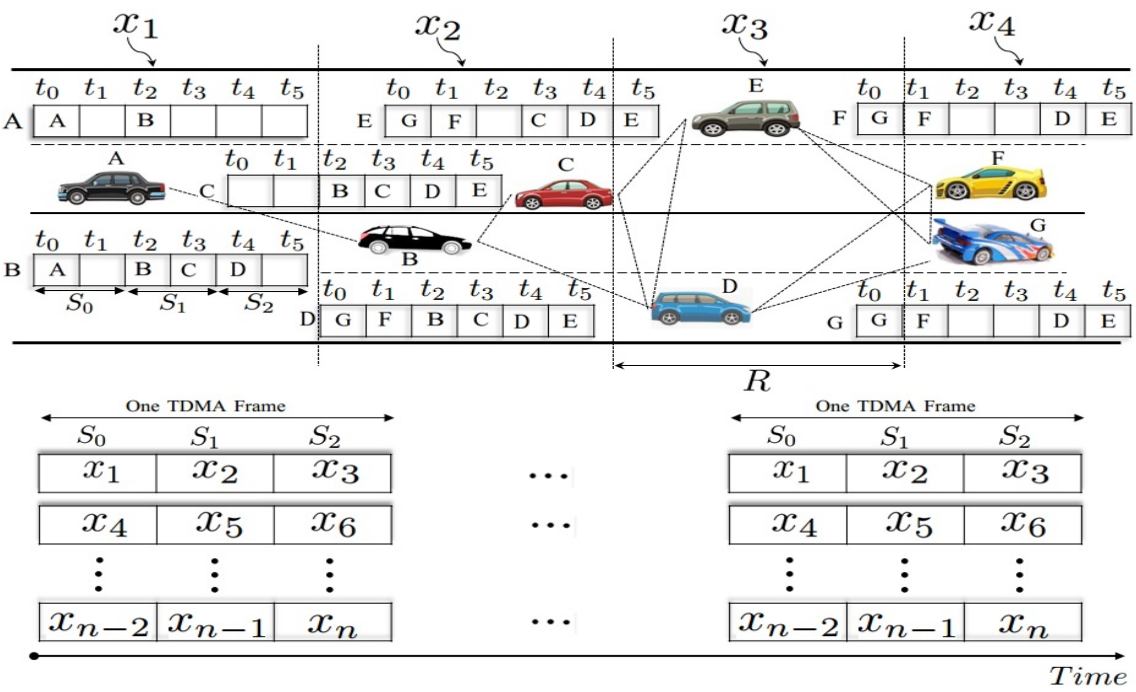

Figure 2. VANET network using DTMAC scheduling scheme.

In the following, the DTMAC protocol has been used by the vehicles to organize the channel access. The TDMA slot scheduling obtained by DTMAC is illustrated in Figure 2. Firstly, vehicle A forwards a packet to B, as vehicle A uses its frame information to choose a vehicle that is accessing the channel during the set S_1. Upon receiving the packet for forwarding, vehicle B will choose by using its frame information a vehicle that's accessing the channel during the set of time slots S_2 (say vehicle D). Then, vehicle D will forward the packet to G, as G is moving in area x_4 (accessing the channel during the set S_0) and it is the direct neighbor of vehicle D. By using DTMAC as the MAC layer, we can note that the path A-B-D-G is the shortest, in terms of the number of hops as well as the end-to-end delay which is equal to 6 time slots (2 time slots between t_0 and t_2 as t_2 is the transmission slot for vehicle B, then 2 time slots between t_2 and t_4 as t_4 is the transmission slot for vehicle D and finally 2 time slots between t_4 and t_0 as t_0 is the transmission slot in which vehicle G will broadcast the message received from vehicle A).

The idea of TRPM is the following. Whenever a vehicle i accessing the channel during the set S_k wants to send/forward an event-driven safety message, it constructs two sets of candidate forwarders based on its frame information FI as follows, where $TS(j)$ indicates the time slot reserved by vehicle j.

- $A_i = \{j \in N(i) \mid TS(j) \in S_{(k+1)\%3}\}$ *// The set of vehicles that are moving in the adjacent right-hand area.*
- $B_i = \{j \in N(i) \mid TS(j) \in S_{(k+2)\%3}\}$ *// The set of vehicles that are moving in the adjacent left-hand area.*

Each source vehicle uses the position of a packet's destination and the TDMA scheduling information to make packet forwarding decisions. In fact, when a source vehicle i is moving behind the destination vehicle, it will select a next hop relay that belongs to set B_i; when the transmitter is moving in front of the destination vehicle, it will select a forwarder vehicle from those in set A_i. For each vehicle i that will send or forward a message, we define the normalized weight function WHS (Weighted next-Hop Selection) which depends on the delay and the distance between each neighboring vehicle j. WHS is calculated as follows:

$$WHS_{i,j} = \alpha * \frac{\Delta t_{i,j}}{\tau} - (1-\alpha) * \frac{d_{i,j}}{R} \qquad (1)$$

Where:

- τ is the length of the TDMA frame (in number of time slots).
- j is one of the neighbors of vehicle i, which represents the potential next hop that will relay the message received from vehicle i.
- $\Delta t_{i,j}$ is the gap between the sending slot of vehicle i and the sending slot of vehicle j.
- $d_{i,j}$ is the distance between the two vehicles i and j, and R is the communication range.
- α is a weighted value in the interval $[0,1]$ that gives more weight to either distance or delay. When α is high, more weight is given to the delay. Otherwise, when α is small, more weight is given to the distance.

We note that the two weight factors $\frac{\Delta t_{i,j}}{\tau}$ and $\frac{d_{i,j}}{R}$ are in conflict. For simplicity, we assume that all the factors should be minimized. In fact, the multiplication of the second weight factor by (-1) allows us to transform a maximization to a minimization. Therefore, the forwarding vehicle for i is the vehicle j that is moving in an adjacent area for which $WHS_{i,j}$ is the lowest value.

The simulation results reveal that our routing protocol significantly outperforms other protocols in terms of average end-to-end delay, average number of relay vehicles and the average delivery ratio.

7.3.1.3. CTMAC: a Centralized TDMA for VANETs

Participants: Mohamed Elhadad Or Hadded, Paul Muhlethaler, Anis Laouiti.

We have designed an infrastructure-based TDMA scheduling scheme which exploits the linear feature of VANET topologies. The vehicles' movements in a highway environment are linear due to the fact that their movements are constrained by the road topology. Our scheduling mechanism is also based on the assumption that the highway is equipped with some RSUs (i.e. one RSU for each $2 \times R\ meters$, where R is the communication range). Note that each area is covered by one RSU installed on the side of the highway and in the middle of the corresponding area. The time slots in each TDMA frame are partitioned into two sets S_1, S_2 associated with vehicles in two adjacent RSU areas (see Figure 3). Each frame consists of a constant number of time slots, denoted by τ and each time slot is of a fixed time duration, denoted by s. Each vehicle can detect the start time of each frame as well as the start time of a time slot.

The CTMAC scheduling mechanism uses a slot reuse concept to ensure that vehicles in adjacent areas covered by two RSUs have a collision-free schedule. The channel time is partitioned into frames and each frame is further partitioned into two sets of time slots S_1 and S_2. These sets are associated with vehicles moving in the adjacent RSU areas. These sets of time slots are reused along the highway in such a way that no vehicles belonging to the same set of two-hop neighbors using the same time slot. As shown in Figure 3, the vehicles in the coverage area of RSU_1 and those in the coverage area of RSU_2 are accessing disjoint sets of time slots. As a result, the scheduling mechanism of CTMAC can decrease the collision rate by avoiding the inter-RSUs interference without using any complex band. Each active vehicle keeps accessing the same time slot on all subsequent frames unless it enters another area covered by another RSU or a merging collision problem occurs. Each vehicle uses only its allocated time slot to transmit its packet on the control channel.

The simulation results reveal that CTMAC significantly outperforms the VeMAC and ADHOC MAC protocols. in terms of transmission collisions and the overhead required to create and maintain the TDMA schedules, see [28].

7.3.1.4. A Flooding-Based Location Service in VANETs

Participants: Selma Boumerdassi, Paul Muhlethaler.

This work was done in collaboration with Eric Renault, Telecom Sud Paris.

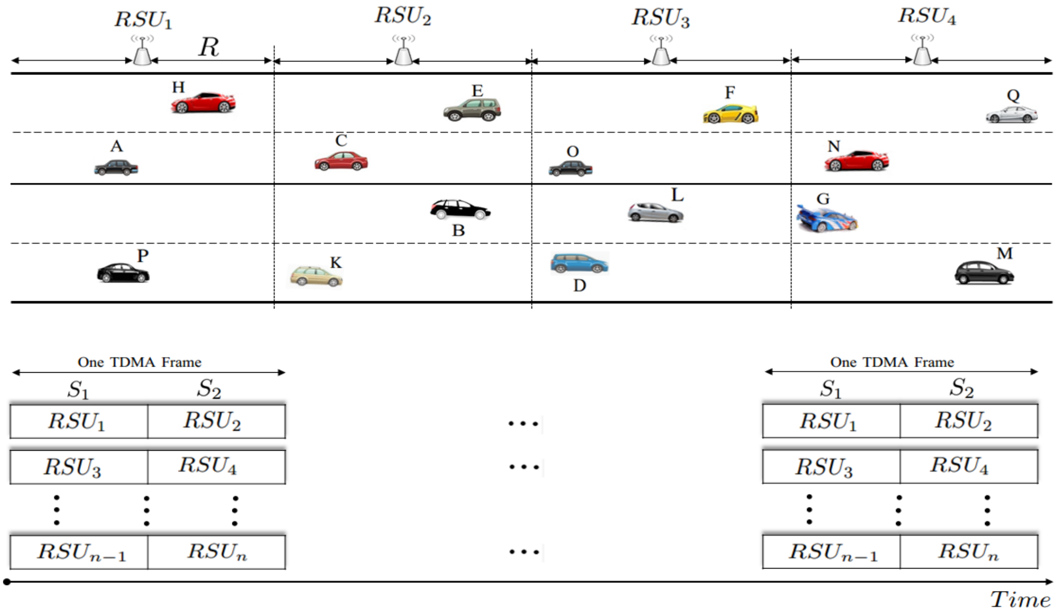

Figure 3. TDMA slots scheduling mechanism of CTMAC

We have designed and analyzed a location service for VANETs; such a service can be used in Location-based routing protocols for VANETs. Our protocol is a proactive flooding-based location service that drastically reduces the number of update packets sent over the network as compared to traditional flooding-based location services. This goal is achieved by partially forwarding location information at each node. A mathematical model and some simulations are proposed to show the effectiveness of this solution. Cases for 1D, 2D and 3D spaces are studied for both deterministic and probabilistic forwarding decisions. We compare our protocol with the Multi-Point Relay (MPR) technique which is used in the OLSR protocol and determine the best technique according to the network conditions.

7.3.2. Models for Wireless Networks and VANETs

7.3.2.1. Performance analysis of IEEE 802.11 broadcast schemes with different inter-frame spacings

Participants: Younes Bouchaala, Paul Muhlethaler, Nadjib Achir.

This work has been in collaboration with Oyunchimeg Shagdar (Vedecom).

We have started to build a model which analyzes the performance of IEEE 802.11p managing different classes of priorities. The differentiation of traffic streams is obtained with different inter-frame spacings: AIFSs (for Arbitration Inter Frame Spacings) and with different back-off windows: CWs (for Collision Windows). This model is based on a Markov model where the state is the remaining number of idle slots that a packet of a given class has to wait before transmission. However, in addition to this Markov model for which we compute a steady state we also consider the Markov chain which counts the number of idle slots after the smallest AIFS. As a matter of fact the probability these states are not evenly distributed since with different AIFSs the arrival rate is not constant when the number of idle slots experienced after the smallest AIFS varies. The resolution of the steady state of these two inter-mixed Markov chains lead to non linear and intertwined equations that can be easily solved with a software such as Maple. With the model we have obtained, we can compute the delivery rate of packets of different classes and show the influence of system parameters: AIFSs and CWs. The preliminary results show a a very strong influence of different AIFSs on the performance for each traffic streams.

7.3.2.2. Model of IEEE 802.11 Broadcast Scheme with Infinite Queue

Participant: Paul Muhlethaler.

This work has been in collaboration with Guy Fayolle (Inria RITS).

We have analyzed the so-called back-off technique of the IEEE 802.11 protocol in broadcast mode with waiting queues. In contrast to existing models, packets arriving when a station (or node) is in back-off state are not discarded, but are stored in a buffer of infinite capacity. As in previous studies, the key point of our analysis hinges on the assumption that the time on the channel is viewed as a random succession of transmission slots (whose duration corresponds to the length of a packet) and mini-slots during which the back-off of the station is decremented. These events occur independently, with given probabilities. The state of a node is represented by a two-dimensional Markov chain in discrete-time, formed by the back-off counter and the number of packets at the station. Two models are proposed both of which are shown to cope reasonably well with the physical principles of the protocol. The stability (ergodicity) conditions are obtained and interpreted in terms of maximum throughput. Several approximations related to these models are also discussed. The results of this study are in [2].

7.3.2.3. Model and optimization of CSMA

Participants: Younes Bouchaala, Paul Muhlethaler, Nadjib Achir.

This work has been in collaboration with Oyunchimeg Shagdar (Vedecom).

We have studied the maximum throughput of CSMA in scenarios with spatial reuse. The nodes of our network form a Poisson Point Process (PPP) of a one- or two-dimensional space. The one-dimensional PPP well represents VANETs. To model the effect of Carrier Sense Multiple Access (CSMA), we give random marks to our nodes and to elect transmitting nodes in the PPP we choose the nodes with the smallest marks in their neighborhood, this is the Matern hardcore selection process. To describe the signal propagation, we use a signal with power-law decay and we add a random Rayleigh fading. To decide whether or not a transmission is successful, we adopt the Signal-over-Interference Ratio (SIR) model in which a packet is correctly received if its transmission power divided by the interference power is above a capture threshold. We assume that each node in our PPP has a random receiver at a typical distance. We choose the average distance to its closest neighbor. We also assume that all the network nodes always have a pending packet. With these assumptions, we analytically study the density of throughput of successful transmissions and we show that it can be optimized with the carrier-sense threshold. The model makes it possible to analytically compute the performance of a CSMA system and gives interesting results on the network performance such as the capture probability when the throughput is optimized, and the effect on a non-optimization of the carrier sense threshold on the throughput. We can also study the influence of the parameters and see their effects on the overall performance. We observe a significant difference between 2D an 1D networks.

We have built two models to compare the spatial density of successful transmissions of CSMA and Aloha. To carry out a fair comparison, we optimize both schemes by adjusting their parameters. For spatial Aloha, we can adapt the transmission probability, whereas for spatial CSMA we have to find the suitable carrier sense threshold. The results obtained show that CSMA, when optimized, outperforms Aloha for nearly all the parameters of the network model values and we evaluate the gain of CSMA over Aloha. We also find interesting results concerning the effect of the model parameters on the performance of both Aloha and CSMA. The closed formulas we have obtained provide immediate evaluation of performance, whereas simulations may take minutes to give their results. Even if Aloha and CSMA are not recent protocols, this comparison of spatial performance is new and provides interesting and useful results.

For Aloha networks, when we study transmissions over the average distance to the closest neighbor, the optimization does not depend on the density of nodes, which is a very interesting property. Thus in Aloha networks, the density of successful transmissions easily scales linearly in λ when we vary λ whereas in CSMA networks the protocol must be carefully tuned to obtain this scaling.

7.3.2.4. Adaptive CSMA

Participants: Nadjib Achir, Younes Bouchaala, Paul Muhlethaler.

This work has been in collaboration with Oyunchimeg Shagdar (Vedecom).

Using the model we have built for CSMA, we have shown that when optimized with the carrier sense detection threshold P_{cs}, the probability p^* of transmission for a node in the CSMA network does not depend on the density of nodes λ. In other words when the CSMA is optimized to obtain the largest density of successful transmissions (communication from nodes to their neighbors), p^* is constant. We have verified this statement on several examples and we think that a formal proof of this remark is possible using scaling arguments. The average access delay is a direct function of the probability of transmission p. Thus the average delay when the carrier sense detection threshold is optimized is a constant D_{target} which does not depend on λ. A stabilization algorithm which adapts P_{cs} to reach the D_{target} can thus be envisioned.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. CNES

Participants: Ines Khoufi, Pascale Minet, Erwan Livolant.

Partners: CNES, Inria.

Following the SAHARA project that ended in 2015, CNES decided to fund a study about the use of wireless sensor networks in space environment. This new project started in November 2015 and will end in November 2016.

For CNES we studied how to use a IEEE 802.15.4e TSCH (Time Slotted Channel Hopping) network in the space launch vehicles. We proposed new solutions and evaluated their performances with the NS3 simulation tool.

8.1.2. OpenMote

Participant: Thomas Watteyne.

Inria-EVA has signed an long-standing Memorandum of Understanding with OpenMote Technology.

8.2. Bilateral Grants with Industry

8.2.1. Gridbee CIFRE

Participants: Jonathan Muñoz, Thomas Watteyne.

- Title: km-scale Industrial Networking
- Type: CIFRE agreement
- Period: Nov 2015 - Oct 2018
- Coordinator: **Thomas Watteyne**
- Goal: CIFRE agreement with Gridbee (http://www.gridbeecom.com/) to apply 6TiSCH-style scheduling on top of long-range IEEE802.15.4g radios. Implementation of those solutions on OpenWSN.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

The Inria-EVA team has not been involved in an ANR project in 2016.

9.1.2. Competitivity Clusters

9.1.2.1. SAHARA

Participants: Pascale Minet, Erwan Livolant.

Period: 2011 - 2016.

Partners: EADS (coordinator), Astrium, BeanAir, CNES, ECE, EPMI, Eurocopter, GlobalSys, Inria, LIMOS, Oktal SE, Reflex CES, Safran Engineering Systems.

SAHARA is a FUI project, labeled by ASTECH and PEGASE, which aims at designing a wireless sensor network embedded in an aircraft. The proposed solution should improve the embedded mass, the end-to-end delays, the cost and performance in the transfers of non critical data.

This project ended in March 2016. After a presentation of the SAHARA project at the IEEE WISEE 2015 conference (Wireless for Space and Extreme Environments), we were selected to write a book chapter entitled "Multichannel Wireless Sensor Networks for Aircraft: Challenges and Issues" in the Wiley book "Wireless sensor systems for extreme environments: space, underwater, underground and industrial".

9.1.2.2. CONNEXION

Participants: Pascale Minet, Ines Khoufi, Erwan Livolant.

Period: 2012 - 2016.

Partners: EDF (coordinator), All4Tec, ALSTOM, AREVA, Atos WorldGrid, CEA, CNRS / CRAN, Corys TESS, ENS Cachan, Esterel Technologies, Inria, LIG, Predict, Rolls-Royce Civil Nuclear, Telecom ParisTech.

The Cluster CONNEXION (Digital Command Control for Nuclear EXport and renovatION) project aims to propose and validate an innovative architecture platforms suitable control systems for nuclear power plants in France and abroad. This architecture integrates a set of technological components developed by the academic partners (CEA, Inria, CNRS / CRAN, ENS Cachan, LIG, Telecom ParisTech) and based on collaborations between major integrators such as ALSTOM and AREVA, the operator EDF in France and "techno-providers" of embedded software (Atos WorldGrid, Rolls-Royce Civil Nuclear, Corys TESS, Esterel Technologies, All4Tec, Predict). With the support of the competitiveness clusters System@tic, Minalogic and Burgundy Nuclear Partnership,the project started in April 2012. The key deliverables of the project covered several topics related demonstration concern-driven engineering models for the design and validation of large technical systems, design environments and evaluation of HMI, the implementation of Wireless Sensor Network context-nuclear, buses business object or real-time middleware facilitating the exchange of heterogeneous data and distributed data models standardized to ensure consistency of digital systems.

The EVA team focuses more particularly on the interconnection of the OCARI wireless sensor network with the industrial facility backbone and deployment algorithms of wireless sensors.

In the Cluster Connexion project, the goal for the EVA team was to design and implement new functionnalities for the OCARI wireless sensor network to allow it to:

- support the mobility of some sensor nodes (targeted application: remote dosimetry to monitor the exposition of people to radiations),
- transmit commands to sensors/actuators (e.g. configuration parameters, regeneration order),
- ensure data gathering during node recoloring,
- remotely manage the parameters of the OCARI network,
- aggregate in a single frame several heterogeneous measures originated from different sensors on a same wireless node,
- use a generic format for the measures: type, length, value.
- integrate this network to the middleware of context-aware services, OPC-UA/ROSA.

The demonstrator “a mobile connected worksite” developed in the Cluster Connexion project meets several objectives:

- Make the wireless sensor networks more reliable in an ionisating environment ionisant;
- Make easier the diagnostic and the repairing by means of the aggregation of data originated from heterogeneous sources;
- Take into account the requirements of information security in the architectures;
- Ensure a continuum of solutions for the industrial involved.

The Industrial IoT (Internet of Things) solution proposed by Connexion is an integrated chain, from the wireless sensor& actuator network up to the surveillance, diagnostic and health infrastructure monitoring applications, using a context-aware middleware fitting the industrial environment.
At the end of the Cluster Connexion project, we made the demonstration of a command/control loop for the regeneration of wireless sensor nodes in collaboration with CEA, Predict, Telecom ParisTech, EDF, ATOS and Inria highlighting the following steps:

- the upstream flow of health indicators from electronic devices,
- detection of an abnormal behavior by a monitoring software (KASEM),
- generation of a regeneration command and transmission of this command to the misbehaving sensor node.
- regeneration of the involved sensor
- insertion of the regenerated sensor in the OCARI network.

When the Cluster Connexion project ended, the results obtained with regard to the OCARI network and the OPC-UA/ROSA middleware have been transferred to the Task Force ConnexSensors hosted by AFNeT. The goals of the ConnexSensors TaskForce are:

- Federate industrial companies around an IoT solution IoT including wireless sensor & actuator networks and a standardized industrial middleware.
- Jointly valorize the OCARI wireless sensor & actuator network and the OPC-UA/ROSA middleware.
- Deploy the Connexion demonstrator in the basemenet of interested industrials.
- Ensure that industrials will keep the mastership of their data.
- Ensure the perennity of the solution proposed.

9.1.3. Other collaborations

EVA has a collaboration with Vedecom. **Paul Muhlethaler** supervises Younes Bouchaala’s PhD funded by Vedecom. This PhD aims at studying vehicle-to-vehicle communication to improve roads safety.

9.2. European Initiatives

9.2.1. H2020 Projects

9.2.1.1. F-Interop

Type: H2020

Objective: Design and implement a cloud-based interoperability testing platform for low-power wireless standards.

Duration: Nov 2015 - Oct 2017

Coordinator: UPMC (FR)

Other partners: iMinds (BE), ETSI (FR), EANTC (DE), Mandat International (CH), DigiCat (UK), UL (LU), Inria (FR), Device Gateway (CH)

Inria contact: **Thomas Watteyne**

9.2.1.2. ARMOUR

Type: H2020

Objective: Security for the IoT

Duration: Dec 2015 – Nov 2017

Coordinator: UPMC (FR)

Other partners: Inria (FR), Synelixis (EL), Smartesting (FR), Unparallel (PT), JRC (BE), Ease Global Market (FR), Odin Solutions (ES)

Inria-EVA contact: **Thomas Watteyne**

9.2.1.3. Project Reviewing

- **Paul Muhlethaler** was reviewer for the E3Network project (E-band transceiver for the backhaul infrastructure of the future networks). The transceiver designed in the E3Network project will use modern digital multi-level modulations to achieve high spectral efficiency. This together with the huge bandwidth will enable high capacities above 10 Gbps.

9.2.2. Collaborations in European Programs, Except H2020

The Inria-EVA team has not participated in non-H2020 European Programs in 2016.

9.2.3. Collaborations with Major European Organizations

European Telecommunications Standards Institute (ETSI)

co-organize two ETSI 6TiSCH plugtests in 2016 (in Paris in February, in Berlin in July).

9.3. International Initiatives

9.3.1. Inria International Labs

9.3.1.1. REALMS Associate Team

Type: Associate Team

Inria International Lab: Inria@SiliconValley

Title: Real-Time Real-World Monitoring Systems

Associate teams: Inria-EVA, Prof. Glaser's team (UC Berkeley), Prof. Kerkez's team (University of Michigan, Ann Arbor)

Duration: 2015-2017

Objective: Prof. Glaser's and Prof. Kerkez's teams are revolutionizing environmental monitoring by using low power wireless TSCH networks to produce continuous environmental data accessible in real time. They are successfully deploying these networks to study mountain hydrology, observe water quality in urban watersheds, and build intelligent urban stormwater grids. The REALMS associate team conducts research across the environmental engineering and networking research domains. Its 3-year goal is to develop easy-to-use real-world network monitoring solutions to provide real-time data for environmental and urban applications. This goal leads to the following objectives: building a long-term large-scale public connectivity dataset of the networks deployed; using that dataset to model TSCH networks; and building an ecosystem of tools around this technology.

website: http://www.snowhow.io/

Inria contact: **Thomas Watteyne**

9.3.1.2. DIVERSITY Associate Team

Type: Associate Team

Inria International Lab: Inria@SiliconValley

Title: Measuring and Exploiting Diversity in Low-Power Wireless Networks

Associate teams: Inria-EVA, Prof. Bhaskar Krishnamachari's team, USC, CA, USA

Duration: 2016-2018

Objective: The Grand Challenge of the DIVERSITY associate team is to develop the networking technology for tomorrow's Smart Factory. The two teams come with a perfectly complementary background on standardization and experimentation (Inria-EVA) and scheduling techniques (USC-ANRG). The key topic addressed by the joint team will be networking solutions for the Industrial Internet of Things (IIoT), with a particular focus on reliability and determinism.

Inria contact: **Thomas Watteyne**

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

9.3.2.1. Tassili

The Tassili project (N° MDU 17MDU988 - Campus France N° 37459VF) "Gestion des caches et orchestration intelligentes dans un environnement réseau virtulaisé" is a project in collaboration with Algeria and France. On the French side, the project is leaded by Samia Bouzefrane (associated professor at CNAM) and **Paul Muhlethaler** (EVA team Inria). On the Algerian side is leaded by the University Mouloud Mammeri of Tizi-Ouzou (UMMTO) représented by Mehammed Daoui (associated professor).

This project will start in January 2017 and will last three years. Three PhD theses will be conducted in co-tutelle between CNAM and UMMTO. This project will support the stay of the tree PhD candidates for a four months visit in France. These two PhD theses will be co-directed by **Paul Muhlethaler**. The first subject is "New intelligent caching and mobility strategies for MEC/ICN based architectures" and the second subject concern the design of a safe architecture for Name Data Networking.

9.3.3. Inria International Partners

9.3.3.1. Declared Inria International Partners

University of California, Berkeley, CA, USA (Glaser)

- Collaboration with Prof. Steven Glaser, Ziran Zhang, Carlos Oroza, Sami Malek and Zeshi Zheng through the REALMS associate team, see Section 9.3.1.1.
- Joint publication in 2016:
  - Long-term Monitoring of the Sierra Nevada Snowpack Using Wireless Sensor Networks. Ziran Zhang, Steven Glaser, Thomas Watteyne, Sami Malek. IEEE Internet of Things Journal, special issue on Large-scale Internet of Things: Theory and Practice, to appear in 2016.
  - Demo: SierraNet: Monitoring the Snow Pack in the Sierra Nevada. Keoma Brun-Laguna, Carlos Oroza, Ziran Zhang, Sami Malek, Thomas Watteyne, Steven Glaser. ACM International Conference on Mobile Computing and Networking (MobiCom), Workshop on Challenged Networks (CHANTS), 7 October 2016, New York, NY, USA.
  - (Not so) Intuitive Results from a Smart Agriculture Low-Power Wireless Mesh Deployment. Keoma Brun-Laguna, Ana Laura Diedrichs, Diego Dujovne, Rémy Léone, Xavier Vilajosana, Thomas Watteyne. ACM International Conference on Mobile Computing and Networking (MobiCom), Workshop on Challenged Networks (CHANTS), 7 October 2016, New York, NY, USA.
  - SOL: An End-to-end Solution for Real-World Remote Monitoring Systems. Keoma Brun-Laguna, Thomas Watteyne, Sami Malek, Ziran Zhang, Carlos Oroza, Steven Glaser, Branko Kerkez. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Valencia, Spain, 4-7 September 2016.

University of Southern California, CA, USA

- Collaboration with Prof. Bhaskar Krishnamachari through the DIVERSITY associate team, see Section 9.3.1.2.
- Joint publication in 2016:
  - Insights into Frequency Diversity from Measurements on an Indoor Low Power Wireless Network Testbed. Pedro Henrique Gomes, Ying Chen, Thomas Watteyne, Bhaskar Krishnamachari. IEEE Global Telecommunications Conference (GLOBECOM), Workshop on Low-Layer Implementation and Protocol Design for IoT Applications (IoT-LINK), Washington, DC, USA, 4-8 December 2016.
  - Reliability through Time-Slotted Channel Hopping and Flooding-based Routing. Pedro Henrique Gomes, Thomas Watteyne, Pradipta Gosh, Bhaskar Krishnamachari. International Conference on Embedded Wireless Systems and Networks (EWSN), Dependability Competition, ACM, Graz, Austria, 14-15 February 2016.

Universidad Tecnologica Nacional, Mendoza, Argentina

- Collaboration with Ana Laura Diedrichs, Juan Carlos Taffernaberry, Gustavo Mercado through the SticAmSud PEACH project.
- Joint publication(s) in 2016:
  - PEACH: Predicting Frost Events in Peach Orchards Using IoT Technology. Thomas Watteyne, Ana Laura Diedrichs, Keoma Brun-Laguna, Javier Emilio Chaar, Diego Dujovne, Juan Carlos Taffernaberry, Gustavo Mercado. EAI Endorsed Transactions on the Internet of Things, to appear in 2016.
  - A Demo of the PEACH IoT-based Frost Event Prediction System for Precision Agriculture. Keoma Brun-Laguna, Ana Laura Diedrichs, Javier Emilio Chaar, Diego Dujovne, Juan Carlos Taffernaberry, Gustavo Mercado, Thomas Watteyne. IEEE International Conference on Sensing, Communication and Networking (SECON), poster and demo session, London, UK, 27-30 June 2016

University of Michigan, Ann Arbor, MI, USA

- Collaboration with Prof. Branko Kerkez through the REALMS associate team, see Section 9.3.1.1.

Linear Technology/Dust Networks, Silicon Valley, USA

- Collaboration with Prof. Kris Pister, Dr. Brett Warneke, Dr. Lance Doherty, Dr. Jonathan Simon and Joy Weiss on SmartMesh IP and 6TiSCH standardization.

9.3.3.2. Informal International Partners

University of California, Berkeley, CA, USA (Pister)

- Collaboration with Prof. Kris Pister through the IETF 6TiSCH working group.
- Joint publication in 2016:
  - Simple Distributed Scheduling with Collision Detection in TSCH Networks. Kazushi Muraoka, Thomas Watteyne, Nicola Accettura, Xavi Vilajosana, Kris Pister. IEEE Sensors Letters, to appear in 2016.

Open University of Catalunya, Spain

- Collaboration with Xavi Vilajosana and Pere Tuset through IETF 6TiSCH working group and the OpenWSN project.
- Joint publication(s) in 2016:
  - Distributed PID-based Scheduling for 6TiSCH Networks. Marc Domingo-Prieto, Tengfei Chang, Xavier Vilajosana, Thomas Watteyne. IEEE Communications Letters, vol PP, Issue 99, March 2016.

  - Poster Abstract: A Benchmark for Low-power Wireless Networking. Simon Duquennoy, Olaf Landsiedel, Carlo Alberto Boano, Marco Zimmerling, Jan Beutel, Mun Choon Chan, Omprakash Gnawali, Mobashir Mohammad, Luca Mottola, Lothar Thiele, Xavier Vilajosana, Thiemo Voigt, Thomas Watteyne. ACM Conference on Embedded Networked Sensor Systems (ACM Sensys), Stanford, CA, USA, 14-16 November 2016.
  - Rover: Poor (but Elegant) Man's Testbed. Zacharie Brodard, Hao Jiang, Tengfei Chang, Thomas Watteyne, Xavier Vilajosana, Pascal Thubert, Geraldine Texier. ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks (PE-WASUN), Valletta, Malta, 13-17 November 2016.
  - Determinism Through Path Diversity: Why Packet Replication Makes Sense. Jesica de Armas, Pere Tuset, Tengfei Chang, Ferran Adelantado, Thomas Watteyne, Xavier Vilajosana. International Conference on Intelligent Networking and Collaborative Systems (INCoS), Ostrava, Czech Republic, 7-9 September 2016.
  - OpenWSN & OpenMote: Demo'ing A Complete Ecosystem for the Industrial Internet of Things. Tengfei Chang, Pere Tuset-Peiro, Jonathan Munoz, Xavier Vilajosana, Thomas Watteyne. IEEE International Conference on Sensing, Communication and Networking (SECON), poster and demo session, London, UK, 27-30 June 2016.
  - OpenMote+: a Range-Agile Multi-Radio Mote. Pere Tuset, Xavier Vilajosana, Thomas Watteyne. International Conference on Embedded Wireless Systems and Networks (EWSN), NexMote Workshop, ACM, Graz, Austria, 14-15 February 2016.
  - Numerous IETF Internet-Drafts.

University of Science and Technology, Beijing, China

- Collaboration with Qin Wang through IETF 6TiSCH working group. Tengfei Chang, engineer at Inria-EVA, comes from her team
- Joint publication(s) in 2016:
  - On-the-Fly Bandwidth Reservation for 6TiSCH Wireless Industrial Networks. Maria-Rita Palattella, Thomas Watteyne, Qin Wang, Kazuki Muraoka, Nicola Accettura, Diego Dujovne, Alfredo Grieco, Thomas Engel. IEEE Sensors Journal, 15 January 2016.
  - LLSF: Low Latency Scheduling Function for 6TiSCH Networks. Tengfei Chang, Thomas Watteyne, Qin Wang, Xavier Vilajosana. IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS), Washington, DC, USA, 26-28 May 2016.

University of Bari, Italy

- Collaboration with Savio Sciancalepore, Giuseppe Piro and Gennaro Boggia through IETF 6TiSCH and OpenWSN.
- Joint publication in 2016:
  - Link-layer Security in TSCH Networks: Effect on Slot Duration. Savio Sciancalepore, Malisa Vucinic, Giuseppe Piro, Gennaro Boggia, Thomas Watteyne. Wiley Transactions on Emerging Telecommunications Technologies (ETT), to appear in 2016.

University of Trento, Italy

- Collaboration with Oana Iova through IETF 6TiSCH working group.

- Joint publication(s) in 2016:
  - The Love-Hate Relationship between IEEE802.15.4 and RPL. Oana Iova, Fabrice Theoleyre, Thomas Watteyne, Thomas Noel. IEEE Communications Magazine, to appear in 2016.

TU Berlin, Germany

- Collaboration with Vlado Handziski, Adam Wolisz through IETF 6TiSCH working group.
- Joint publication(s) in 2016:
  - Industrial Wireless IP-based Cyber Physical Systems. Thomas Watteyne, Vlado Handziski, Xavier Vilajosana, Simon Duquennoy, Oliver Hahm, Emmanuel Baccelli, Adam Wolisz. Proceedings of the IEEE, Vol. PP, Issue 99, pp. 1-14, March 2016.

Mandat International, Switzerland

- Collaboration with Sebastien Ziegler through the H2020 F-Interop project
- Joint publication(s) in 2016:
  - F-Interop – Online Platform of Interoperability and Performance Tests for the Internet of Things. Sébastien Ziegler, Serge Fdida, Cesar Viho, Thomas Watteyne. Conference on Interoperability in IoT (InterIoT), Paris, France, 26-28 October 2016.

KU Leuven, Belgium

- Collaboration with Prof. Danny Hughes, Prof. Wouter Joosen, Dr. Nelson Matthys, Fan Yang, Wilfried Daniels on MicroPnP.

Inria-EVA has a strong relationship with ENSI (Tunisia) and ENSIAS (Morocco). A significant part of our PhD students come from these engineering schools.

9.3.4. Participation in Other International Programs

9.3.4.1. PEACH

Program: STIC-AmSud 2015

Title: PEACH - PrEcision Agriculture through Climate researcH

Inria principal investigator: **Thomas Watteyne**

International Partners (Institution - Laboratory - Researcher):

Escuela de Informática y Telecomunicaciones, Universidad Diego Portales, Santiago, Chile. Coordinator: Prof. Diego Dujovne

Universidad Tecnológica Nacional - Facultad Regional Mendoza, Grupo de I&D en Tecnologías de la Información y Comunicaciones (GridTICS). Coordinator: Prof. Gustavo Mercado

DHARMa Lab, Universidad Tecnológica Nacional, Facultad Regional Mendoza, Argentina.

Cátedra de Fisiología Vegetal, Facultad de Ciencias Agrarias, Universidad Nacional de Cuyo, Mendoza, Argentina.

Duration: 2016-2017

Goal: Propose a design methodology for a lowpower wireless IoT sensing network, given the requirements and restrictions of a Machine Learning model to predict frost events in peach orchards and vineyards.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Professors/Researchers:
  - **Mario Gerla**, Professor, UCLA, USA, visit 10-20 December 2016
  - **Leila Saidane**, Professor, ENSI, Tunis, Tunisia, visit November 2016
  - **Felipe Lalanne**, Reseacher, Inria Chile, Chile, visit 19–26 October 2016
  - **Mario Gerla**, Professor, UCLA, USA, visit 31 August - 23 September 2016
  - **Diego Dujovne**, Professor, Universidad Diego Portales, Chile, visit 22-31 July 2016
  - **Ruben Milocco**, Universidad Nacional Comahue, Argentina, visit July 2016
  - **Branko Kerkez**, Professor, U. Michigan, USA, visit 17-22 June 2016
  - **Steven Glaser**, Professor, UC Berkeley, USA, visit 21-25 March 2016
  - **Xavi Vilajosana**,Professor, UOC/OpenMote, Spain, visit 2-4 February 2016
- PhD Students:
  - **Travis Massey**, PhD Student, UC Berkeley, USA, visit 22 July 2016
  - **Carlos Oroza**, PhD Student, UC Berkeley, USA, visit 23-29 July 2016
  - **David Burnett**, PhD Student, UC Berkeley, USA, visit 13-15 June 2016
  - **Filip Barac**, PhD Student, Mid Sweden University, Sweden, visit 8-19 February 2016

9.4.2. Internships

- **Jiangnan Yang**, internship on simulation of wireless TDMA networks with NS3, March-August 2016.

9.4.3. Visits to International Teams

9.4.3.1. Research Stays Abroad

- **Keoma-Brun Laguna**, stay in Prof. Glaser's lab at UC Berkeley, USA, August 2016.
- **Thomas Watteyne**, stay in Prof. Glaser and Prof. Pister's labs at UC Berkeley, USA, August 2016.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Event Organization

10.1.1.1. General Chair, Scientific Chair

- **Pascale Minet**
  1. co-chair with Leila Saidane of PEMWN 2016, organized in Paris (at the CNAM), France, November 2016.
- **Thomas Watteyne**
  1. Co-chair with Fabrice Theoleyre and Antoine Gallais of a special session on "Industrial Internet of Things: Constraints, Guarantees, and Resiliency" at the 23rd IEEE International Conference on Telecommunications (ICT), Thessaloniki, Greece, 16-18 May 2016.
  2. Technical Program Committee Chair and Local Chair, EAI Conference on Interoperability in IoT (InterIoT), Paris, France, October 2016.
  3. Demo Chair, IEEE International Conference on Sensing, Communication and Networking (SECON), London, UK, 27-30 June 2016.
  4. General Chair, Second ETSI 6TiSCH plugtests, Paris, France, 2-4 February 2016.
- Nadjib Achir

1. track chair of the Internet of Things (IoT) track of the Selected Areas in Communications Symposium of IEEE Global Telecommunications Conference 2014

- Selma Boumerdassi
  1. chair of the International Workshop on Energy Management for Sustainable Internet-of-Things and Cloud Computing (EMSICC 2016), August 2016.
  2. chair of the International Conference on Mobile, Secure and Programmable Networking (MSPN 2016), June 2016.

10.1.1.2. Member of the Organizing Committee

- **Pascale Minet**
  1. chaired the session entitled "Wireless Sensor networks II" at the IEEE MASS 2016 conference held in Brasilia in October 2016.
- **Paul Muhlethaler**
  1. have invited two presenters Laurent Georges (ESIEE) and Nadjib Ait Saadi (Paris XII) at EVA-MIMOVE-RITs seminar. Laurent Georges presented his work on software radio and Nadjib Ait Saadi on virtualisation and communication in data centers.
- **Thomas Watteyne**
  1. Demo Chair, IEEE International Conference on Sensing, Communication and Networking (SECON), London, UK, 27-30 June 2016.
- Christine Anocq
  1. member of the organizing committee of the international conference PEMWN 2016

10.1.2. Scientific Events Selection

The list of conferences Inria-EVA researcher participated in as Technical Program Committee (TPC) members:

- **Nadjib Achir**
  1. TPC Member IEEE Global Telecommunications Conference, GLOBECOM 2016.
  2. TPC Member IEEE International Conference on Communications, ICC 2016.
  3. TPC Member Personal, Indoor and Mobile Radio Communications, PIMRC 2016.
  4. TPC Member IEEE Wireless Communications and Networking Conference, WCNC 2016.
  5. TPC Member IEEE Consumer Communications & Networking Conference, CCNC 2016.
  6. TPC Member Global Information Infrastructure and Networking Symposium, GIIS 2016.
  7. TPC Member International Conference On Network of the Future, NoF 2016.
  8. TPC Member of the fourth International conference on Performance Evaluation and Modeling in Wireless Networks, PEMWN 2016.
- **Selma Boumerdassi**
  1. TPC Member IEEE Global Telecommunications Conference, GLOBECOM 2016
  2. TPC Member IEEE International Conference on Communications, ICC 2016
  3. TPC Member International Conference on Open and Big Data, OBD, 2016
  4. TPC Member International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks, PEMWN, 2016
  5. TPC Member International Conference on Platform Technology and Service, PlatCon, 2016
- **Pascale Minet**
  1. CSCN 2016, IEEE Conference on Standards for Communications and Networking, October 2016.

2. CCNC 2016, 13th Annual IEEE Consumer Communications and Networking Conference, January 2016.
3. DCNET 2016, 7th International Conference on Data Communication Networking, July 2016.
4. PEMWN 2016, 5th International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks, November 2016.
5. PECCS 2016, 6th International conference on Pervasive and Embedded Computing and Communication Systems, July 2016.
6. RTNS 2016, 24th International Conference on Real-Time and Network Systems, November 2016.
7. Wireless Days, IFIP international conference, March 2016.
8. WiSEE 2016, IEEE International Conference on Wireless for Space and Extreme Environments, September 2016.

- **Paul Muhlethaler**
  1. 3rd International Workshop on Energy Management for Sustainable Internet-of-Things and Cloud Computing (EMSICC 2016) 24 August 2016 Vienna.
  2. Technical Committee of the International Conference on Mobile, Secure and Programmable Networking MSPN' 2016, June 1 - 3 2016, Paris, France.
  3. Steering Committee Member of MobileHealth 2016, 6th EAI International Conference on Wireless Mobile Communication and Healthcare, November 14–16, 2016, Milan, Italia
  4. Technical Committee of the fourth International conference on Performance Evaluation and Modeling in Wireless Networks, PEMWN 2016, 22-24 November 2016, Paris, France.
- **Thomas Watteyne**
  1. TPC Member IFIP/IEEE International Symposium on Integrated Network Management, workshop on Future Networks for Secure Smart Cities (FNSSC), 2017.
  2. TPC Member IEEE International Conference on Communications (ICC), Selected Areas in Communications (SAC), 2015, 2016, 2017.
  3. TPC Member IEEE International Conference on Telecommunications (ICT), Thessaloniki, Greece, 16-18 May 2016.
  4. TPC Member IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2008, 2009, 2013, 2016.
  5. TPC Member International Conference on Embedded Wireless Systems and Networks (EWSN), 2016.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- **Nadjib Achir**
  1. guest editor of the special issue "Planning and Deployment of Wireless Sensor Networks", of the International Journal of Distributed Sensor Networks.
- **Thomas Watteyne**
  1. Editor, EAI Transactions on Internet of Things since 2015.
  2. Editor, IEEE Internet of Things (IoT) Journal 2014-2016.

10.1.3.2. Reviewer - Reviewing Activities

- **Nadjib Achir**
  1. Sensor Networks (MDPI)

2. Wireless Communications and Mobile Computing (Wiley)
3. Internet of Things Journal (IEEE)
4. Ad Hoc Networks Journal (Elsevier)

- **Selma Boumerdassi**
  1. The Journal of Human-centric Computing and Information Sciences (Springer).
- **Pascale Minet**
  1. Annals of Telecommunications,
  2. Ad Hoc Networks,
  3. International Journal of Distributed Sensor Networks,
  4. International Journal of Networked and Distributed Computing,
  5. Journal of Sensor and Actuator Networks,
  6. Computer Communications Journal,
  7. Sensors Journal.
- **Paul Muhlethaler**
  1. Annals of telecommunications
  2. Ad Hoc Networks,
  3. Journal of Distributed Sensor Networks,
  4. Journal of Networked and Distributed Computing,
  5. IEEE Transactions on Wireless Communications
  6. IEEE Transactions on Vehicular Technology
  7. IEEE Transactions on Information Theory
  8. International Journal of Distributed Sensor Networks. Hindawi.

10.1.4. Invited Talks and Panels

- **Thomas Watteyne**: SmartMesh IP: a ready-to-use IoT Solution. Inria-Chile, 2 December 2016.
- **Thomas Watteyne**: From Smart Dust to 6TiSCH: Academic and Commercial Background on TSCH Technology, Sensor Platform for HEalthcare in a Residential Environment (SPHERE) seminar, University of Bristol, Bristol, UK, 6 October 2016.
- **Thomas Watteyne** was panelist on “A Not So Politically Correct Reality Check about the IoT”, ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc), Paderborn, Germany, 5-8 July 2016.
- **Thomas Watteyne**: A Not So Politically Correct Reality Check about the IoT. ACM MobiHoc, Paderborn, Germany, 6 July 2016.
- **Thomas Watteyne** was panelist on “IoT & quantified self” at the open Inria conference on “Digital Sciences and Technologies for Health”, Futur en Seine Festival, Paris, France, 10 June 2016.
- **Thomas Watteyne**: HeadsUp! Long-Term Real-Time Patient Position Monitoring. International Conference on Digital Sciences and Technologies for Health, Futur en Seine Festival, Paris, France, 10 June 2016.
- **Thomas Watteyne** was panelist on “IoT & Smart City” at the Inria@Silicon Valley annual workshop (BIS), Paris, France, 9 June 2016.
- **Thomas Watteyne**: PEACH: Predicting Frost Events in Peach Orchards Using IoT Technology, Inria@Silicon Valley annual workshop (BIS), “IoT & Smart City” panel, Paris, France, 9 June 2016.
- **Thomas Watteyne**: Overview (Industrial) IoT Standardization Efforts at IETF. ITU Meeting, Geneva, Switzerland, 10 May 2016.

- **Thomas Watteyne**: Having fun with Industrial IoT, University of Southern California, Los Angeles, CA, USA, 1 April 2016.
- **Thomas Watteyne**: Not-so-Politically Correct Food for Thought on NGIoT, US-Europe Invited Workshop on Next-Generation IoT (NGIOT), Los Angeles, CA, USA, 31 March 2016.
- **Thomas Watteyne**: Industrial IoT Standards and not-to-industrial applications, GDR SOC-SIP/RSD, Paris, France, 30 March 2016.

10.1.5. Leadership within the Scientific Community

- **Thomas Watteyne** is co-chair of IETF 6TiSCH standardization working group.
- **Thomas Watteyne** is Senior member of the IEEE since 2015.

10.1.6. Scientific Expertise

- **Paul Muhlethaler**is a reviewer for the European Commission. He regularly reviews project; this year he was at the last review meeting of E3Network a project dedicated to high-speed radio links in the E-band.
- **Thomas Watteyne** and Tengfei Chang consulted for ETSI as part of the "panel of experts" to prepare the Test Description and Golden Image for the ETSI 6TiSCH 3 plugtests, held on 15-16 July 2016 in Berlin, Germany.
- **Thomas Watteyne** and Tengfei Chang consulted for ETSI as part of the "panel of experts" to prepare the Test Description and Golden Image for the ETSI 6TiSCH 2 plugtests, held on 2-4 February 2016 in Paris, France.

10.1.7. Research Administration

- **Pascale Minet** was member of the evaluation committee in charge of recruiting an Assistant Professor at the University of Auvergne, Clermont-Ferrand in April and May 2016.
- **Pascale Minet** was a member of the EDITE Doctoral School's Grant Allocation Committee (June 2016).
- **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!

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

License:

- **Thomas Watteyne** taught a 2-h course on industrial IoT at University of Southern California (USC), USA, 1 April 2016.

Master:

- **Thomas Watteyne** and Keoma Brun-Laguna taught a 4-day course on IoT, with associated hands-on labs. ENSTA ParisTech. Together with Dominique Barthel, 12-15 December 2016.
- **Thomas Watteyne** taught a 1/2-day crash course on the Industrial IoT, ParisTech, 5 October 2016.

10.2.2. Supervision

PhD:

- Mohamed Hadded, "Design and Optimizaion of Access Control Protocols in Vehicular Ad Hoc Networks (VANETs)", EDITE - Télécom Sud-Paris, November 2016, **Paul Muhlethaler** adviser, Anis Laouiti, co-adviser.

PhD in progress:

- Nesrine Ben Hassine “Learning algorithms for network resource management”, UVSQ, Sept 2017 (expected), **Pascale Minet** adviser, co-adviser Dana Marinca.
- Younes Bouchaala “Handling Safety Messages in Vehicular Ad Hoc Networks (VANETs)”, UVSQ, Sept 2017 (expected), **Paul Muhlethaler** adviser.
- Keoma Brun-Laguna “Limits of Time Synchronized Channel Hopping” *(tentative title)*, EDITE - UPMC, Dec 2018 (expected), **Thomas Watteyne** adviser.
- Jonathan Muñoz “km-scale Industrial Networks” *(tentative title)*, EDITE - UPMC, Oct 2018 (expected), **Thomas Watteyne** adviser.

10.2.3. Juries

HdR:

- Gerard Chalhoub, “Enhanced communications in data collection multihop wireless sensor networks”, University of Auvergne, June 2016, **Pascale Minet** examinator.

PhD:

- Guillaume Gaillard, “Opérer les réseaux de l’Internet des Objets”, University of Lyon - INSA. Viva on 19 December 2016. **Pascale Minet** reviewer.
- Mohamed Hadded “Design and Optimization of Access Control Protocols in Vehicular Ad Hoc Networks (VANETs)” Thèse de Telecom Sud Paris. Viva on 30 Novembre 2016. Anis Laouiti examiner, **Paul Muhlethaler** examiner.
- Naourez Mejri, “Vers une ville intelligente au service du citoyen mobile: découverte de l’infrastructure d’accès et gestion intelligente de parkings”, University of La Manouba, ENSI, Tunis. Viva in November 2016. **Pascale Minet** examiner, **Paul Muhlethaler** reviewer.
- Nour Brinis Khay, “Stratégies de collecte de données dans les réseaux de capteurs”, University of La Manouba, ENSI, Tunis. Viva in November 2016. **Pascale Minet** examiner.
- Alexandre Ragaleux, “Mécanismes D’Accès Multiple dans les Réseaux Sans Fil Large Bande”, Université Pierre et Marie Curie. Viva on 22 September 2016. **Nadjib Achir** reviewer.
- Remy Leone “Intelligent Gateway for Low-Power Wireless Networks” (“Passerelle intelligente pour réseaux de capteurs sans fil contraints” in French), Telecom ParisTech, France, under the supervision of Jean Louis Rougier and Vania Conan. Viva on 24 June 2016. **Thomas Watteyne** examiner.
- Kevin Roussel, “Évaluation et amélioration des plates-formes logicielles pour réseaux de capteurs sans-fil, pour optimiser la qualité de service et l’énergie”, Inria Nancy/ Université de Lorraine, France, under the supervision of Ye-Qiong Song and Olivier Zendra. Viva on 3 June 2016. **Thomas Watteyne** examiner.
- Samira Chouikhi “Tolérance aux pannes dans un réseau de capteurs sans fil multi-canal ” University of Paris East Marne-La-Vallée. Viva in June 2016. **Pascale Minet** examiner, **Paul Muhlethaler** reviewer.
- Jin Cui, “Data aggregation in wireless sensor networks”, University of Lyon - INSA. Viva in June 2016. **Pascale Minet** reviewer.
- Mohamed Nidhal Mejri “Securing vehicular networks against denial of service attacks”, Université Paris 13. Viva on 19 May 2016. **Paul Muhlethaler** president, **Nadjib Achir** examiner.
- Chiraz Houaida, “Vers des mécanismes de routage robustes et optimisés dans un réseau sans-fil métropolitain et collaboratif”, University of Toulouse. Viva in May 2016. **Pascale Minet** reviewer.

- Antonio O. Gonga, “Mobility and Multi-channel Communications in Low-power Wireless Networks”, KTH Electrical Engineering, under the supervision of Prof. Mikael Johansson. Viva on 14 January 2016. **Thomas Watteyne** examiner.

10.3. Popularization

10.3.1. Demos

Inria-EVA is working with the Inria-Paris management team to install a permanent demo on the Inria-Paris premises in Q1 2017. Pieces of the same demo have been presented to all people and teams visiting the lab. For example, the high school students from the Lycee Louis-Le-Grand on 1 June 2016.

10.3.2. Videos

The following videos were prepared with the audio/video team of Inria.

1. Interview of **Thomas Watteyne** during the “Nouveaux Arrivants” workshop. To be published.
2. “Save The Peaches” details the Smart Agriculture deployment. Featuring **Thomas Watteyne** and Keoma Brun-Laguna.
   - English version: https://www.youtube.com/watch?v=_qGSH8l0Vkk
   - French version: https://www.youtube.com/watch?v=cZvGw7DyIzI

10.3.3. In the News

1. **Thomas Watteyne** featured in “Experto en IoT realiza webinar con Inria Chile”. Inria-Chile newsletter, December 2017.
2. Interview of **Thomas Watteyne**: “In Argentina alone, connected agriculture could have saved 10,000 jobs in one year”, Inria Homepage, November 2016.
3. Interview of **Thomas Watteyne** in “La Gazette”, the Inria-Paris newsletter, on Smart Agriculture, October 2016.
4. **Thomas Watteyne** featured in “Outcome Report from the US-Europe Invited Workshop on Next-Generation Internet of Things @USC”, Inria-SiliconValley Newsletter, 26 July 2016.
5. Keoma Brun-Laguna presented the PEACH project during the annual Math exhibition in Paris, on 28 May 2016.
6. **Thomas Watteyne** featured in “Inria deploys a wireless sensor network in Argentina to save peaches”, Inria-Paris homepage, 22 April 2016.
7. **Thomas Watteyne** featured in “Great fun at the OpenWSN & OpenMote tutorial at IEEE GLOBECOM in San Diego!”, Inria-SiliconValley Newsletter, 17 March 2016.
8. **Thomas Watteyne** featured in “6TiSCH cherche à concilier protocole IPv6, réseau radio maillé et performance industrielle”, L’embarque, 8 February 2016.

10.3.4. Web Presence & Social Media

- The Inria-EVA team maintains the following websites:
  - https://team.inria.fr/eva/
  - http://www.savethepeaches.com/
  - http://www.snowhow.io/
  - http://www.solsystem.io/
  - http://www.headsup.tech/
- The Inria-EVA team maintains the Twitter accounts:
  - EVA (https://twitter.com/InriaEVA), 180 tweets
  - OpenWSN (https://twitter.com/openWSN), 88 tweets

- savethepeaches (https://twitter.com/peachesthesave), 73 tweets
- HeadsUp! (https://twitter.com/heads_up_tech), 28 tweets
- REALMS team (https://twitter.com/realms_team), 28 tweets
- SOLsystem (https://twitter.com/S0Lsystem), 16 tweets
- snowhow (https://twitter.com/snowhow_io), 14 tweets

11. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] M. DOMINGO-PRIETO, T. CHANG, X. VILAJOSANA, T. WATTEYNE.*Distributed PID-based Scheduling for 6TiSCH Networks*, in "IEEE Communications Letters", March 2016, https://hal.inria.fr/hal-01289628.

[2] G. FAYOLLE, P. MUHLETHALER.*A Markovian Analysis of IEEE 802.11 Broadcast Transmission Networks with Buffering*, in "Probability in the Engineering and Informational Sciences", June 2016, vol. 30, n^{o} 3, 19 [*DOI :* 10.1017/S0269964816000036], https://hal.inria.fr/hal-01166082.

[3] O. IOVA, F. THEOLEYRE, T. WATTEYNE, T. NOEL.*The Love-Hate Relationship between IEEE802.15.4 and RPL*, in "IEEE Communications Magazine", 2016, https://hal.archives-ouvertes.fr/hal-01206377.

[4] K. MURAOKA, T. WATTEYNE, N. ACCETTURA, X. VILAJOSANA, K. PISTER.*Simple Distributed Scheduling with Collision Detection in TSCH Networks*, in "IEEE Sensors Letters", July 2016, https://hal.inria.fr/hal-01319765.

[5] M.-R. PALATTELLA, T. WATTEYNE, Q. WANG, K. MURAOKA, N. ACCETTURA, D. DUJOVNE, L. A. GRIECO, T. ENGEL.*On-the-Fly Bandwidth Reservation for 6TiSCH Wireless Industrial Networks*, in "IEEE Sensors Journal", January 2016, 10 [*DOI :* 10.1109/JSEN.2015.2480886], https://hal.inria.fr/hal-01208256.

[6] S. SCIANCALEPORE, M. VUCINIC, G. PIRO, G. BOGGIA, T. WATTEYNE.*Link-layer Security in TSCH Networks: Effect on Slot Duration*, in "Transactions on Emerging Telecommunications Technologies", July 2016 [*DOI :* 10.1002/ETT.3089], https://hal.inria.fr/hal-01342664.

[7] T. WATTEYNE, A. L. DIEDRICHS, K. BRUN-LAGUNA, J. E. CHAAR, D. DUJOVNE, J. C. TAFFERNABERRY, G. MERCADO.*PEACH: Predicting Frost Events in Peach Orchards Using IoT Technology*, in "EAI Endorsed Transactions on the Internet of Things", June 2016, https://hal.inria.fr/hal-01312685.

[8] T. WATTEYNE, V. HANDZISKI, X. VILAJOSANA, S. DUQUENNOY, O. HAHM, E. BACCELLI, A. WOLISZ.*Industrial Wireless IP-based Cyber Physical Systems*, in "Proceedings of the IEEE", March 2016, p. 1-14 [*DOI :* 10.1109/JPROC.2015.2509186], https://hal.inria.fr/hal-01282597.

[9] 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", December 2016, https://hal.inria.fr/hal-01388391.

Invited Conferences

[10] Z. Brodard, H. Jiang, T. Chang, T. Watteyne, X. Vilajosana, P. Thubert, G. Texier.*Rover: Poor (but Elegant) Man's Testbed*, in "ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks (PE-WASUN)", Valletta, Malta, November 2016, https://hal.inria.fr/hal-01359812.

[11] K. Brun-Laguna, A. L. Diedrichs, D. Dujovne, R. Leone, X. Vilajosana, T. Watteyne.*(Not so) Intuitive Results from a Smart Agriculture Low-Power Wireless Mesh Deployment*, in "CHANTS'16", New York City, United States, September 2016 [*DOI :* 10.1145/2979683.2979696], https://hal.inria.fr/hal-01361333.

[12] P. Muhlethaler, Y. Bouchaala, O. Shagdar, N. Achir.*A simple Stochastic Geometry Model to test a simple adaptive CSMA Protocol: Application for VANETs*, in "PEMWN 2016 - 5th IFIP International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks", Paris, France, Proceedings of the 5th IFIP International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks, November 2016, https://hal.archives-ouvertes.fr/hal-01412607.

International Conferences with Proceedings

[13] N. Achir, Y. Bouchaala, P. Muhlethaler, O. Shagdar.*Comparison of Spatial Aloha and CSMA using Simple Stochastic Geometry Models for 1D and 2D Networks*, in "ICT 2016 - 23rd International Conference on Telecommunications, 2016", Thessalonique, Greece, May 2016 [*DOI :* 10.1109/ICT.2016.7500470], https://hal.inria.fr/hal-01368875.

[14] N. Achir, Y. Bouchaala, P. Muhlethaler, O. Shagdar.*Optimisation of spatial CSMA using a simple stochastic geometry model for 1D and 2D networks*, in "IWCMC 2016 - 12th International Wireless Communications & Mobile Computing Conference", Paphos, Cyprus, Proceedings of the 12th International Wireless Communications & Mobile Computing Conference, September 2016, p. 558 - 563 [*DOI :* 10.1109/IWCMC.2016.7577118], https://hal.archives-ouvertes.fr/hal-01379975.

[15] I. Amdouni, C. Adjih, N. Aitsaadi, P. Muhlethaler.*Experiments with ODYSSE: Opportunistic Duty cYcle based routing for wirelesS Sensor nEtworks*, in "IEEE LCN 2016: The 41st IEEE Conference on Local Computer Networks (LCN), November 7-10, 2016, Dubai, UAE", Dubai, United Arab Emirates, November 2016, https://hal.inria.fr/hal-01407525.

[16] I. Amdouni, C. Adjih, P. Minet, T. Plesse.*Delay analysis of STDMA in grid wireless sensor networks*, in "ICMCIS 2016: International Conference on Military Communications and Information Systems ICMCIS (former MCC) Brussels, Belgium, 23rd - 24th May 2016", Brussels, Belgium, May 2016, p. 1 - 8 [*DOI :* 10.1109/ICMCIS.2016.7496580], https://hal.inria.fr/hal-01405540.

[17] K. Avrachenkov, P. Jacquet, J. K. Sreedharan.*Distributed Spectral Decomposition in Networks by Complex Diffusion and Quantum Random Walk*, in "IEEE Infocom 2016", San Francisco, United States, April 2016, https://hal.inria.fr/hal-01263811.

[18] N. Ben Hassine, D. Marinca, P. Minet, D. Barth.*Expert-based on-line learning and prediction in Content Delivery Networks*, in "IWCMC 2016 - The 12th International Wireless Communications & Mobile Computing Conference", Paphos, Cyprus, September 2016, p. 182 - 187 [*DOI :* 10.1109/IWCMC.2016.7577054], https://hal.inria.fr/hal-01411119.

[19] M. N. BOUATIT, S. BOUMERDASSI, P. MINET, A. DJAMA.*Fault-Tolerant Mechanism for Multimedia Transmission in Wireless Sensor Networks*, in "VTC-Fall 2016 - IEEE 84th Vehicular Technology Conference", Montreal, Canada, September 2016, https://hal.inria.fr/hal-01417601.

[20] Y. BOUCHAALA, P. MUHLETHALER, O. SHAGDAR, N. ACHIR.*Optimized Spatial CSMA for VANETs: A Comparative Study using a Simple Stochastic Model and Simulation Results*, in "CCNC 2017. 8-11 january 2017. Las Vegas.", Las Vegas, United States, Proceedings of CCNC 2017, January 2017, https://hal.archives-ouvertes.fr/hal-01379978.

[21] S. BOUMERDASSI, E. RENAULT, P. MUHLETHALER.*A Stateless Time-based Authenticated-Message Protocol for Wireless Sensor Networks (STAMP)*, in "WCNC 2016 - IEEE Wireless Communications and Networking Conference", Doha, Qatar, April 2016 [*DOI :* 10.1109/WCNC.2016.7564905], https://hal.inria.fr/hal-01251917.

[22] K. BRUN-LAGUNA, T. WATTEYNE, S. MALEK, Z. ZHANG, C. OROZA, S. GLASER, B. KERKEZ.*SOL: An End-to-end Solution for Real-World Remote Monitoring Systems*, in "IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)", Valencia, Spain, IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), September 2016, https://hal.inria.fr/hal-01327798.

[23] T. CHANG, J. ARMAS, X. VILAJOSANA, F. ADELANTADO, T. WATTEYNE.*Determinism Through Path Diversity: Why Packet Replication Makes Sense*, in "Intelligent Networking and Collaborative Systems (INCoS), 2016 International Conference on", Ostrava, Czech Republic, September 2016, https://hal.inria.fr/hal-01419846.

[24] T. CHANG, T. WATTEYNE, W. QIN, X. VILAJOSANA.*LLSF: Low Latency Scheduling Function for 6TiSCH Networks*, in "International Conference on Distributed Computing in Sensor Systems (DCOSS)", Washington, DC, United States, IEEE, May 2016, https://hal.inria.fr/hal-01297645.

[25] P. H. GOMES, Y. CHEN, T. WATTEYNE, B. KRISHNAMACHARI.*Insights into Frequency Diversity from Measurements on an Indoor Low Power Wireless Network Testbed*, in "IEEE Global Telecommunications Conference (GLOBECOM), Workshop on Low-Layer Implementation and Protocol Design for IoT Applications (IoT-LINK)", Washington, DC, United States, December 2016, https://hal.inria.fr/hal-01355060.

[26] P. H. GOMES, T. WATTEYNE, P. GOSH, B. KRISHNAMACHARI.*Reliability through Time-Slotted Channel Hopping and Flooding-based Routing*, in "International Conference on Embedded Wireless Systems and Networks (EWSN), Dependability Competition", Graz, Austria, ACM, February 2016, https://hal.inria.fr/hal-01217190.

[27] M. HADDED, A. LAOUITI, P. MUHLETHALER, L. A. SAIDANE.*An Infrastructure-Free Slot Assignment Algorithm for Reliable Broadcast of Periodic Messages in Vehicular Ad hoc Networks*, in "VTC Fall 2016", Montréal , Canada, Proceedings of VTC Fall 2016, September 2016, https://hal.archives-ouvertes.fr/hal-01379216.

[28] M. HADDED, P. MUHLETHALER, A. LAOUITI, L. A. SAIDANE.*A Centralized TDMA based Scheduling Algorithm for Real-Time Communications in Vehicular Ad Hoc Networks*, in "SoftCom 2016", Split , Croatia, Proceedings of SoftCom 2016, September 2016, https://hal.archives-ouvertes.fr/hal-01379219.

[29] M. HADDED, P. MUHLETHALER, A. LAOUITI, L. A. SAIDANE.*A Novel Angle-based Clustering Algorithm for Vehicular Ad Hoc Networks*, in "IWVSC 2016", Kuala Lumpur , Malaysia, Proceedings of IWVSC 2016, August 2016, https://hal.archives-ouvertes.fr/hal-01379221.

[30] M. T. HAMMI, E. LIVOLANT, P. BELLOT, A. SERHOUCHNI, P. MINET.*MAC Sub-Layer Node Authentication in OCARI*, in "PEMWN 2016 - 5th IFIP International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks", Paris, France, November 2016, https://hal.inria.fr/hal-01417671.

[31] N. B. HASSINE, D. MARINCA, P. MINET, D. BARTH.*Caching strategies based on popularity prediction in content delivery networks*, in "The 12th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications", New York City, United States, October 2016, p. 1 - 8 [*DOI :* 10.1109/WIMOB.2016.7763215], https://hal.inria.fr/hal-01411122.

[32] I. KHOUFI, P. MINET, N. ACHIR.*Unmanned Aerial Vehicles Path Planning for Area Monitoring*, in "PEMWN 2016 : The 5th IFIP International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks", paris, France, November 2016, https://hal.inria.fr/hal-01410071.

[33] I. KHOUFI, P. MINET, M.-A. KOULALI, A. KOBBANE.*Path Planning of Mobile Sinks in Charge of Data Gathering: a Coalitional Game Theory Approach*, in "35th IEEE – International Performance Computing and Communications Conference", Las Vegas, United States, December 2016, https://hal.inria.fr/hal-01410082.

[34] I. KHOUFI, P. MINET, A. LAOUITI.*Fault-Tolerant and Constrained Relay Node Placement in Wireless Sensor Networks*, in "The 13th IEEE International Conference on Mobile Ad hoc and Sensor Systems (IEEE MASS 2016)", Brasilia, Brazil, October 2016, https://hal.inria.fr/hal-01410069.

[35] I. KHOUFI, P. MINET, A. LAOUITI.*OA-DVFA: A Distributed Virtual Forces-based Algorithm to Monitor an Area with Unknown Obstacles*, in "Consumer Communications & Networking Conference (CCNC)", Las-Vegas, United States, Proceedings of Consumer Communications & Networking Conference 2016, January 2016, https://hal.inria.fr/hal-01244802.

[36] E. LIVOLANT, P. MINET, T. WATTEYNE.*The Cost of Installing a 6TiSCH Schedule*, in "AdHoc-Now 2016 - International Conference on Ad Hoc Networks and Wireless", Lille, France, July 2016, https://hal.inria.fr/hal-01302966.

[37] P. TUSET, X. VILAJOSANA, T. WATTEYNE.*OpenMote+: a Range-Agile Multi-Radio Mote*, in "International Conference on Embedded Wireless Systems and Networks (EWSN)", Graz, Austria, ACM, February 2016, p. 333-334, https://hal.inria.fr/hal-01239662.

[38] S. ZIEGLER, S. FDIDA, T. WATTEYNE, C. VIHO.*F-Interop - Online Conformance, Interoperability and Performance Tests for the IoT*, in "Conference on Interoperability in IoT (InterIoT)", Paris, France, October 2016, https://hal.inria.fr/hal-01363427.

[39] J. DE ARMAS, P. TUSET, T. CHANG, F. ADELANTADO, T. WATTEYNE, X. VILAJOSANA.*Determinism Through Path Diversity: Why Packet Replication Makes Sense*, in "International Conference on Intelligent Networking and Collaborative Systems (INCoS)", Ostrava, Czech Republic, France, September 2016, https://hal.inria.fr/hal-01333316.

Scientific Books (or Scientific Book chapters)

[40] T. DANG, P. MINET, P. BELLOT, C. MOZZATI, E. LIVOLANT.*Industrial IoT: Mobility Support and Service Discovery for Industrial Process Monitoring*, in "Internet of Things: Novel Advances and Envisioned Applications", D. P. ACHARJYA (editor), Springer Verlag, December 2016, https://hal.inria.fr/hal-01416381.

[41] P. MINET, G. CHALHOUB, R. SOUA, M. MISSON, E. LIVOLANT, R. DIAB, B. RMILI, J.-F. PERELGRITZ.*Multichannel wireless sensor networks for aircraft: challenges and issues*, in "Wireless sensor systems for extreme environments: space, underwater, underground and industrial", H. RASHVAND, A. ABEDI (editors), John Wiley, December 2016, https://hal.inria.fr/hal-01416367.

Research Reports

[42] I. AMDOUNI, C. ADJIH, N. AITSAADI, P. MUHLETHALER.*ODYSSE: A routing Protocol for Wireless Sensor Networks*, UPEC ; Inria Saclay ; Inria – Centre Paris-Rocquencourt, February 2016, n° RR-8873, https://hal.inria.fr/hal-01292479.

Other Publications

[43] K. BRUN-LAGUNA, A. L. DIEDRICHS, J. E. CHAAR, D. DUJOVNE, J. C. TAFFERNABERRY, G. MERCADO, T. WATTEYNE.*A Demo of the PEACH IoT-based Frost Event Prediction System for Precision Agriculture*, June 2016, SECON 2016 - 13th Annual IEEE International Conference on Sensing, Communication and Networking , Poster [*DOI :* 10.1109/SAHCN.2016.7732963], https://hal.inria.fr/hal-01311527.

[44] K. BRUN-LAGUNA, C. OROZA, Z. ZHANG, S. MALEK, T. WATTEYNE, S. D. GLASER.*Demo: SierraNet: Monitoring the Snowpack in the Sierra Nevada*, October 2016, ACM International Conference on Mobile Computing and Networking (MobiCom), Workshop on Challenged Networks (CHANTS), Poster [*DOI :* 10.1145/2979683.2979698], https://hal.inria.fr/hal-01364041.

[45] T. CHANG, P. TUSET, J. MUNOZ, X. VILAJOSANA, T. WATTEYNE.*OpenWSN & OpenMote: Demo'ing A Complete Ecosystem for the Industrial Internet of Things*, June 2016, SECON 2016 - 13th Annual IEEE International Conference on Sensing, Communication and Networking, Poster [*DOI :* 10.1109/SAHCN.2016.7733004], https://hal.inria.fr/hal-01311264.

[46] T. CHANG, T. WATTEYNE, Q. WANG, X. VILAJOSANA.*Demo: Scheduling Function Zero on a 6TiSCH Network*, February 2017, International Conference on Embedded Wireless Systems and Networks (EWSN), Poster, https://hal.inria.fr/hal-01419913.

[47] S. DUQUENNOY, O. LANDSIEDEL, C. A. BOANO, M. ZIMMERLING, J. BEUTEL, M. C. CHAN, O. GNAWALI, M. MOHAMMAD, L. MOTTOLA, L. THIELE, X. VILAJOSANA, T. VOIGT, T. WATTEYNE.*Poster Abstract: A Benchmark for Low-power Wireless Networking*, November 2016, ACM Conference on Embedded Networked Sensor Systems (SenSys), Poster, https://hal.inria.fr/hal-01377942.

[48] I. KHOUFI, P. MINET, E. LIVOLANT, B. RMILI.*Building an IEEE 802.15.4e TSCH network*, December 2016, 35th IEEE – International Performance Computing and Communications Conference, Poster, https://hal.inria.fr/hal-01410080.

[49] M.-R. PALATTELLA, P. THUBERT, T. WATTEYNE, Q. WANG.*Terminology in IPv6 over the TSCH mode of IEEE 802.15.4e*, December 2016, draft-ietf-6tisch-terminology-08 [work-in-progress], https://hal.inria.fr/hal-01419812.

[50] X. VILAJOSANA, P. TUSET, T. CHANG, T. WATTEYNE.*Standards for the Industrial IoT: a Hands-on Tutorial with OpenWSN and OpenMote*, September 2016, Tutorial at the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Valencia, Spain, https://hal.inria.fr/hal-01312689.

[51] T. WATTEYNE, X. VILAJOSANA, P. TUSET, T. CHANG.*Introduction to the IETF 6TiSCH stack with OpenWSN & OpenMote*, May 2016, Tutorial at the International Conference on Telecommunications (ICT), Thessaloniki, Greece, https://hal.inria.fr/hal-01312691.

Project-Team GALLIUM

Programming languages, types, compilation and proofs

RESEARCH CENTER
Paris

THEME
Proofs and Verification

Table of contents

Project-Team GALLIUM

Creation of the Project-Team: 2006 May 01

Keywords:

Computer Science and Digital Science:

1.1.3. - Memory models
2.1.1. - Semantics of programming languages
2.1.2. - Object-oriented programming
2.1.3. - Functional programming
2.1.6. - Concurrent programming
2.1.11. - Proof languages
2.2.1. - Static analysis
2.2.2. - Memory models
2.2.3. - Run-time systems
2.2.4. - Parallel architectures
2.4.1. - Analysis
2.4.3. - Proofs
2.5.4. - Software Maintenance & Evolution
4.5. - Formal methods for security
7.1. - Parallel and distributed algorithms
7.4. - Logic in Computer Science

Other Research Topics and Application Domains:

5.2.3. - Aviation
6.1. - Software industry
6.3.1. - Web
6.5. - Information systems
6.6. - Embedded systems
9.4.1. - Computer science

1. Members

Research Scientists

Xavier Leroy [Team leader, Senior Researcher, Inria]
Umut Acar [Advanced Research position, Carnegie Mellon University]
Arthur Charguéraud [Researcher, Inria, 40%]
Damien Doligez [Researcher, Inria]
Fabrice Le Fessant [Researcher, Inria]
Luc Maranget [Researcher, Inria]
Michel Mauny [Senior Researcher, Inria]
François Pottier [Senior Researcher, Inria, HDR]
Michael Rainey [Starting Research position, Inria]
Didier Rémy [Senior Researcher, Inria, HDR]

Faculty Member

Pierre Courtieu [Associate Professor on délégation, CNAM, until Aug 2016]

Technical Staff
Sébastien Hinderer [Research Engineer, Inria, 70%, from Apr 2016]

PhD Students
Vitalii Aksenov [Inria]
Armaël Guéneau [ENS Lyon, from Sep 2016]
Jacques-Henri Jourdan [Inria, until Mar 2016, granted by ANR VERASCO project]
Gabriel Scherer [ENS Paris and Inria, until Jan 2016]
Thomas Williams [ENS Paris]

Post-Doctoral Fellows
Adrien Guatto [Inria, until Sep 2016]
Filip Sieczkowski [Inria, until Sep 2016]

Visiting Scientist
Andrea Parri [Sant'Anna School of Advanced Studies, Pisa, Italy, from May 2016 until Nov 2016]

Administrative Assistant
Laurence Bourcier [Inria]

Others
Jacques-Pascal Deplaix [Student Intern, Epitech, from Mar 2016 until Aug 2016]
Felipe Garay [Student Intern, Universidad de Santiago de Chile, from Feb 2016 until Apr 2016]
Ambroise Lafont [Student Intern, École Polytechnique, from Apr 2016 until August 2016]

2. Overall Objectives

2.1. Research at Gallium

The research conducted in the Gallium group aims at improving the safety, reliability and security of software through advances in programming languages and formal verification of programs. Our work is centered on the design, formalization and implementation of functional programming languages, with particular emphasis on type systems and type inference, formal verification of compilers, and interactions between programming and program proof. The OCaml language and the CompCert verified C compiler embody many of our research results. Our work spans the whole spectrum from theoretical foundations and formal semantics to applications to real-world problems.

3. Research Program

3.1. Programming languages: design, formalization, implementation

Like all languages, programming languages are the media by which thoughts (software designs) are communicated (development), acted upon (program execution), and reasoned upon (validation). The choice of adequate programming languages has a tremendous impact on software quality. By "adequate", we mean in particular the following four aspects of programming languages:

- **Safety.** The programming language must not expose error-prone low-level operations (explicit memory deallocation, unchecked array access, etc) to programmers. Further, it should provide constructs for describing data structures, inserting assertions, and expressing invariants within programs. The consistency of these declarations and assertions should be verified through compile-time verification (e.g. static type-checking) and run-time checks.

- **Expressiveness.** A programming language should manipulate as directly as possible the concepts and entities of the application domain. In particular, complex, manual encodings of domain notions into programmatic notations should be avoided as much as possible. A typical example of a language feature that increases expressiveness is pattern matching for examination of structured data (as in symbolic programming) and of semi-structured data (as in XML processing). Carried to the extreme, the search for expressiveness leads to domain-specific languages, customized for a specific application area.
- **Modularity and compositionality.** The complexity of large software systems makes it impossible to design and develop them as one, monolithic program. Software decomposition (into semi-independent components) and software composition (of existing or independently-developed components) are therefore crucial. Again, this modular approach can be applied to any programming language, given sufficient fortitude by the programmers, but is much facilitated by adequate linguistic support. In particular, reflecting notions of modularity and software components in the programming language enables compile-time checking of correctness conditions such as type correctness at component boundaries.
- **Formal semantics.** A programming language should fully and formally specify the behaviours of programs using mathematical semantics, as opposed to informal, natural-language specifications. Such a formal semantics is required in order to apply formal methods (program proof, model checking) to programs.

Our research work in language design and implementation centers on the statically-typed functional programming paradigm, which scores high on safety, expressiveness and formal semantics, complemented with full imperative features and objects for additional expressiveness, and modules and classes for compositionality. The OCaml language and system embodies many of our earlier results in this area [49]. Through collaborations, we also gained experience with several domain-specific languages based on a functional core, including distributed programming (JoCaml), XML processing (XDuce, CDuce), reactive functional programming, and hardware modeling.

3.2. Type systems

Type systems [52] are a very effective way to improve programming language reliability. By grouping the data manipulated by the program into classes called types, and ensuring that operations are never applied to types over which they are not defined (e.g. accessing an integer as if it were an array, or calling a string as if it were a function), a tremendous number of programming errors can be detected and avoided, ranging from the trivial (misspelled identifier) to the fairly subtle (violation of data structure invariants). These restrictions are also very effective at thwarting basic attacks on security vulnerabilities such as buffer overflows.

The enforcement of such typing restrictions is called type-checking, and can be performed either dynamically (through run-time type tests) or statically (at compile-time, through static program analysis). We favor static type-checking, as it catches bugs earlier and even in rarely-executed parts of the program, but note that not all type constraints can be checked statically if static type-checking is to remain decidable (i.e. not degenerate into full program proof). Therefore, all typed languages combine static and dynamic type-checking in various proportions.

Static type-checking amounts to an automatic proof of partial correctness of the programs that pass the compiler. The two key words here are *partial*, since only type safety guarantees are established, not full correctness; and *automatic*, since the proof is performed entirely by machine, without manual assistance from the programmer (beyond a few, easy type declarations in the source). Static type-checking can therefore be viewed as the poor man's formal methods: the guarantees it gives are much weaker than full formal verification, but it is much more acceptable to the general population of programmers.

3.2.1. Type systems and language design.

Unlike most other uses of static program analysis, static type-checking rejects programs that it cannot prove safe. Consequently, the type system is an integral part of the language design, as it determines which programs are acceptable and which are not. Modern typed languages go one step further: most of the language design is determined by the *type structure* (type algebra and typing rules) of the language and intended application area. This is apparent, for instance, in the XDuce and CDuce domain-specific languages for XML transformations [46], [43], whose design is driven by the idea of regular expression types that enforce DTDs at compile-time. For this reason, research on type systems – their design, their proof of semantic correctness (type safety), the development and proof of associated type-checking and inference algorithms – plays a large and central role in the field of programming language research, as evidenced by the huge number of type systems papers in conferences such as Principles of Programming Languages.

3.2.2. Polymorphism in type systems.

There exists a fundamental tension in the field of type systems that drives much of the research in this area. On the one hand, the desire to catch as many programming errors as possible leads to type systems that reject more programs, by enforcing fine distinctions between related data structures (say, sorted arrays and general arrays). The downside is that code reuse becomes harder: conceptually identical operations must be implemented several times (say, copying a general array and a sorted array). On the other hand, the desire to support code reuse and to increase expressiveness leads to type systems that accept more programs, by assigning a common type to broadly similar objects (for instance, the `Object` type of all class instances in Java). The downside is a loss of precision in static typing, requiring more dynamic type checks (downcasts in Java) and catching fewer bugs at compile-time.

Polymorphic type systems offer a way out of this dilemma by combining precise, descriptive types (to catch more errors statically) with the ability to abstract over their differences in pieces of reusable, generic code that is concerned only with their commonalities. The paradigmatic example is parametric polymorphism, which is at the heart of all typed functional programming languages. Many forms of polymorphic typing have been studied since then. Taking examples from our group, the work of Rémy, Vouillon and Garrigue on row polymorphism [55], integrated in OCaml, extended the benefits of this approach (reusable code with no loss of typing precision) to object-oriented programming, extensible records and extensible variants. Another example is the work by Pottier on subtype polymorphism, using a constraint-based formulation of the type system [53]. Finally, the notion of "coercion polymorphism" proposed by Cretin and Rémy[3] combines and generalizes both parametric and subtyping polymorphism.

3.2.3. Type inference.

Another crucial issue in type systems research is the issue of type inference: how many type annotations must be provided by the programmer, and how many can be inferred (reconstructed) automatically by the type-checker? Too many annotations make the language more verbose and bother the programmer with unnecessary details. Too few annotations make type-checking undecidable, possibly requiring heuristics, which is unsatisfactory. OCaml requires explicit type information at data type declarations and at component interfaces, but infers all other types.

In order to be predictable, a type inference algorithm must be complete. That is, it must not find *one*, but *all* ways of filling in the missing type annotations to form an explicitly typed program. This task is made easier when all possible solutions to a type inference problem are *instances* of a single, *principal* solution.

Maybe surprisingly, the strong requirements – such as the existence of principal types – that are imposed on type systems by the desire to perform type inference sometimes lead to better designs. An illustration of this is row variables. The development of row variables was prompted by type inference for operations on records. Indeed, previous approaches were based on subtyping and did not easily support type inference. Row variables have proved simpler than structural subtyping and more adequate for type-checking record update, record extension, and objects.

Type inference encourages abstraction and code reuse. A programmer's understanding of his own program is often initially limited to a particular context, where types are more specific than strictly required. Type inference can reveal the additional generality, which allows making the code more abstract and thus more reuseable.

3.3. Compilation

Compilation is the automatic translation of high-level programming languages, understandable by humans, to lower-level languages, often executable directly by hardware. It is an essential step in the efficient execution, and therefore in the adoption, of high-level languages. Compilation is at the interface between programming languages and computer architecture, and because of this position has had considerable influence on the design of both. Compilers have also attracted considerable research interest as the oldest instance of symbolic processing on computers.

Compilation has been the topic of much research work in the last 40 years, focusing mostly on high-performance execution ("optimization") of low-level languages such as Fortran and C. Two major results came out of these efforts: one is a superb body of performance optimization algorithms, techniques and methodologies; the other is the whole field of static program analysis, which now serves not only to increase performance but also to increase reliability, through automatic detection of bugs and establishment of safety properties. The work on compilation carried out in the Gallium group focuses on a less investigated topic: compiler certification.

3.3.1. Formal verification of compiler correctness.

While the algorithmic aspects of compilation (termination and complexity) have been well studied, its semantic correctness – the fact that the compiler preserves the meaning of programs – is generally taken for granted. In other terms, the correctness of compilers is generally established only through testing. This is adequate for compiling low-assurance software, themselves validated only by testing: what is tested is the executable code produced by the compiler, therefore compiler bugs are detected along with application bugs. This is not adequate for high-assurance, critical software which must be validated using formal methods: what is formally verified is the source code of the application; bugs in the compiler used to turn the source into the final executable can invalidate the guarantees so painfully obtained by formal verification of the source.

To establish strong guarantees that the compiler can be trusted not to change the behavior of the program, it is necessary to apply formal methods to the compiler itself. Several approaches in this direction have been investigated, including translation validation, proof-carrying code, and type-preserving compilation. The approach that we currently investigate, called *compiler verification*, applies program proof techniques to the compiler itself, seen as a program in particular, and use a theorem prover (the Coq system) to prove that the generated code is observationally equivalent to the source code. Besides its potential impact on the critical software industry, this line of work is also scientifically fertile: it improves our semantic understanding of compiler intermediate languages, static analyses and code transformations.

3.4. Interface with formal methods

Formal methods collectively refer to the mathematical specification of software or hardware systems and to the verification of these systems against these specifications using computer assistance: model checkers, theorem provers, program analyzers, etc. Despite their costs, formal methods are gaining acceptance in the critical software industry, as they are the only way to reach the required levels of software assurance.

In contrast with several other Inria projects, our research objectives are not fully centered around formal methods. However, our research intersects formal methods in the following two areas, mostly related to program proofs using proof assistants and theorem provers.

3.4.1. Software-proof codesign

The current industrial practice is to write programs first, then formally verify them later, often at huge costs. In contrast, we advocate a codesign approach where the program and its proof of correctness are developed in interaction, and we are interested in developing ways and means to facilitate this approach. One possibility that we currently investigate is to extend functional programming languages such as OCaml with the ability to state logical invariants over data structures and pre- and post-conditions over functions, and interface with automatic or interactive provers to verify that these specifications are satisfied. Another approach that we practice is to start with a proof assistant such as Coq and improve its capabilities for programming directly within Coq.

3.4.2. Mechanized specifications and proofs for programming languages components

We emphasize mathematical specifications and proofs of correctness for key language components such as semantics, type systems, type inference algorithms, compilers and static analyzers. These components are getting so large that machine assistance becomes necessary to conduct these mathematical investigations. We have already mentioned using proof assistants to verify compiler correctness. We are also interested in using them to specify and reason about semantics and type systems. These efforts are part of a more general research topic that is gaining importance: the formal verification of the tools that participate in the construction and certification of high-assurance software.

4. Application Domains

4.1. High-assurance software

A large part of our work on programming languages and tools focuses on improving the reliability of software. Functional programming, program proof, and static type-checking contribute significantly to this goal.

Because of its proximity with mathematical specifications, pure functional programming is well suited to program proof. Moreover, functional programming languages such as OCaml are eminently suitable to develop the code generators and verification tools that participate in the construction and qualification of high-assurance software. Examples include Esterel Technologies's KCG 6 code generator, the Astrée static analyzer, the Caduceus/Jessie program prover, and the Frama-C platform. Our own work on compiler verification combines these two aspects of functional programming: writing a compiler in a pure functional language and mechanically proving its correctness.

Static typing detects programming errors early, prevents a number of common sources of program crashes (null dereferences, out-of bound array accesses, etc), and helps tremendously to enforce the integrity of data structures. Judicious uses of generalized abstract data types (GADTs), phantom types, type abstraction and other encapsulation mechanisms also allow static type checking to enforce program invariants.

4.2. Software security

Static typing is also highly effective at preventing a number of common security attacks, such as buffer overflows, stack smashing, and executing network data as if it were code. Applications developed in a language such as OCaml are therefore inherently more secure than those developed in unsafe languages such as C.

The methods used in designing type systems and establishing their soundness can also deliver static analyses that automatically verify some security policies. Two examples from our past work include Java bytecode verification [50] and enforcement of data confidentiality through type-based inference of information flow and noninterference properties [54].

4.3. Processing of complex structured data

Like most functional languages, OCaml is very well suited to expressing processing and transformations of complex, structured data. It provides concise, high-level declarations for data structures; a very expressive pattern-matching mechanism to destructure data; and compile-time exhaustiveness tests. Therefore, OCaml is an excellent match for applications involving significant amounts of symbolic processing: compilers, program analyzers and theorem provers, but also (and less obviously) distributed collaborative applications, advanced Web applications, financial modeling tools, etc.

4.4. Rapid development

Static typing is often criticized as being verbose (due to the additional type declarations required) and inflexible (due to, for instance, class hierarchies that must be fixed in advance). Its combination with type inference, as in the OCaml language, substantially diminishes the importance of these problems: type inference allows programs to be initially written with few or no type declarations; moreover, the OCaml approach to object-oriented programming completely separates the class inheritance hierarchy from the type compatibility relation. Therefore, the OCaml language is highly suitable for fast prototyping and the gradual evolution of software prototypes into final applications, as advocated by the popular "extreme programming" methodology.

4.5. Teaching programming

Our work on the Caml language family has an impact on the teaching of programming. Caml Light 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.

5. Highlights of the Year

5.1. Highlights of the Year

Xavier Leroy received the 2016 Royal Society Milner Award "in recognition of his exceptional achievements in computer programming which includes the design and implementation of the OCaml programming language".

Xavier Leroy received one of the two 2016 Van Wijngaarden Awards from Centrum Wiskunde & Informatica (Amsterdam).

Xavier Leroy received the ACM SIGPLAN Most Influential POPL Paper Award for his POPL 2006 paper, *Formal certification of a compiler back-end or: programming a compiler with a proof assistant* [51].

6. New Software and Platforms

6.1. CompCert

Participants: Xavier Leroy [**contact**], Sandrine Blazy [team Celtique], Jacques-Henri Jourdan, Bernhard Schommer [AbsInt GmbH].

The CompCert project investigates the formal verification of realistic compilers usable for critical embedded software. Such verified compilers come with a mathematical, machine-checked proof that the generated executable code behaves exactly as prescribed by the semantics of the source program. By ruling out the possibility of compiler-introduced bugs, verified compilers strengthen the guarantees that can be obtained by applying formal methods to source programs. AbsInt Angewandte Informatik GmbH sells a commercial version of CompCert with long-term maintenance.

- URL: http://compcert.inria.fr/ (academic), http://www.absint.com/compcert/ (commercial).

6.2. Diy

Participants: Luc Maranget [**contact**], Jade Alglave [Microsoft Research, Cambridge].

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 or generic (LISA) assembler 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. Recent versions also take a subset of the C language as input, so as to test and simulate the C11 model. Recent releases ("Seven") provide a new license (Cecill-B), a simplified build process and numerous features, including a simple macro system that connects the C input language and LISA annotations.

- URL: http://diy.inria.fr/

6.3. Menhir

Participants: François Pottier [**contact**], Yann Régis-Gianas [Université Paris Diderot].

Menhir is a LR(1) parser generator for the OCaml programming language. That is, Menhir compiles LR(1) grammar specifications down to OCaml code.

- URL: http://gallium.inria.fr/~fpottier/menhir/

6.4. OCaml

Participants: Damien Doligez [**contact**], Alain Frisch [LexiFi], Jacques Garrigue [Nagoya University], Fabrice Le Fessant, Xavier Leroy, Luc Maranget, Gabriel Scherer, Mark Shinwell [Jane Street], Leo White [Jane Street], Jeremy Yallop [OCaml Labs, Cambridge University].

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 SPARC), a debugger, a documentation generator, a compilation manager, a package manager, and many libraries contributed by the user community.

- URL: http://ocaml.org/

6.5. OPAM Builder

Participant: Fabrice Le Fessant.

OPAM Builder checks in real time the installability on a computer of all packages after any modification of the OPAM repository. To achieve this result, it uses smart mechanisms to compute incremental differences between package updates, to be able to reuse cached compilations, and go down from quadratic complexity to linear complexity.

- URL: http://github.com/OCamlPro/opam-builder

6.6. PASL

Participants: Michael Rainey [**contact**], Arthur Charguéraud, Umut Acar.

PASL is a C++ library for writing parallel programs targeting the broadly available multicore computers. The library provides a high level interface and can still guarantee very good efficiency and performance, primarily due to its scheduling and automatic granularity control mechanisms.

- URL: http://deepsea.inria.fr/pasl/

6.7. TLAPS

Participants: Damien Doligez [**contact**], Stefan Merz [team Veridis], Martin Riener [team Veridis].

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 independent proof steps. 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+ 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.

- URL: https://tla.msr-inria.inria.fr/tlaps/content/Home.html

6.8. Zenon

Participants: Damien Doligez [**contact**], Guillaume Bury [CNAM], David Delahaye [CNAM], Pierre Halmagrand [team Deducteam], Olivier Hermant [MINES ParisTech].

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 easily retargeted to output scripts for different frameworks (for example, Isabelle and Dedukti).

- URL: http://zenon-prover.org/

7. New Results

7.1. Formal verification of compilers and static analyzers

7.1.1. The CompCert formally-verified compiler

Participants: Xavier Leroy, Bernhard Schommer [AbsInt GmbH], Jacques-Henri Jourdan.

In the context of our work on compiler verification (§3.3.1), since 2005 we have been developing and formally verifying a moderately-optimizing compiler for a large subset of the C programming language, generating assembly code for the PowerPC, ARM, and x86 architectures [7]. This compiler comprises a back-end, which translates the Cminor intermediate language to PowerPC assembly, and is reusable for source languages other than C [6]; 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, out of which Coq's extraction facility generates executable OCaml code. The compiler comes with a 50000-line, machine-checked Coq proof of semantic preservation, establishing that the generated assembly code executes exactly as prescribed by the semantics of the source C program.

This year, the CompCert C compiler was improved in several directions:

- The proof of semantic preservation was extended to account for separate compilation and linking. (See section 7.1.2.)
- Support for 64-bit target processors was added, while keeping the original support for 32-bit processors. The x86 code generator, initially 32-bit only, was extended to handle x86 64-bit as well.
- The generation of DWARF debugging information in `-g` mode, developed last year for PowerPC, is now available for ARM and x86 as well.
- The semantics of conversions from pointer types to the `_Bool` type is fully defined again. (It was made temporarily undefined while addressing issues with comparisons between the null pointer and out-of-bound pointers.)
- More features of ISO C 2011 are supported, such as the `_Noreturn` attribute, or anonymous members of struct and union types.
- As a result of his research on implementing a correct parser for the C language (§7.1.5), Jacques-Henri Jourdan improved the implementation of the parser.

Version 2.7 of CompCert was released in June 2016, incorporating most of these enhancements, with the exception of 64-bit processor support and anonymous members, which will be released Q1 2017.

7.1.2. Separate compilation and linking in CompCert

Participants: Xavier Leroy, Chung-Kil Hur [KAIST, Seoul], Jeehoon Kang [KAIST, Seoul].

Separate compilation (of multiple C source files into multiple object files, followed by linking of the object files to produce the final executable program) has been supported for a long time by the CompCert implementation, but it was not accounted for by CompCert's correctness proof. That proof established semantic preservation in the case of a single, monolithic C source file which is compiled at once to produce the final executable, but not in the more general case of separate compilation and linking.

Version 2.7 of CompCert, released this year, extends the proof of semantic preservation in order to account for separate compilation and linking. It follows the approach described by Kang, Kim, Hur, Dreyer and Vafeiadis in their POPL 2016 paper [47] and prototyped by Kang on CompCert 2.4. In this approach, the proof considers a set of C compilation units, separately compiled to assembly then linked, and shows that the resulting assembly program preserves the semantics of the C program that would be obtained by syntactic linking of the source C compilation units. The simplicity of this approach follows from the fact that semantic preservation is still shown between whole programs (after linking); there is no need to give semantics to individual compilation units. Xavier Leroy integrated the approach of Kang *et al.* into the CompCert development, and extended it to several new optimization passes that were not present in Kang's prototype implementation.

7.1.3. Separation logic assertions for compiler verification

Participants: Xavier Leroy, Timothy Bourke [EPI Parkas], Lélio Brun [EPI Parkas], Maxime Dénès [EPI Marelle].

Separation logic is a powerful tool to reason about imperative programs. It is a Hoare-style program logic where preconditions and postconditions are assertions about the contents of mutable state. Those assertions are built in a compositional manner using a separating conjunction operator.

While effective to prove the correctness of a given program, separation logic and program logics in general are less effective to prove the correctness of a compiler or of a program transformation, in particular because it is difficult to show preservation of termination. The alternative approach that we investigated this year consists in using the assertion language of separation logic, and in particular its separating conjunction, in the context of a conventional, CompCert-style proof of semantic preservation based on simulation diagrams. Assertions from separation logic make it possible to state the invariant that relates the memory states of the program before and after the transformation in a compositional manner, simplifying the proof that this invariant is preserved through execution steps.

This approach was developed and experimentally evaluated in in three case studies.

The first case study was part of project CEEC and consisted in verifying a code generator from a domain-specific, purely-functional intermediate language down to the Clight language of CompCert. Xavier Leroy and Maxime Dénès used ad-hoc separation logic assertions to describe the memory states of the generated Clight programs, and in particular the use of pointers to return multiple function results via "out" parameters.

The second case study was a complete rewrite of the Stacking pass of the CompCert back-end and of its correctness proof, as part of the new support for 64-bit architectures (§7.1.2). For this new proof, Xavier Leroy reused and improved the separation logic assertions of the previous project, using a shallow embedding into Coq instead of a deep embedding. Separating conjunctions are used to specify the layout and current contents of the stack frames for every compiled function, in a way that accommodates 32- and 64-bit registers and pointer values equally well.

The third use takes place in the context of the verified Lustre-to-C compiler in development at team Parkas (see their activity report). The final pass of this compiler translates a simple object-oriented intermediate language, Obc, to CompCert's Clight. Timothy Bourke and Lélio Brun used the separation logic assertions from the second project to specify and reason about the Clight memory layout of the Obc nested objects. Timothy Bourke and Xavier Leroy also extended the separation logic with a "magic wand" operator. A paper on this compiler verification project is under review.

7.1.4. Formal verification of static analyzers based on abstract interpretation

Participants: Jacques-Henri Jourdan, Xavier Leroy, Sandrine Blazy [team Celtique], David Pichardie [team Celtique], Sylvain Boulmé [Grenoble INP, VERIMAG], Alexis Fouilhé [Université Joseph Fourier de Grenoble, VERIMAG], Michaël Périn [Université Joseph Fourier de Grenoble, VERIMAG].

In the context of the Verasco ANR project, we are investigating the formal specification and verification in Coq of a realistic static analyzer based on abstract interpretation. This static analyzer handles a large subset of the C language (the same subset as the CompCert compiler, minus recursion and dynamic allocation); supports a combination of abstract domains, including relational domains; and should produce usable alarms. The long-term goal is to obtain a static analyzer that can be used to prove safety properties of real-world embedded C code.

This year, Jacques-Henri Jourdan published in his PhD thesis [11] an in-depth description of the mode of operation of the current version of the Verasco static analyzer. He also presented at the NSAD workshop [24] the new algorithms used in Verasco for the abstract domain of Octagons that he developed in 2015.

7.1.5. Correct parsing of C using LR(1)

Participants: Jacques-Henri Jourdan, François Pottier.

The C programming language cannot be parsed directly using LR technology. Indeed, the grammar described in the C standard exhibits ambiguities which are addressed in English prose. On the implementation side, it is known from the folklore that one can in fact use an LALR(1) parser to parse C, provided one sets up a so-called "lexer hack" to perform on-the-fly disambiguation of tokens, guided by the current state of the parser.

However, Jacques-Henri Jourdan and François Pottier found that a correct implementation of the "lexer hack" is, surprisingly, difficult. To clarify this situation, they implemented a reference C11 parser using Menhir. They invented new techniques that improve and simplify the "lexer hack", so as to write correct yet reasonably simple C11 parsers. They also created a test suite of C programs that exhibit particularly challenging corner cases. This work is described in a paper that is currently under review.

7.1.6. A SPARK front-end for CompCert

Participants: Pierre Courtieu, Zhi Zang [Kansas University].

SPARK is a language, and a platform, dedicated to developing and verifying critical software. It is a subset of the Ada language. It shares with Ada a strict typing discipline and gives strict guarantees in terms of safety. SPARK goes one step further by disallowing certain "dangerous" features, that is, those that are too difficult to statically analyze (aliasing, references, etc). Given its dedication to safety critical software, we think that the SPARK platform can benefit from a certified compiler. We are working on adding a SPARK front-end to the CompCert verified compiler.

Defining a semantics for SPARK in Coq is previous joint work with Zhi Zang. The current front-end is based on this semantics. The compiler has been written and tested and the proofs of correctness are nearing completion.

7.2. Language design and type systems

7.2.1. Types with unique inhabitants for code inference

Participants: Gabriel Scherer [Northeastern University], Didier Rémy.

Some programming language features (coercions, type-classes, implicits) rely on inferring a part of the code that is determined by its usage context. In order to better understand the theoretical underpinnings of this mechanism, we ask: when is it the case that there is a unique program that could have been guessed, or in other words, that all possible guesses result in equivalent program fragments? Which types have a unique inhabitant?

To approach the question of uniqueness, we build on work in proof theory on canonical representations of proofs. Using the proofs-as-programs correspondence, we adapt the logical technique of focusing to obtain canonical program representations.

In the setting of simply-typed lambda-calculus with sums, equipped with the strong $\beta\eta$-equivalence, we show that uniqueness is decidable. We present a saturating focused logic that introduces irreducible cuts on positive types "as soon as possible". Goal-directed proof search in this logic gives an effective algorithm that returns either zero, one or two distinct inhabitants for any given type.

This work, which was previously presented at a conference [56] and was the main part of Scherer's PhD dissertation [12], has been submitted for journal publication.

7.2.2. Refactoring with ornaments in ML

Participants: Thomas Williams, Didier Rémy.

Thomas Williams and Didier Rémy continued working on ornaments for program refactoring and program transformation in ML. Ornaments have been introduced as a way to describe some changes in data type definitions that preserve their recursive structure, reorganizing, adding, or dropping some pieces of data. After a new data structure has been described as an ornament of an older one, some functions operating on the bare structure can be partially or sometimes totally lifted into functions operating on the ornamented structure.

We have continued working on the decomposition of the algorithm in several steps. Using ornament inference, we first elaborate an ML program into a generic program, which can be seen as a template for all possible liftings of the original program. The generic program is defined in a superset of ML. It can then be instantiated with specific ornaments, and simplified back into an ML program. We studied the semantics of this intermediate language and used them to prove the correctness of the lifting, using logical relations techniques. A paper describing this process was submitted to PLDI.

On the practical side, we updated our prototype implementation to match our theoretical presentation: we create the generic program, then instantiate it. We then simplify the resulting term so that it remains readable to the programmer, and output an ML program. In the case of refactoring (the representation of a data type is modified without adding any data), the transformation is still fully automatic.

7.3. Shared-memory parallelism

7.3.1. Weak memory models

Participants: Luc Maranget, Jade Alglave [University College London–Microsoft Research, UK], Patrick Cousot [New York University], Andrea Parri [Sant'Anna School of Advanced Studies, Pisa, Italy].

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 took 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 focus on devices.

More precisely, in 2016, Luc Maranget pursued his collaboration with Jade Alglave and Patrick Cousot to extend "Cats", a domain-specific language for defining and executing weak memory models. Last year, a long article that presents a precise semantics for "Cats" and a study and formalisation of the HSA memory model was submitted. (The Heterogeneous System Architecture foundation is an industry standards body targeting heterogeneous computing devices.) As this article was rejected, a new paper, focused on the "Cats" semantics, was submitted this year, while the definition of the HSA memory model was made available on the web site of the HSA foundation (http://www.hsafoundation.com/standards/).

This year, our team hosted Andrea Parri, a Ph.D. student (supervised by Mauro Marinoni at Sant'Anna School of Advanced Studies, Pisa, Italy), for six months. Luc Maranget and Andrea Parri collaborated with Paul McKenney (IBM), Alan Stern (Harvard University) and Jade Alglave on the definition of a memory model for the Linux kernel. A preliminary version of this work was presented by Paul McKenney at the *2016 Linux Conference Europe*. While invited at the Dagstuhl seminar "*Concurrency with Weak Memory Models...*", Luc Maranget demonstrated the Diy toolsuite and the "Cats" language. It is worth noting that Cats models are being used independently of us by other researchers, most notably by Yatin Manerkar and Caroline J. Trippel (Princeton University) who discovered an anomaly in the published compilation scheme of the C11 language down to the Power architecture.

Luc Maranget also co-authored a paper that will be presented at POPL 2017 [23]. This work describes memory-model-aware "mixed-size" semantics for the ARMv8 architecture and for the C11 and Sequential Consistency models. A mixed-size semantics accounts for the behaviour of systems that access memory at different granularity levels (bytes, words, etc.) This is joint work with many researchers, including Shaked Flur and other members of Peter Sewell's team (University of Cambridge) as well as Mark Batty (University of Kent).

7.3.2. Algorithms and data structures for parallel computing

Participants: Umut Acar, Vitalii Aksenov, Arthur Charguéraud, Adrien Guatto, Michael Rainey, Filip Sieczkowski.

The ERC Deepsea project, with principal investigator Umut Acar, started in June 2013 and is hosted by the Gallium team. This project aims at developing techniques for parallel and self-adjusting computation in the context of shared-memory multiprocessors (i.e., multicore platforms). The project is continuing work that began at Max Planck Institute for Software Systems between 2010 and 2013. As part of this project, we are developing a C++ library, called PASL, for programming parallel computations at a high level of abstraction. We use this library to evaluate new algorithms and data structures. We obtained four main results this year.

Our first result is a calculus for parallel computing on hardware shared-memory computers such as modern multicores. Many languages for writing parallel programs have been developed. These languages offer several distinct abstractions for parallelism, such as fork-join, async-finish, futures, etc. While they may seem similar, these abstractions lead to different semantics, language design and implementation decisions. In this project, we consider the question of whether it would be possible to unify these approaches to parallelism. To this end, we propose a calculus, called the *DAG-calculus*, which can encode existing approaches to parallelism based on fork-join, async-finish, and futures, and possibly others. We have shown that the approach is realistic by presenting an implementation in C++ and by performing an empirical evaluation. This work was presented at ICFP 2016 [18].

Our second result is a concurrent data structure that may be used to efficiently determine when a concurrently-updated counter reaches the value zero. Our data structure extends an existing data structure called SNZI [44]. While the latter imposes a fixed number of threads, our structure is able to dynamically grow in response to the increasing degree of concurrency in the system. We use our dynamic non-zero indicator data structure to derive an efficient runtime representation of async/finish programs. The async/finish paradigm for expressing parallelism is one that, in the past decade, has become a part of many research-language implementations (e.g. X10) and is now gaining traction in a number of mainstream languages, most notably Java. The implementation of async/finish is challenging because the finish-block mechanism permits, and even encourages, computations in which a large number of threads are required to synchronize on shared barriers, and this number is not statically known. We present an implementation of async/finish and prove that, in a model that takes contention into account, the cost of synchronization of the async-ed threads is amortized constant time, regardless of the number of threads. We also present experimental evaluation suggesting that the approach performs well in practice. This work has been accepted for publication at PPoPP [17].

Our third result is an extended, polished presentation of our prior work on granularity control for parallel algorithms using user-provided complexity functions. Granularity control denotes the problem of controlling the size of parallel threads created in implicitly parallel programs. If small threads are executed in parallel, the overheads due to thread creation can overwhelm the benefits of parallelism. If large threads are executed sequentially, processors may spin idle. In our work, we show that, if we have an oracle able to approximately predict the execution time of every sub-task, then there exists a strategy that delivers provably good performance. Moreover, we present empirical results showing that, for simple recursive divide-and-conquer programs, we are able to implement such an oracle simply by requiring the user to annotate functions with their asymptotic complexity. The idea is to estimate the constant factors that apply by conducting measures at runtime. This work is described in depth in an article published in the Journal of Functional Programming (JFP) [13].

Our fourth result is an extension of our aforementioned granularity control approach, with three major additions. First, we have developed an algorithm that ensures convergence of the estimators associated with the constant factors for all fork-join programs, and not just for a small class of programs. Second, we have built a theoretical analysis establishing bounds for the overall overheads of the convergence phase. Third, we have developed a C++ implementation accompanied with an extensive experimental study covering several benchmarks from the Problem Based Benchmark Suite (PBBS), a collection of high-quality parallel algorithms that delivers state-of-the-art performance. Even though our approach does not leverage a specific compiler and does not require any magic constant to be hard-coded in the source programs, our code either matches or exceeds the performance of the authors' original, hand-tuned codes. An article describing this work is in preparation.

7.4. The OCaml language and system

7.4.1. OCaml

Participants: Damien Doligez, Alain Frisch [Lexifi SAS], Jacques Garrigue [University of Nagoya], Sébastien Hinderer, Fabrice Le Fessant, Xavier Leroy, Luc Maranget, Gabriel Scherer, Mark Shinwell [Jane Street], Leo White [Jane Street], Jeremy Yallop [OCaml Labs, Cambridge University].

This year, we released versions 4.03.0 and 4.04.0 of the OCaml system. These are major releases that introduce a large number of new features. The most important features are:

- A new optimization subsystem called *flambda*, which does inlining and specialization of functions as well as static allocation of some data structures, etc.
- *ephemerons*: a generalization of weak pointers that is better suited for memoization of mutually-recursive functions.
- A fine-grained memory profiler to help programmers understand the allocation behavior of their programs.
- *unboxed types*: a user-controlled optimized representation for some simple data types.

7.4.2. Infrastructure for OCaml

Participant: Sébastien Hinderer.

Sébastien Hinderer worked on improving the test infrastructure of the OCaml compiler. These tests aim at verifying that the compiler works as expected. Currently, they are driven by a set of Makefiles which are hard to maintain and extend and make it difficult to add new tests. Sébastien developed the `ocamltest` driver, which parses test descriptions written in a domain-specific language and runs the appropriate tests.

Sébastien Hinderer also worked on merging the Makefiles used for building the compiler under Unix and Windows. The existence of separate sets of Makefiles, which is the result of a long development history, makes it especially hard to maintain and extend the compiler's build system. Sébastien worked on eliminating this redundancy, so that a single build system can be used on every platform. This is a prerequisite for using the GNU `autoconf` tools and for building easy-to-use cross-compilers for OCaml. A cross-compiler is required, for instance, to build iOS apps using OCaml.

7.4.3. Continuous integration of OCaml packages

Participant: Fabrice Le Fessant.

OPAM is a repository of OCaml source packages. It is now advertised as the official way of installing the OCaml distribution. To maintain a high level of quality for the thousands of source packages distributed in the repository, it is crucial to provide feedback to the developers on the impact of their modifications to the repository, in real-time, despite the high churn and the cascading costs of package recompilations.

We have designed and prototyped a simple modular architecture for a service that monitors the OPAM repository, and triggers recompilation of packages that are impacted by the latest modifications to the repository, for all major and minor OCaml versions since 3.12.1. Previous attempts to design such a system have failed to scale, although they targeted cloud systems of thousands of virtual machines. On the contrary, the new prototype has been deployed on a single quadcore server, and has been able to follow the OPAM repository for eight months, providing feedback in almost real-time. To achieve such a result, it uses many optimizations and caching techniques, to make recompilations as incremental as possible [37].

7.4.4. Global analyses of OCaml programs

Participants: Thomas Blanc [ENSTA-ParisTech & OCamlPro], Pierre Chambart [OCamlPro], Vincent Laviron [OCamlPro], Fabrice Le Fessant, Michel Mauny.

Exception handling in OCaml can be used for managing and reporting errors, as well as to express complex control flow constructs. As such, exceptions can be the source of errors, when, for instance, a function that may raise an exception is called in a context where this exception cannot be handled. In such situations, the program may fail unexpectedly, and the source of the error can be difficult to identify.

This work aims at performing global static analyses of OCaml programs using abstract interpretation techniques, with a particular focus on the detection of uncaught exceptions. Starting from one of the OCaml intermediate languages, we produce a hypergraph that represents the program to be analyzed. Each node of this hypergraph is a program state and each edge is an operation. Operations that may or may not raise an exception (such as function calls) have one or two successors. A fixpoint iteration is then performed on the graph, where function application edges are dynamically replaced by the corresponding subgraphs. In essence, environment information is propagated through the graph, adding at each node a superset of all possible values of each variable, until no additional information can be found. A description of the framework was presented at the 2015 OCaml workshop. We expect concrete results as well as Thomas Blanc's thesis manuscript during 2017.

7.4.5. Type-checking the OCaml intermediate languages

Participants: Pierrick Couderc [ENSTA-ParisTech & OCamlPro], Grégoire Henry [OCamlPro], Fabrice Le fessant, Michel Mauny.

This work aims at propagating type information through the intermediate languages used by the OCaml compiler. We started by the design and implementation of a consistency checker of the type-annotated abstract syntax trees (TASTs) produced by the OCaml compiler. It appears that, when presented as inference rules, the different cases of this TAST checker can be read as the rules of the OCaml type system. Proving the correctness of (part of) the checker would prove the soundness of the corresponding part of the OCaml type system. A preliminary report on this work has been presented at the 17th Symposium on Trends in Functional Programming (TFP 2016).

7.4.6. Optimizing OCaml for satisfiability problems

Participants: Sylvain Conchon [LRI, Univ. Paris Sud], Albin Coquereau [ENSTA-ParisTech], Fabrice Le fessant, Michel Mauny.

This work aims at improving the performance of the Alt-Ergo SMT solver, implemented in OCaml. For safety reasons, the implementation of Alt-Ergo uses as much as possible a functional programming style and persistent data structures, which are sometimes less efficient that the imperative style and mutable data structures. We would like to first obtain a better understanding of the OCaml memory and cache behavior, so as to understand where efficiency could be gained, and then design dedicated data structures (for instance, semi-persistent data structures) and compare their efficiency to the current ones. This work is still at a preliminary stage: we have selected benchmarks and profiled their execution in order to discover sources of inefficiency.

7.4.7. Type compatibility checking for dynamically loaded OCaml data

Participants: Florent Balestrieri [ENSTA-ParisTech], Michel Mauny.

The SecurOCaml project (FUI 18) aims at enhancing the OCaml language and environment in order to make it more suitable for building secure applications, following recommendations published by the French ANSSI in 2013. Michel Mauny and Florent Balistrieri (ENSTA-ParisTech) represent ENSTA-Paristech in this project for the two-year period 2016-2017.

The goal of this first year was to design and produce an effective OCaml implementation that checks whether a memory graph – typically the result obtained by un-marshalling some data – is compatible with a given OCaml type, following the algorithm designed by Henry *et al.* in 2012. As the algorithm needs a runtime representation of OCaml types, Florent Balestrieri implemented a library for generic programming in OCaml [21]. He also implemented a type-checker which, when given a type and a memory graph, checks whether the former could be the type of the latter. The algorithm handles sharing and polymorphism, but currently supports neither functional values nor existential types.

7.4.8. Pattern matching

Participants: Luc Maranget, Gabriel Scherer [Northeastern University, Boston], Thomas Réfis [Jane Street LLC].

A new pattern matching diagnostic message, which should help OCaml programmers to detect rare but vicious programming errors, was integrated in the yearly release of the OCaml compiler, and was presented at the OCaml Users and Developers Workshop [39].

7.4.9. Error diagnosis in Menhir parsers

Participant: François Pottier.

In 2015, François Pottier proposed a reachability algorithm for LR automata, which he implemented in the Menhir parser generator. He applied this approach to the C grammar in the front-end of the CompCert compiler, therefore allowing CompCert to produce better syntax error messages. This work has been presented at the conferences JFLA 2016 [31] and CC 2016 [26].

7.5. Software specification and verification

7.5.1. Step-indexing in program logics

Participant: Filip Sieczkowski.

Filip Sieczkowski pursued a line of work focused on techniques for formal reasoning about programs, in joint work with Lars Birkedal (Aarhus University) and Kasper Svendsen (Cambridge University). A modern and successful approach to grounding programs logics is to rely on so-called step-indexed models. Filip and his co-authors solved a problem that arises in most step-indexed models, due to a tight coupling between the unfoldings of a recursive domain equation and evaluation steps. Their approach is based on the use of transfinite step-indexing. This work appeared at ESOP 2016 [29].

7.5.2. TLA+

Participants: Damien Doligez, Leslie Lamport [Microsoft Research], Martin Riener [team VeriDis], Stephan Merz [team VeriDis].

Damien Doligez is 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 [48], and to build tools for writing TLA+ specifications and mechanically checking the proofs.

Our rewrite of the TLAPS tools is almost done and we hope to do a first release in the first quarter of 2017.

7.5.3. Hash tables and iterators: a case study in program verification

Participant: François Pottier.

In the setting of the Vocal ANR project, François Pottier developed the the specification and proof of an (imperative, sequential) hash table implementation, as found in the module `Hashtbl` of OCaml's standard library. This data structure supports the usual dictionary operations (insertion, lookup, and so on), as well as iteration via folds and iterators. The code was verified using higher-order separation logic, embedded in Coq, via Charguéraud's CFML tool and library. This work was presented at CPP 2017 [27]. It can be viewed as a case study that should help prepare the way for verifying other modules in the Vocal library.

7.5.4. Read-only permissions in separation logic

Participants: Arthur Charguéraud, François Pottier.

Separation Logic, as currently implemented in Charguéraud's CFML tool and library, imposes a simple ownership discipline on mutable heap-allocated data structures: a thread either has full read-write access to a data structure, or has no access at all. This implies, for instance, that two threads cannot temporarily share read-only access to a data structure. There exist more flexible disciplines in the literature, such as "fractional permissions" and "share algebras", but they are much more complex.

In the setting of the Vocal ANR project, Arthur Charguéraud and François Pottier noted that it would be desirable to define an extension of Separation Logic that allows temporary shared read-only access, yet remains very simple. They proposed a general mechanism for temporarily converting any assertion (or "permission") to a read-only form. The metatheory of this proposal has been verified in Coq. This work will be presented at ESOP 2017 [42].

Charguéraud and Pottier believe that this mechanism should allow more concise specifications and proofs. This remains to be confirmed, in future work, via an implementation in CFML and case studies in the Vocal project.

7.5.5. Formal reasoning about asymptotic complexity

Participants: Armaël Guéneau, Arthur Charguéraud, François Pottier.

Armaël Guéneau started his Ph.D. at Gallium in September 2016, supervised by Arthur Charguéraud and François Pottier. In the line of his previous M2 internship at Gallium, he continued his work on asymptotic reasoning in Coq. The challenge is to give a formal definition of the well-known big-O notation, covering both single-variable and multiple-variable scenarios, to establish its fundamental properties, and to define tactics that make asymptotic reasoning as convenient in Coq as it seemingly is on paper. The ultimate goal is to apply these techniques to machine-checked proofs of the asymptotic time complexity of programs.

7.5.6. Certified distributed algorithms for autonomous mobile robots

Participant: Pierre Courtieu.

The variety and complexity of the tasks that can be performed by autonomous robots are increasing. Many applications envision groups of mobile robots that self-organise and cooperate toward the resolution of common objectives, in the absence of any central coordinating authority.

Pierre Courtieu is elaborating a verification platform, based on Coq, for distributed algorithms for autonomous robots. (This is joint work with Xavier Urbain, Sebastien Tixeuil and Lionel Rieg.) As part of this effort, Pierre Courtieu designed and verified a protocol for mobile robots that achieves the "gathering" task in all cases where it has not been proved impossible [34], [35].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. The Caml Consortium

Participants: Xavier Leroy [**contact**], Damien Doligez, Didier Rémy.

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 Caml. 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.

The Consortium currently has 14 member companies:

- Aesthetic Integration
- Ahrefs
- Bloomberg
- CEA
- Citrix
- Dassault Aviation
- Esterel Technologies
- Facebook
- Jane Street
- Kernelyze
- LexiFi
- Microsoft
- OCamlPro
- SimCorp

For a complete description of this structure, refer to http://caml.inria.fr/consortium/. Xavier Leroy chairs the scientific committee of the Consortium.

8.1.2. Scientific Advisory for OCamlPro

Participant: Fabrice Le Fessant.

OCamlPro is a startup company founded in 2011 by Fabrice Le Fessant to promote the use of OCaml in the industry, by providing support, services and tools for OCaml to software companies. OCamlPro performs a lot of research and development, in close partnership with academic institutions such as IRILL, Inria and Univ. Paris Sud, and is involved in several collaborative projects with Gallium, such as the Bware ANR, the Vocal ANR and the Secur-OCaml FUI.

Since 2011, Fabrice Le Fessant is a scientific advisor at OCamlPro, as part of a collaboration contract for Inria, to transfer his knowledge on the internals of the OCaml runtime and the OCaml compilers.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR projects

9.1.1.1. BWare

Participants: Damien Doligez, Fabrice Le Fessant.

The "BWare" project (2012–2016) is coordinated by David Delahaye at Conservatoire National des Arts et Métiers and funded by the *Ingénierie Numérique et Sécurité* programme of *Agence Nationale de la Recherche*. BWare is an industrial research project that aims to provide a mechanized framework to support the automated verification of proof obligations coming from the development of industrial applications using the B method and requiring high guarantees of confidence.

9.1.1.2. Verasco

Participants: Jacques-Henri Jourdan, Xavier Leroy.

The "Verasco" project (2012–2016) is coordinated by Xavier Leroy and funded by the *Ingéniérie Numérique et Sécurité* programme of *Agence Nationale de la Recherche*. The objective of this 4.5-year project is to develop and formally verify a static analyzer based on abstract interpretation, and interface it with the CompCert C verified compiler.

9.1.1.3. Vocal

Participants: 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".

The library will be made available to all OCaml programmers and will be of particular interest to implementors of safety-critical OCaml programs, such as Coq, Astrée, Frama-C, CompCert, Alt-Ergo, as well as new projects. By offering verified program components, our work will provide the essential building blocks that are needed to significantly decrease the cost of developing new formally verified programs.

9.1.2. FSN projects

9.1.2.1. ADN4SE

Participants: Damien Doligez, Martin Riener.

The "ADN4SE" project (2012–2016) is coordinated by the Sherpa Engineering company and funded by the *Briques Génériques du Logiciel Embarqué* programme of *Fonds national pour la Société Numérique*. The aim of this project is to develop a process and a set of tools to support the rapid development of embedded software with strong safety constraints. Gallium is involved in this project to provide tools and help for the formal verification in TLA+ of some important aspects of the PharOS real-time kernel, on which the whole project is based.

9.1.3. FUI Projects

9.1.3.1. Secur-OCaml

Participants: Damien Doligez, Fabrice Le Fessant.

The "Secur-OCaml" project (2015–2018) is coordinated by the OCamlPro company, with a consortium focusing on the use of OCaml in security-critical contexts, while OCaml is currently mostly used in safety-critical contexts. Gallium is invoved in this project to integrate security features in the OCaml language, to build a new independant interpreter for the language, and to update the recommendations for developers issued by the former LaFoSec project of ANSSI.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. Deepsea

Participants: Umut Acar, Vitalii Aksenov, Arthur Charguéraud, Michael Rainey, Filip Sieczkowski.

The Deepsea project (2013–2018) is coordinated by Umut Acar and funded by FP7 as an ERC Starting Grant. Its objective is to develop abstractions, algorithms and languages for parallelism and dynamic parallelism, with applications to problems on large data sets.

9.2.2. ITEA3 Projects

9.2.2.1. Assume

Participants: Xavier Leroy, Luc Maranget.

ASSUME (2015–2018) is an ITEA3 project involving France, Germany, Netherlands, Turkey and Sweden. The French participants are coordinated by Jean Souyris (Airbus) and include Airbus, Kalray, Sagem, ENS Paris, and Inria Paris. The goal of the project is to investigate the usability of multicore and manycore processors for critical embedded systems. Our involvement in this project focuses on the formalisation and verification of memory models and of automatic code generators from reactive languages.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

- Princeton University: interactions between the CompCert verified C compiler and the Verified Software Toolchain developed at Princeton.
- Cambridge University and Microsoft Research Cambridge: formal modeling and testing of weak memory models.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

Michel Mauny is a member of the steering committee of the OCaml workshop.

Didier Rémy was a member of the steering committee of the OCaml workshop until September 2017. He is a member of the steering committee of the ML Family workshop.

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

Xavier Leroy was a member of the program committees of the Compiler Construction conference (CC 2016), of the conference on Interactive Theorem Proving (ITP 2016), and on the external review committee of the symposium on Principles of Programming Languages (POPL 2017).

François Pottier was a member of the program committees of the conferences Journées Francophones des Langages Applicatifs (JFLA 2017) and Compiler Construction (CC 2017).

10.1.2.2. Reviewer

In 2016, the members of Gallium reviewed at least 30 conference submissions.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Xavier Leroy is area editor (programming languages) for the Journal of the ACM. He is on the editorial board for the Research Highlights column of Communications of the ACM. He is a member of the editorial board of the Journal of Automated Reasoning.

François Pottier is an editor for the Journal of Functional Programming.

10.1.4. Invited Talks

Xavier Leroy was an invited speaker at the ICALP conference (Rome, July 2016).

10.1.5. Research Administration

Xavier Leroy is *délégué scientifique adjoint* of Inria Paris and appointed member of Inria's *Commission d'Évaluation*. He participated in the following Inria hiring and promotion committees: *jury d'admissibilité DR2*, *promotions CR1*, and *promotions DR1*.

Xavier Leroy was a member of the hiring committee for a professor position at Université de Lorraine.

Xavier Leroy was a member of the HCERES evaluation panel for the LORIA laboratory.

François Pottier is a member of the *Commission de Développement Technologique* and (as of January 2016) chairs the *Comité de Suivi Doctoral* of Inria Paris.

Didier Rémy is *Deputy Scientific Director* (ADS) in change of *Algorithmics, Programming, Software and Architecture*.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Xavier Leroy and Didier Rémy, "Functional programming languages", 15+18h, M2 (MPRI), Université Paris Diderot, France.

Master: Luc Maranget, "Semantics, languages and algorithms for multi-core programming", 13.5h, M2 (MPRI), Université Paris Diderot, France.

Master: "Principles of Programming Languages", 32h, M1, ENSTA-ParisTech, France.

Licence: François Pottier, "Programmation avancée" (INF441), 20h, L3, École Polytechnique, France.

Master: François Pottier, "Compilation" (INF564), 20h, M1, École Polytechnique, France.

Licence: Michael Rainey and Umut Acar, "Theory and practice of parallel computing" (part of a longer course entitled 15-210, "Parallel and Sequential Data Structures and Algorithms"), 9h, L3, Carnegie Mellon University, USA.

Michel Mauny has been a Professor at ENSTA-ParisTech from August 1st, 2005 to July 31st, 2016. While at ENSTA-ParisTech, Michel Mauny was in charge of the specialization "Architecture and Security of Information Systems" (MSc. 2nd year).

François Pottier has been a Professeur Chargé de Cours at École Polytechnique from September 1st, 2004 to August 31st, 2016.

Didier Rémy is Inria's delegate in the pedagogical team of the MPRI.

Fabrice Le Fessant has been involved in the second edition of the OCaml MOOC on the FUN platform, in coordination with the OCamlPro team in charge of the development of the exercise platform [33].

10.2.2. Supervision

M2 (Master Pro): Jacques-Pascal Deplaix, Epitech, supervised by François Pottier.

M2 (MPRI): Ambroise Lafont, École Polytechnique, supervised by Xavier Leroy.

PhD: Pierre Halmagrand, "Automated Deduction and Proof Certification for the B Method" [45], Conservatoire National des Arts et Métiers, defended December 10, 2016, supervised by David Delahaye, Damien Doligez and Olivier Hermant.

PhD: Jacques-Henri Jourdan, "Verasco: a formally verified C static analyzer" [11], Université Paris Diderot, defended May 2016, supervised by Xavier Leroy.

PhD: Gabriel Scherer, "Which types have a unique inhabitant?" [12], Université Paris Diderot, defended March 2016, supervised by Didier Rémy.

PhD in progress: Vitalii Aksenov, "Parallel Dynamic Algorithms", Université Paris Diderot, since September 2015, supervised by Umut Acar (co-advised with Anatoly Shalyto, ITMO University of Saint Petersburg, Russia).

PhD in progress: Thomas Blanc (ENSTA-ParisTech & OCamlPro), "Analyses de programmes complets, application à OCaml", Université Paris-Saclay, since February 2014, supervised by Michel Mauny and Pierre Chambart (OCamlPro).

PhD in progress: Pierrick Couderc (ENSTA-ParisTech & OCamlPro), "Typage modulaire du langage intermédiaire du compilateur OCaml," Université Paris-Saclay, since December 2014, supervised by Michel Mauny, Grégoire Henry (OCamlPro) and Fabrice Le Fessant.

PhD in progress: Albin Coquereau (ENSTA-ParisTech), "Amélioration de performances pour le solveur SMT Alt-Ergo: conception d'outils d'analyse, optimisations et structures de données efficaces pour OCaml," Université Paris-Saclay, since October 2015, supervised by Michel Mauny, Sylvain Conchon (LRI, Université Paris-Sud) and Fabrice Le Fessant.

PhD in progress: Armaël Guéneau, "Towards Machine-Checked Time Complexity Analyses", Université Paris Diderot, since September 2016, supervised by Arthur Charguéraud and François Pottier.

PhD in progress: Thomas Williams, "Putting Ornaments into practice", Université Paris Diderot, since September 2014, supervised by Didier Rémy.

10.2.3. Juries

François Pottier was a reviewer for the Ph.D. thesis of Benoît Vaugon, Université Paris-Saclay, March 2016. He was a reviewer for the Habilitation of Damien Pous, ENS Lyon, September 2016. He was a member of the jury for the Ph.D. thesis of Léon Gondelman, Université Paris-Saclay, December 2016.

Xavier Leroy was on the Ph.D. committee of Pierre Wilke, Université Rennes 1, November 2016.

Didier Rémy was chair of the Ph.D. committee of Raphaël Cauderlier, Conservatoire National des Arts et Métiers (CNAM), October 2016.

10.3. Popularization

Xavier Leroy gave a popularization talk on formal methods at the plenary days of Inria's DGD-T (may 2016) and another on critical avionics software for first-year students at École Polytechnique (june 2016).

11. 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, article no 7, http://dx.doi.org/10.1145/2627752.

[2] 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.

[3] J. CRETIN, D. RÉMY.*System F with Coercion Constraints*, in "CSL-LICS 2014: Computer Science Logic / Logic In Computer Science", ACM, 2014, article no 34, http://dx.doi.org/10.1145/2603088.2603128.

[4] 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.

[5] D. LE BOTLAN, D. RÉMY.*Recasting MLF*, in "Information and Computation", 2009, vol. 207, n^{o} 6, p. 726–785, http://dx.doi.org/10.1016/j.ic.2008.12.006.

[6] 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.

[7] 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.

[8] F. POTTIER.*Hiding local state in direct style: a higher-order anti-frame rule*, in "Proceedings of the 23rd Annual IEEE Symposium on Logic In Computer Science (LICS'08)", IEEE Computer Society Press, June 2008, p. 331-340, http://dx.doi.org/10.1109/LICS.2008.16.

[9] F. POTTIER, J. PROTZENKO.*Programming with permissions in Mezzo*, in "Proceedings of the 18th International Conference on Functional Programming (ICFP 2013)", ACM Press, 2013, p. 173–184, http://dx.doi.org/10.1145/2500365.2500598.

[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] J.-H. JOURDAN.*Verasco: a Formally Verified C Static Analyzer*, Universite Paris Diderot-Paris VII, May 2016, https://hal.archives-ouvertes.fr/tel-01327023.

[12] G. SCHERER.*Which types have a unique inhabitant?: Focusing on pure program equivalence*, Université Paris-Diderot, March 2016, https://hal.inria.fr/tel-01309712.

Articles in International Peer-Reviewed Journal

[13] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY.*Oracle-Guided Scheduling for Controlling Granularity in Implicitly Parallel Languages*, in "Journal of Functional Programming", November 2016, vol. 26 [*DOI :* 10.1017/S0956796816000101], https://hal.inria.fr/hal-01409069.

[14] 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 (TOPLAS)", August 2016, vol. 38, n° 4, 94 [*DOI :* 10.1145/2837022], https://hal.inria.fr/hal-01246534.

[15] M.-K. RIVIERE, J.-H. JOURDAN, S. ZOHAR.*dfcomb: An R-package for phase I/II trials of drug combinations*, in "Computer Methods and Programs in Biomedicine", 2016, vol. 125, p. 117–133 [*DOI :* 10.1016/J.CMPB.2015.10.018], http://hal.upmc.fr/hal-01297367.

[16] M.-K. RIVIERE, Y. YUAN, J.-H. JOURDAN, F. DUBOIS, S. ZOHAR.*Phase I/II dose-finding design for molecularly targeted agent: Plateau determination using adaptive randomization*, in "Statistical Methods in Medical Research", March 2016 [*DOI :* 10.1177/0962280216631763], http://hal.upmc.fr/hal-01298681.

International Conferences with Proceedings

[17] U. A. ACAR, N. BEN-DAVID, M. RAINEY.*Contention in Structured Concurrency: Provably Efficient Dynamic Non-Zero Indicators for Nested Parallelism*, in "22nd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming", Austin, United States, February 2017 [*DOI :* 10.1145/3018743.3018762], https://hal.inria.fr/hal-01416531.

[18] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY, F. SIECZKOWSKI.*Dag-calculus: a calculus for parallel computation*, in "Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming (ICFP)", Nara, Japan, September 2016, p. 18 - 32 [*DOI :* 10.1145/2951913.2951946], https://hal.inria.fr/hal-01409022.

[19] 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)", Unknown, Unknown or Invalid Region, ACM, 2017, p. 515-529, https://hal.archives-ouvertes.fr/hal-01424794.

[20] S. AZAIEZ, D. DOLIGEZ, M. LEMERRE, T. LIBAL, S. MERZ.*Proving Determinacy of the PharOS Real-Time Operating System*, in "Abstract State Machines, Alloy, B, TLA, VDM, and Z - 5th International Conference, ABZ 2016", Linz, Austria, M. J. BUTLER, K.-D. SCHEWE, A. MASHKOOR, M. BIRÓ (editors), LNCS - Lecture Notes in Computer Science, Springer, May 2016, vol. 9675, p. 70-85 [*DOI :* 10.1007/978-3-319-33600-8_4], https://hal.inria.fr/hal-01322335.

[21] F. BALESTRIERI, M. MAUNY.*Generic Programming in OCaml*, in "OCaml 2016 - The OCaml Users and Developers Workshop", Nara, Japan, September 2016, https://hal.inria.fr/hal-01413061.

[22] S. FLUR, K. E. GRAY, C. PULTE, S. SARKAR, A. SEZGIN, L. MARANGET, W. DEACON, P. SEWELL.*Modelling the ARMv8 Architecture, Operationally: Concurrency and ISA*, in "Principles of Programming Languages 2016 (POPL 2016)", Saint Petersburg, United States, January 2016, https://hal.inria.fr/hal-01244776.

[23] S. FLUR, S. SARKAR, C. PULTE, K. NIENHUIS, L. MARANGET, K. E. GRAY, A. SEZGIN, M. BATTY, P. SEWELL.*Mixed-size Concurrency: ARM, POWER, C/C++11, and SC*, in "44th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2017)", Paris, France, ACM, January 2017, https://hal.inria.fr/hal-01413221.

[24] J.-H. JOURDAN.*Sparsity Preserving Algorithms for Octagons*, in "NSAD 2016 - Numerical and symbolic abstract domains workshop", Edinburgh, United Kingdom, I. MASTROENI (editor), Elsevier, September 2016, 14, https://hal.inria.fr/hal-01406795.

[25] D. KÄSTNER, X. LEROY, S. BLAZY, B. SCHOMMER, M. SCHMIDT, C. FERDINAND.*Closing the Gap – The Formally Verified Optimizing Compiler CompCert*, in "SSS'17: Safety-critical Systems Symposium 2017", Bristol, United Kingdom, Proceedings of the Twenty-fifth Safety-Critical Systems Symposium, February 2017, https://hal.inria.fr/hal-01399482.

[26] F. POTTIER.*Reachability and Error Diagnosis in LR(1) Parsers*, in "CC 2016 - 25th International Conference on Compiler Construction", Barcelone, Spain, Proceedings of the 25th International Conference on Compiler Construction (CC 2016), March 2016, 11 [*DOI :* 10.1145/2892208.2892224], https://hal.inria.fr/hal-01417004.

[27] F. POTTIER.*Verifying a Hash Table and Its Iterators in Higher-Order Separation Logic*, in "Certified Programs and Proofs", Paris, France, Proceedings of the 6th ACM SIGPLAN Conference on Certified Programs and Proofs (CPP 2017), January 2017, https://hal.inria.fr/hal-01417102.

[28] R. A. RAGHUNATHAN, S. A. MULLER, U. A. ACAR, G. A. BLELLOCH.*Hierarchical Memory Management for Parallel Programs*, in "Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming", Nara, Japan, September 2016 [*DOI :* 10.1145/3022670.2951935], https://hal.inria.fr/hal-01416237.

[29] K. SVENDSEN, F. SIECZKOWSKI, L. BIRKEDAL.*Transfinite Step-Indexing: Decoupling Concrete and Logical Steps*, in "25th European Symposium on Programming Languages and Systems", Eindhoven, Netherlands, December 2016, vol. 9632, p. 727 - 751 [*DOI :* 10.1007/978-3-662-49498-1_28], https://hal.inria.fr/hal-01408649.

[30] B. VAUGON, M. MAUNY.*A Type Inference System Based on Saturation of Subtyping Constraints*, in "Trends in Functional Programming", College Park (MD), United States, June 2016, https://hal.inria.fr/hal-01413043.

National Conferences with Proceeding

[31] F. POTTIER.*Reachability and error diagnosis in LR(1) automata*, in "Journées Francophones des Langages Applicatifs", Saint-Malo, France, January 2016, https://hal.inria.fr/hal-01248101.

Conferences without Proceedings

[32] Ç. BOZMAN, T. HUFFSCHMITT, M. LAPORTE, F. LE FESSANT.*ocp-lint, A Plugin-based Style-Checker with Semantic Patches*, in "OCaml Users and Developers Workshop 2016", Nara, Japan, September 2016, https://hal.inria.fr/hal-01352013.

[33] B. CANOU, G. HENRY, Ç. BOZMAN, F. LE FESSANT.*Learn OCaml, An Online Learning Center for OCaml*, in "OCaml Users and Developers Workshop 2016", Nara, Japan, September 2016, https://hal.inria.fr/hal-01352015.

[34] P. COURTIEU, L. RIEG, S. TIXEUIL, X. URBAIN.*A Certified Universal Gathering Algorithm for Oblivious Mobile Robots*, in "Distributed Computing (DISC)", Paris, France, September 2016, http://hal.upmc.fr/hal-01349061.

[35] P. Courtieu, L. Rieg, S. Tixeuil, X. Urbain.*Certified Universal Gathering in R2 for Oblivious Mobile Robots*, in "ACM Conference on Principles of Distributed Computing (PODC)", Chicago, United States, ACM, July 2016, http://hal.upmc.fr/hal-01349084.

[36] J.-H. Jourdan.*Statistically profiling memory in OCaml*, in "OCaml 2016", Nara, Japan, September 2016, https://hal.inria.fr/hal-01406809.

[37] F. Le Fessant.*OPAM-builder: Continuous Monitoring of OPAM Repositories*, in "OCaml Users and Developers Workshop 2016", Nara, Japan, September 2016, https://hal.inria.fr/hal-01352008.

[38] X. Leroy, S. Blazy, D. Kästner, B. Schommer, M. Pister, C. Ferdinand.*CompCert - A Formally Verified Optimizing Compiler*, in "ERTS 2016: Embedded Real Time Software and Systems, 8th European Congress", Toulouse, France, SEE, January 2016, https://hal.inria.fr/hal-01238879.

[39] G. Scherer, L. Maranget, T. Réfis.*Ambiguous pattern variables*, in "OCaml 2016: The OCaml Users and Developers Workshop", Nara, Japan, September 2016, 2, https://hal.inria.fr/hal-01413241.

Research Reports

[40] X. Leroy, D. Doligez, A. Frisch, J. Garrigue, D. Rémy, J. Vouillon.*The OCaml system release 4.04: Documentation and user's manual*, Inria, November 2016, https://hal.inria.fr/hal-00930213.

[41] X. Leroy.*The CompCert C verified compiler: Documentation and user's manual: Version 2.7*, Inria, June 2016, https://hal.inria.fr/hal-01091802.

Other Publications

[42] A. Charguéraud, F. Pottier.*Temporary Read-Only Permissions for Separation Logic*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01408657.

References in notes

[43] V. Benzaken, G. Castagna, A. Frisch.*CDuce: an XML-centric general-purpose language*, in "Proceedings of the Eighth ACM SIGPLAN International Conference on Functional Programming", C. Runciman, O. Shivers (editors), ACM, 2003, p. 51–63, https://www.lri.fr/~benzaken/papers/icfp03.ps.

[44] F. Ellen, Y. Lev, V. Luchangco, M. Moir.*SNZI: Scalable NonZero Indicators*, in "Proceedings of the Twenty-sixth Annual ACM Symposium on Principles of Distributed Computing", 2007, p. 13–22, http://dl.acm.org/citation.cfm?id=1281106.

[45] P. Halmagrand.*Automated Deduction and Proof Certification for the B Method*, Conservatoire National des Arts et Métiers, December 2016.

[46] H. Hosoya, B. C. Pierce.*XDuce: A Statically Typed XML Processing Language*, in "ACM Transactions on Internet Technology", 2003, vol. 3, n^o 2, p. 117–148, http://doi.acm.org/10.1145/767193.767195.

[47] J. Kang, Y. Kim, C. Hur, D. Dreyer, V. Vafeiadis.*Lightweight verification of separate compilation*, in "Proceedings of the 43rd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages", 2016, p. 178–190, http://doi.acm.org/10.1145/2837614.2837642.

[48] 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.

[49] X. LEROY, D. DOLIGEZ, J. GARRIGUE, D. RÉMY, J. VOUILLON.*The Objective Caml system, documentation and user's manual – release 4.02*, Inria, August 2014, http://caml.inria.fr/pub/docs/manual-ocaml-4.02/.

[50] X. LEROY.*Java bytecode verification: algorithms and formalizations*, in "Journal of Automated Reasoning", 2003, vol. 30, n^{o} 3–4, p. 235–269, http://dx.doi.org/10.1023/A:1025055424017.

[51] X. LEROY.*Formal certification of a compiler back-end, or: programming a compiler with a proof assistant*, in "33rd ACM symposium on Principles of Programming Languages", ACM Press, 2006, p. 42–54, http://doi.acm.org/10.1145/1111037.1111042.

[52] B. C. PIERCE.*Types and Programming Languages*, MIT Press, 2002.

[53] F. POTTIER.*Simplifying subtyping constraints: a theory*, in "Information and Computation", 2001, vol. 170, n^{o} 2, p. 153–183, http://gallium.inria.fr/~fpottier/publis/fpottier-ic01.ps.gz.

[54] F. POTTIER, V. SIMONET.*Information Flow Inference for ML*, in "ACM Transactions on Programming Languages and Systems", January 2003, vol. 25, n^{o} 1, p. 117–158, http://dx.doi.org/10.1145/596980.596983.

[55] D. RÉMY, J. VOUILLON.*Objective ML: A simple object-oriented extension to ML*, in "24th ACM Conference on Principles of Programming Languages", ACM Press, 1997, p. 40–53, http://gallium.inria.fr/~remy/ftp/objective-ml!popl97.pdf.

[56] G. SCHERER, D. RÉMY.*Which simple types have a unique inhabitant?*, in "ICFP'15: 20th International Conference on Functional Programming", ACM Press, 2015, p. 243–255, http://dx.doi.org/10.1145/2784731.2784757.

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:

1.2. - Networks
1.2.3. - Routing
1.2.9. - Social Networks
1.3. - Distributed Systems
3.5. - Social networks
3.5.1. - Analysis of large graphs
6.1.3. - Discrete Modeling (multi-agent, people centered)
7.1. - Parallel and distributed algorithms
7.2. - Discrete mathematics, combinatorics
7.3. - Optimization
7.9. - Graph theory
7.10. - Network science
7.13. - Quantum algorithms

Other Research Topics and Application Domains:

1.1.8. - Evolutionnary biology
1.1.11. - Systems biology
6.3.2. - Network protocols
6.3.4. - Social Networks
7.2. - Smart travel

1. Members

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, HDR]

Faculty Members

Yacine Boufkhad [Univ. Paris VII, Associate Professor]
Pierre Charbit [Univ. Paris VII, Associate Professor]
Fabien de Montgolfier [Univ. Paris VII, Associate Professor]
Hugues Fauconnier [Univ. Paris VII, Professor, HDR]
Carole Gallet Delporte [Univ. Paris VII, Professor, HDR]
Michel Habib [Univ. Paris VII, Professor, HDR]

PhD Students

Leonardo Linguaglossa [Inria, until Jun 2016, granted by Alcatel-Lucent Bell Labs]
Simon Collet [ERC BDA]
Lucas Boczkowski [ERC BDA]
Laurent Feuilloley [ENS Cachan]

Mengchuan Zou [Inria, from Oct 2016]

Administrative Assistant

Christine Anocq [Inria]

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 Decompositions

We study new decompositions schemes such as 2-join, skew partitions and others partition problems. These graph decompositions appeared in the structural graph theory and are the basis of some well-known theorems such as the Perfect Graph Theorem. For these decompositions there is a lack of efficient algorithms. We aim at designing algorithms working in $O(nm)$ since we think that this could be a lower bound for these decompositions.

3.1.2. 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 do not need extra memory. We already have 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.

3.1.3. 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 community can be viewed as the union of two sub-communities. This is true even in our team. Even though they are not completely disjoint, they are disjoint enough not to leverage each others' 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 complexities, 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 the 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 from 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 research program is ongoing to carry many important notions of Distributed Computing into a *standard* computational complexity.

3.3. Network Algorithms and Analysis

Based on our scientific foundation on 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, where noise can alter information as it propagates or where memory of nodes is limited for example.

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 target is multiple path routing: how to design forwarding tables providing multiple disjoint paths to a 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 and analysis can be performed. We especially target online social networks if 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 forwarding information becomes critical and has recently gained in interest. Some problems that arise in network verification such as loop detection for example, may be naturally encoded as Boolean Satisfiability problems. Beside the theoretical interest of this encoding in complexity proofs, it has also a practical value for solving these problems by taking advantage of the many 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 need to collect data sets from 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 represent 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 answer although fast computation time is not guaranteed for all networks. We already have designed such heuristics for diameter computation; understanding the structural properties that enable fast computation time in practice is still an open question.

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. New Software and Platforms

5.1. big-graph-tools

FUNCTIONAL DESCRIPTION

Gang is developing software for big graph manipulation. A preliminary library offering diameter and skeleton computation is available online [0]. This library was used to compute the diameters of the worldwide road network (200M edges), skeleton subtrees of the shortest-path trees of continental-sized road networks, as well as 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/

[0] https://who.rocq.inria.fr/Laurent.Viennot/dev/big-graph-tools/

6. New Results

6.1. Graph and Combinatorial Algorithms

6.1.1. New Results in Multi-sweep Graph Search

A theoretical model to describe a series of successive graph searches is proposed in [7]. We apply this model to deal with cocomparability graphs (i.e., complement of comparability graphs) in [6] and in [48] or [44]. In this series of papers we provide a general algorithmic framework for many optimization problems on cocomparability graphs, such as Minimum Path Cover, Maximum Independent Set, Maximal interval subgraph, etc.

We also provide a new very simple algorithm for the recognition of cocomparability graphs. This algorithm is also based on a series of successive graph searches in [13].

We mainly use the two well-known Lexicographic graph searches: LBFS and LDFS, but not only. In [48], we also introduced a new graph search LocalMNS which seems to behave nicely on cocomparability graphs.

6.1.2. Studies of Read Networks and Laminar Graphs

In the context of biological networks, in [50] we introduce k-laminar graphs — a new class of graphs which extends the idea of Asteroidal-triple-free graphs. A graph is k-laminar if it admits a diametral path that is k-dominating. This bio-inspired class of graphs was motivated by a biological application dealing with sequence similarity networks of reads. We briefly develop the context of the biological application in which this graph class appeared and then we consider the relationships of this new graph class among known graph classes and then we study its recognition problem. For the recognition of k-laminar graphs, we develop polynomial algorithms when k is fixed. For $k = 1$, our algorithm improves a Deogun and Krastch's algorithm (1999). We finish by an NP-completeness result when k is unbounded.

6.1.3. Further Studies into Shortest Paths, Eccentricity, and Laminarity

From our recent research on diameter computations on graphs we also investigated some reductions between polynomial problems on graphs [3].

We also extend the well-known multisweep BFS to give a better polynomial-time approximation for the Maximum Eccentricity Shortest Path Problem, in relation with the k-Laminarity Problem [20].

6.1.4. Clique Colourings of Perfect Graphs

A *clique-coloring* of a graph G is an assignment of colors to the vertices of G in such a way that no inclusion-wise maximal clique of size at least two of G is monochromatic (as usual, a set of vertices is *monochromatic* if all vertices in the set received the same color). The *clique-chromatic number* of G, denoted by $\chi_C(G)$, is the smallest integer k such that G is admits a clique-coloring using at most k colors. Note that every proper coloring of G is also a clique-coloring of G, and so $\chi_C(G) \leq \chi(G)$. Furthermore, if G is triangle-free, then $\chi_C(G) = \chi(G)$ (since there are triangle-free graphs of arbitrarily large chromatic number, this implies that there are triangle-free graphs of arbitrarily large clique-chromatic number). However, if G contains triangles, $\chi_C(G)$ may be much smaller than $\chi(G)$. For instance, if G contains a dominating vertex, then $\chi_C(G) \leq 2$ (we assign the color 1 to the dominating vertex and the color 2 to all other vertices of G), while $\chi(G)$ may be arbitrarily large. Note that this implies that the clique-chromatic number is not monotone with respect to induced subgraphs, that is, there exist graphs H and G such that H is an induced subgraph of G, but $\chi_C(H) > \chi_C(G)$. (In particular, the restriction of a clique-coloring of G to an induced subgraph H of G need not be a clique-coloring of H.)

A graph G is *perfect* if all its induced subgraphs H satisfy $\chi(H) = \omega(H)$, where $\omega(H)$ denotes the size of a maximum clique. It was asked by Duffus, Sands, Sauer, and Woodrow in a paper from 1991 whether perfect graphs have a bounded clique-chromatic number and indeed it has been proven since that for many sublasses of the class of perfect graphs, this holds. Even more, until now it was not known whether there were any perfect graphs of clique-chromatic number greater than three. The main result of [4] is to prove that there exist perfect graphs of arbitrarily large clique-chromatic number, which gives a negative answer for the question of Duffus et al. mentioned above.

6.1.5. Algorithmic Aspects of Switch Cographs

Cographs are the graphs totally decomposable using series and parallel operations, in [5] we introduced an interesting generalization, namely the class of switch cographs. These are the class of graphs that are totally decomposable w.r.t involution modular decomposition — a generalization of the modular decomposition of 2-structure, which has a unique linear-sized decomposition tree. We use our new decomposition tool to design three practical algorithms for the maximum cut, vertex cover and vertex separator problems. The complexity of these problems was previously unknown for this class of graphs.

6.1.6. Shrinking Maxima, Decreasing Costs: New Online Packing and Covering Problems

In [16], we consider two new variants of online integer programs that are duals. In the packing problem we are given a set of items and a collection of knapsack constraints over these items that are revealed over time in an online fashion. Upon arrival of a constraint we may need to remove several items (irrevocably) so as to maintain feasibility of the solution. Hence, the set of packed items becomes smaller over time. The goal is to maximize the number, or value, of packed items. The problem originates from a buffer-overflow model in communication networks, where items represent information units broken into multiple packets. The other problem considered is online covering: there is a universe to be covered. Sets arrive online, and we must decide for each set whether we add it to the cover or give it up. The cost of a solution is the total cost of sets taken, plus a penalty for each uncovered element. The number of sets in the solution grows over time, but its cost goes down. This problem is motivated by team formation, where the universe consists of skills, and sets represent candidates we may hire. The packing problem was introduced in Emek et al. (SIAM J Comput 41(4):728-746, 2012) for the special case where the matrix is binary; in this paper we extend the solution to general matrices with non-negative integer entries. The covering problem is introduced in this paper; we present matching upper and lower bounds on its competitive ratio.

6.1.7. The Complexity of the Shortest-path Broadcast Problem

In [8], we study the shortest-path broadcast problem in graphs and digraphs, where a message has to be transmitted from a source node s to all the nodes along shortest paths, in the classical telephone model. For both graphs and digraphs, we show that the problem is equivalent to the broadcast problem in layered directed graphs. We then prove that this latter problem is NP-hard, and therefore that the shortest-path broadcast problem is NP-hard in graphs as well as in digraphs. Nevertheless, we prove that a simple polynomial-time algorithm, called MDST-broadcast, based on min-degree spanning trees, approximates the optimal broadcast time within a multiplicative factor $3/2$ in 3-layer digraphs, and $O(\log n/\log\log n)$ in arbitrary multi-layer digraphs. As a consequence, one can approximate the optimal shortest-path broadcast time in polynomial time within a multiplicative factor $3/2$ whenever the source has eccentricity at most 2, and within a multiplicative factor $O(\log n/\log\log n)$ in the general case, for both graphs and digraphs. The analysis of MDST-broadcast is tight, as we prove that this algorithm cannot approximate the optimal broadcast time within a factor smaller than $\Omega(\log n/\log\log n)$.

6.1.8. Setting Ports in an Anonymous Network: How to Reduce the Level of Symmetry

A fundamental question in the setting of anonymous graphs concerns the ability of nodes to spontaneously break symmetries, based on their local perception of the network. In contrast to previous work, which focuses on symmetry breaking under arbitrary port labelings, in [37] we study the following design question: Given an anonymous graph G without port labels, how to assign labels to the ports of G, in interval form at each vertex, so that symmetry breaking can be achieved using a message-passing protocol requiring as few rounds of synchronous communication as possible?

More formally, for an integer $l > 0$, the *truncated view* $\mathcal{V}_l(v)$ of a node v of a port-labeled graph is defined as a tree encoding labels encountered along all walks in the network which originate from node v and have length at most l, and we ask about an assignment of labels to the ports of G so that the views $\mathcal{V}_l(v)$ are distinct for all nodes $v \in V$, with the goal being to minimize l.

We present such efficient port labelings for any graph G, and we exhibit examples of graphs showing that the derived bounds are asymptotically optimal in general. More precisely, our results imply the following statements.

1. For any graph G with n nodes and diameter D, a uniformly random port labeling achieves $l = O(\min(D, \log n))$, w.h.p.
2. For any graph G with n nodes and diameter D, it is possible to construct in polynomial time a labeling that satisfies $l = O(\min(D, \log n))$.
3. For any integers $n \geq 2$ and $D \leq \log_2 n - \log_2 \log_2 n$, there exists a graph G with n nodes and diameter D which satisfies $l \geq \frac{1}{2}D - \frac{5}{2}$.

6.1.9. Robustness of the Rotor-Router Mechanism

The *rotor-router model*, also called the *Propp machine*, was first considered as a deterministic alternative to the random walk. The edges adjacent to each node v (or equivalently, the exit ports at v) are arranged in a fixed cyclic order, which does not change during the exploration. Each node v maintains a *port pointer* π_v which indicates the exit port to be adopted by an agent on the conclusion of the next visit to this node (the "next exit port"). The rotor-router mechanism guarantees that after each consecutive visit at the same node, the pointer at this node is moved to the next port in the cyclic order. It is known that, in an undirected graph G with m edges, the route adopted by an agent controlled by the rotor-router mechanism forms eventually an Euler tour based on arcs obtained via replacing each edge in G by two arcs with opposite direction. The process of ushering the agent to an Euler tour is referred to as the *lock-in problem*. In [Yanovski et al., Algorithmica 37(3), 165–186 (2003)], it was proved that, independently of the initial configuration of the rotor-router mechanism in G, the agent locks-in in time bounded by $2mD$, where D is the diameter of G.

In [2], we examine the dependence of the lock-in time on the initial configuration of the rotor-router mechanism. Our analysis is performed in the form of a game between a player p intending to lock-in the agent in an Euler tour as quickly as possible and its adversary a with the counter objective. We consider all cases of who decides the initial cyclic orders and the initial values π_v. We show, for example, that if a provides its own port numbering after the initial setup of pointers by p, the complexity of the lock-in problem is $O(m \cdot \min\{\log m, D\})$.

We also investigate the robustness of the rotor-router graph exploration in presence of faults in the pointers π_v or dynamic changes in the graph. We show, for example, that after the exploration establishes an Eulerian cycle, if k edges are added to the graph, then a new Eulerian cycle is established within $O(km)$ steps.

6.1.10. The Multi-Agent Rotor-Router on the Ring: A Deterministic Alternative to Parallel Random Walks

Continuing the line of research on the rotor-router model, in [18] we consider the setting in which multiple, indistinguishable agents are deployed in parallel in the nodes of the graph, and move around the graph in synchronous rounds, interacting with a single rotor-router system. We propose new techniques which allow us to perform a theoretical analysis of the multi-agent rotor-router model, and to compare it to the scenario of parallel independent random walks in a graph. Our main results concern the n-node ring, and suggest a strong similarity between the performance characteristics of this deterministic model and random walks.

We show that on the ring the rotor-router with k agents admits a cover time of between $\Theta(n^2/k^2)$ in the best case and $\Theta(n^2/\log k)$ in the worst case, depending on the initial locations of the agents, and that both these bounds are tight. The corresponding expected value of the cover time for k random walks, depending on the initial locations of the walkers, is proven to belong to a similar range, namely between $\Theta(n^2/(k^2/\log^2 k))$ and $\Theta(n^2/\log k)$.

Finally, we study the limit behavior of the rotor-router system. We show that, once the rotor-router system has stabilized, all the nodes of the ring are always visited by some agent every $\Theta(n/k)$ steps, regardless of how the system was initialized. This asymptotic bound corresponds to the expected time between successive visits to a node in the case of k random walks. All our results hold up to a polynomially large number of agents ($1 \leq k < n^{1/11}$).

6.1.11. Bounds on the Cover Time of Parallel Rotor Walks

In [12], we study the parallel rotor-router model in the case of general graphs. We consider the cover time of such a system, i.e., the number of steps after which each node has been visited by at least one walk, regardless of the initialization of the walks. We show that for any graph with m edges and diameter D, this cover time is at most $\Theta(mD/\log k)$ and at least $\Theta(mD/k)$, which corresponds to a speedup of between $\Theta(\log k)$ and $\Theta(k)$ with respect to the cover time of a single walk.

6.2. Distributed Computing

6.2.1. Local Conflict Coloring

Locally finding a solution to symmetry-breaking tasks such as vertex-coloring, edge-coloring, maximal matching, maximal independent set, etc., is a long-standing challenge in distributed network computing. More recently, it has also become a challenge in the framework of centralized local computation. In [30], we introduce conflict coloring as a general symmetry-breaking task that includes all the aforementioned tasks as specific instantiations — conflict coloring includes all locally checkable labeling tasks from [Naor&Stockmeyer, STOC 1993]. Conflict coloring is characterized by two parameters l and d, where the former measures the amount of freedom given to the nodes for selecting their colors, and the latter measures the number of constraints which colors of adjacent nodes are subject to. We show that, in the standard LOCAL model for distributed network computing, if $l/d > \Delta$, then conflict coloring can be solved in $\widetilde{O}(\sqrt{\Delta}) + \log^* n$ rounds in n-node graphs with maximum degree Δ, where $\widetilde{O}$ ignores the polylog factors in Δ. The dependency in n is optimal, as a consequence of the $\Omega(\log^* n)$ lower bound by [Linial, SIAM J. Comp. 1992] for $(\Delta + 1)$-coloring. An important special case of our result is a significant improvement over the best known algorithm for distributed $(\Delta + 1)$-coloring due to [Barenboim, PODC 2015], which required $\widetilde{O}(\Delta^{3/4}) + \log^* n$ rounds. Improvements for other variants of coloring, including $(\Delta + 1)$-list-coloring, $(2\Delta - 1)$-edge-coloring, T-coloring, etc., also follow from our general result on conflict coloring. Likewise, in the framework of centralized local computation algorithms (LCAs), our general result yields an LCA which requires a smaller number of probes than the previously best known algorithm for vertex-coloring, and works for a wide range of coloring problems.

6.2.2. A Hierarchy of Local Decision

In [29], we extend the notion of *distributed decision* in the framework of distributed network computing, inspired by recent results on so-called *distributed graph automata*. We show that, by using distributed decision mechanisms based on the interaction between a *prover* and a *disprover*, the size of the certificates distributed to the nodes for certifying a given network property can be drastically reduced. For instance, we prove that minimum spanning tree can be certified with $O(\log n)$-bit certificates in n-node graphs, with just one interaction between the prover and the disprover, while it is known that certifying MST requires $\Omega(\log^2 n)$-bit certificates if only the prover can act. The improvement can even be exponential for some simple graph properties. For instance, it is known that certifying the existence of a nontrivial automorphism requires $\Omega(n^2)$ bits if only the prover can act. We show that there is a protocol with two interactions between the prover and the disprover enabling to certify nontrivial automorphism with $O(\log n)$-bit certificates. These results are achieved by defining and analysing a *local hierarchy* of decision which generalizes the classical notions of *proof-labelling schemes* and *locally checkable proofs*.

6.2.3. Distributed Testing of Excluded Subgraphs

In [35], we study property testing in the context of distributed computing, under the classical CONGEST model. It is known that testing whether a graph is triangle-free can be done in a constant number of rounds, where the constant depends on how far the input graph is from being triangle-free. We show that, for every connected 4-node graph H, testing whether a graph is H-free can be done in a constant number of rounds too. The constant also depends on how far the input graph is from being H-free, and the dependence is identical to the one in the case of testing triangle-freeness. Hence, in particular, testing whether a graph is K_4-free, and testing whether a graph is C_4-free can be done in a constant number of rounds (where K_k denotes the k-node clique, and C_k denotes the k-node cycle). On the other hand, we show that testing K_k-freeness and C_k-freeness for $k \geq 5$ appear to be much harder. Specifically, we investigate two natural types of generic algorithms for testing H-freeness, called DFS tester and BFS tester. The latter captures the previously known algorithm to test the presence of triangles, while the former captures our generic algorithm to test the presence of a 4-node graph pattern H. We prove that both DFS and BFS testers fail to test K_k-freeness and C_k-freeness in a constant number of rounds for $k \geq 5$.

6.2.4. Asynchronous Coordination Under Preferences and Constraints

Adaptive renaming can be viewed as a coordination task involving a set of asynchronous agents, each aiming at grabbing a single resource out of a set of resources totally ordered by their desirability. Similarly, musical chairs is also defined as a coordination task involving a set of asynchronous agents, each aiming at picking one of a set of available resources, where every agent comes with an a priori preference for some resource. In [22], we foresee instances in which some combinations of resources are allowed, while others are disallowed. We model these constraints, i.e., the restrictions on the ability to use some combinations of resources, as an undirected graph whose nodes represent the resources, and an edge between two resources indicates that these two resources cannot be used simultaneously. In other words, the sets of resources that are allowed are those which form independent sets in the graph. E.g., renaming and musical chairs are specific cases where the graph is stable (i.e., it the empty graph containing no edges). As for musical chairs, we assume that each agent comes with an a priori preference for some resource. If an agent's preference is not in conflict with the preferences of the other agents, then this preference can be grabbed by the agent. Otherwise, the agents must coordinate to resolve their conflicts, and potentially choose non preferred resources. We investigate the following problem: given a graph, what is the maximum number of agents that can be accommodated subject to non-altruistic behaviors of early arriving agents? We entirely solve this problem under the restriction that agents which cannot grab their preferred resources must then choose a resource among the nodes of a predefined independent set. However, the general case, where agents which cannot grab their preferred resource are then free to choose any resource, is shown to be far more complex. In particular, just for cyclic constraints, the problem is surprisingly difficult. Indeed, we show that, intriguingly, the natural algorithm inspired from optimal solutions to adaptive renaming or musical chairs is sub-optimal for cycles, but proven to be at most 1 to the optimal. The main message of this paper is that finding optimal solutions to the coordination with constraints and preferences task requires to design " dynamic " algorithms, that is, algorithms of a completely different nature than the "static" algorithms used for, e.g., renaming.

6.2.5. Making Local Algorithms Wait-Free: The Case of Ring Coloring

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. This work proposes a new DECOUPLED model in an attempt to reconcile these two worlds. It consists of a synchronous and reliable communication graph of n nodes, and on top a set of asynchronous crash-prone processes, each attached to a communication node. To illustrate the DECOUPLED model, the paper [21] 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 vertex coloring algorithm in which the processes are not required to start simultaneously.

6.2.6. t-Resilient Immediate Snapshot Is Impossible

An immediate snapshot object is a high level communication object, built on top of a read/write distributed system in which all except one processes may crash. It allows each process to write a value and obtains a set of pairs (process id, value) such that, despite process crashes and asynchrony, the sets obtained by the processes satisfy noteworthy inclusion properties. Considering an n-process model in which up to t processes are allowed to crash (t-crash system model), the paper [25] is on the construction of t-resilient immediate snapshot objects. In the t-crash system model, a process can obtain values from at least $(n-t)$ processes, and, consequently, t-immediate snapshot is assumed to have the properties of the basic $(n-1)$-resilient immediate snapshot plus the additional property stating that each process obtains values from at least $(n-t)$ processes. The main result of the work is the following. While there is a (deterministic) $(n-1)$-resilient algorithm implementing the basic $(n-1)$-immediate snapshot in an $(n-1)$-crash read/write system, there is no t-resilient algorithm in a t-crash read/write model when $t \in [1...(n-2)]$. This means that, when $t < n-1$, the notion of t-resilience is inoperative when one has to implement t-immediate snapshot for these values of t: the model assumption "at most $t < n-1$ processes may crash" does not provide us with additional computational power allowing for the design of a genuine t-resilient algorithm (genuine meaning that such an algorithm would work in the t-crash model, but not in the $(t+1)$-crash model). To show these results, we rely on well-known distributed computing agreement problems such as consensus and k-set agreement.

6.2.7. Perfect Failure Detection with Very Few Bits

A *failure detector* is a distributed oracle that provides the processes with information about failures. The *perfect* failure detector provides accurate and eventually complete information about process failures. In [34], we show that, in asynchronous failure-prone message-passing systems, perfect failure detection can be achieved by an oracle that outputs at most $\lceil \log \alpha(n) \rceil + 1$ bits per process in n-process systems, where α denotes the inverse-Ackermann function. This result is essentially optimal, as we also show that, in the same environment, no failure detectors outputting a constant number of bit per process can achieve perfect failure detection.

6.2.8. Decentralized Asynchronous Crash-Resilient Runtime Verification

Runtime Verification (RV) is a lightweight method for monitoring the formal specification of a system during its execution. It has recently been shown that a given state predicate can be monitored consistently by a set of crash-prone asynchronous *distributed* monitors, only if sufficiently many different verdicts can be emitted by each monitor. In [27], we revisit this impossibility result in the context of LTL semantics for RV. We show that employing the four-valued logic RVLTL will result in inconsistent distributed monitoring for some formulas. Our first main contribution is a family of logics, called LTL(k), that refines RVLTL incorporating $2k+4$ truth values, for each $k \geq 0$. The truth values of LTL(k) can be effectively used by each monitor to reach a consistent global set of verdicts for each given formula, provided k is sufficiently large. Our second main contribution is an algorithm for monitor construction enabling fault-tolerant distributed monitoring based on the aggregation of the individual verdicts by each monitor.

6.2.9. Asynchronous Consensus with Bounded Memory

The paper [11] presents a bounded memory size Obstruction-Free consensus algorithm for the asynchronous shared memory model. More precisely for a set of n processes, this algorithm uses $n+2$ multi-writer multi-reader registers, each of these registers being of size $O(\log n)$ bits. From this, we get a bounded memory size space complexity consensus algorithm with single-writer multi-reader registers and a bounded memory size space complexity consensus algorithm in the asynchronous message passing model with a majority of correct processes. As it is easy to ensure the Obstruction-Free assumption with randomization (or with leader election failure detector Ω) we obtain a bounded memory size randomized consensus algorithm and a bounded memory size consensus algorithm with failure detector.

6.2.10. Implementing Snapshot Objects on Top of Crash-Prone Asynchronous Message-Passing Systems

Distributed snapshots, as introduced by Chandy and Lamport in the context of asynchronous failure-free message-passing distributed systems, are consistent global states in which the observed distributed application might have passed through. It appears that two such distributed snapshots cannot necessarily be compared (in the sense of determining which one of them is the "first"). Differently, snapshots introduced in asynchronous crash-prone read/write distributed systems are totally ordered, which greatly simplify their use by upper layer applications. In order to benefit from shared memory snapshot objects, it is possible to simulate a read/write shared memory on top of an asynchronous crash-prone message-passing system, and build then snapshot objects on top of it. This algorithm stacking is costly in both time and messages. To circumvent this drawback, the paper [24] presents algorithms building snapshot objects directly on top of asynchronous crash-prone message-passing system. "Directly" means here "without building an intermediate layer such as a read/write shared memory". To the authors knowledge, the proposed algorithms are the first providing such constructions. Interestingly enough, these algorithms are efficient and relatively simple.

6.2.11. Set-Consensus Collections are Decidable

A natural way to measure the power of a distributed-computing model is to characterize the set of tasks that can be solved in it. In general, however, the question of whether a given task can be solved in a given model is undecidable, even if we only consider the wait-free shared-memory. In [23], we address this question for restricted classes of models and tasks. We show that the question of whether a collection C of (ℓ, j)-set consensus objects, for various ℓ (the number of processes that can invoke the object) and j (the number of distinct outputs the object returns), can be used by n processes to solve wait-free k-set consensus is decidable. Moreover, we provide a simple $O(n^2)$ decision algorithm, based on a dynamic programming solution to the Knapsack optimization problem. We then present an adaptive wait-free set-consensus algorithm that, for each set of participating processes, achieves the best level of agreement that is possible to achieve using C. Overall, this gives us a complete characterization of a read-write model defined by a collection of set-consensus objects through its set-consensus power.

6.2.12. Minimizing the Number of Opinions for Fault-Tolerant Distributed Decision Using Well-Quasi Orderings

The notion of deciding a *distributed language*L is of growing interest in various distributed computing settings. Each process p_i is given an input value x_i, and the processes should collectively decide whether their set of input values $x = (x_i)_i$ is a valid state of the system w.r.t. to some specification, i.e., if $x \in L$. In *non-deterministic* distributed decision each process p_i gets a local certificate c_i in addition to its input x_i. If the input $x \in L$then there exists a certificate $c = (c_i)_i$ such that the processes collectively accept x, and if $x \notin L$, then for every c, the processes should collectively reject x. The collective decision is expressed by the set of *opinions* emitted by the processes, and one aims at minimizing the number of possible opinions emitted by each process. In [33], we study non-deterministic distributed decision in asynchronous systems where processes may crash. In this setting, it is known that the number of opinions needed to deterministically decide a language can grow with n, the number of processes in the system.We prove that every distributed language L can be non-deterministically decided using only three opinions, with certificates of size $\lceil \log \alpha(n) \rceil + 1$ bits, where α grows at least as slowly as the inverse of the Ackerman function. The result is optimal, as we show that there are distributed languages that cannot be decided using just two opinions, even with arbitrarily large certificates. To prove our upper bound, we introduce the notion of *distributed encoding of the integers*, that provides an explicit construction of a long *bad sequence* in the *well-quasi-ordering*$(\{0,1\}^*, \leq_*)$ controlled by the successor function. Thus, we provide a new class of applications for well-quasi-orderings that lies outside logic and complexity theory. For the lower bound we use combinatorial topology techniques.

6.2.13. Collision-Free Network Exploration

In [9], we consider a network exploration setting in which mobile agents start at different nodes of an n-node network. The agents synchronously move along the network edges in a *collision-free* way, i.e., in no round two agents may occupy the same node. An agent has no knowledge of the number and initial positions of other agents. We are looking for the shortest time required to reach a configuration in which each agent has visited all nodes and returned to its starting location. In the scenario when each mobile agent knows the map of the network, we provide tight (up to a constant factor) lower and upper bounds on the collision-free exploration time in arbitrary graphs, and the exact bound for the trees. In the second scenario, where the network is unknown to the agents, we propose collision-free exploration strategies running in $O(n^2)$ rounds in tree networks and in $O(n^5 \log n)$ rounds in networks with an arbitrary topology.

6.2.14. When Patrolmen Become Corrupted: Monitoring a Graph Using Faulty Mobile Robots

In [10], we consider a setting in which a team of k mobile robots is deployed on a weighted graph whose edge weights represent distances. The robots perpetually move along the domain, represented by all points belonging to the graph edges, not exceeding their maximal speed. The robots need to patrol the graph by regularly visiting all points of the domain. We consider a team of robots (patrolmen), at most f of which may be unreliable, failing to comply with their patrolling duties. What algorithm should be followed so as to minimize the maximum time between successive visits of every edge point by a reliable patrolmen? The corresponding measure of efficiency of patrolling called *idleness* has been widely accepted in the robotics literature. We extend it to the case of untrusted patrolmen; we denote by $I_k^f(G)$ the maximum time that a point of the domain may remain unvisited by reliable patrolmen. The objective is to find patrolling strategies minimizing $I_k^f(G)$.

We investigate this problem for various classes of graphs. We design optimal algorithms for line segments, which turn out to be surprisingly different from strategies for related patrolling problems proposed in the literature. We then use these results to provide algorithms for general graphs. For Eulerian graphs G, we give an optimal patrolling strategy with idleness $I_k^f(G) = (f+1)E(G)/k$, where $E(G)$ is the sum of the lengths of the edges of G. For arbitrary graphs and given ratio r of faulty robots, $r := f/k < 1/2$, we design a strategy which is a $(1+\epsilon)$ approximation of the optimal one, for sufficiently large k. Further, we show the hardness of the problem of computing the idle time for three robots, at most one of which is faulty, by reduction from 3-edge-coloring of cubic graphs — a known NP-hard problem. A byproduct of our proof is the investigation of classes of graphs minimizing idle time (with respect to the total length of edges); an example of such a class is known in the literature under the name of Kotzig graphs.

6.2.15. Noisy Rumor Spreading and Plurality Consensus

Error-correcting codes are efficient methods for handling noisy communication channels in the context of technological networks. However, such elaborate methods differ a lot from the unsophisticated way biological entities are supposed to communicate. Yet, it has been recently shown by Feinerman, Haeupler, and Korman [PODC 2014] that complex coordination tasks such as rumor spreading and majority consensus can plausibly be achieved in biological systems subject to noisy communication channels, where every message transferred through a channel remains intact with small probability $1+\varepsilon$, without using coding techniques. This result is a considerable step towards a better understanding of the way biological entities may cooperate. It has nevertheless been established only in the case of 2-valued opinions: rumor spreading aims at broadcasting a single-bit opinion to all nodes, and majority consensus aims at leading all nodes to adopt the single-bit opinion that was initially present in the system with (relative) majority. In [32], we extend this previous work to k-valued opinions, for any constant $k \geq 2$. Our extension requires to address a series of important issues, some conceptual, others technical. We had to entirely revisit the notion of noise, for handling channels carrying k-valued messages. In fact, we precisely characterize the type of noise patterns for which plurality consensus is solvable. Also, a key result employed in the bivalued case by Feinerman et al. is an estimate of the probability of observing the most frequent opinion from observing the mode of a small sample. We generalize this result to the multivalued case by providing a new analytical proof for the bivalued case that is amenable to be extended, by induction, and that is of independent interest.

6.3. Models and Algorithms for Networks

6.3.1. Beyond Highway Dimension: Small Distance Labels Using Tree Skeletons

The goal of a hub-based distance labeling scheme for a network $G = (V, E)$ is to assign a small subset $S(u) \subseteq V$ to each node $u \in V$, in such a way that for any pair of nodes u, v, the intersection of hub sets $S(u) \cap S(v)$ contains a node on the shortest uv-path. The existence of small hub sets, and consequently efficient shortest path processing algorithms, for road networks is an empirical observation. A theoretical explanation for this phenomenon was proposed by Abraham et al. (SODA 2010) through a network parameter they called highway dimension, which captures the size of a hitting set for a collection of shortest paths of length at least r intersecting a given ball of radius $2r$. In [38], we revisit this explanation, introducing a more tractable (and directly comparable) parameter based solely on the structure of shortest-path spanning trees, which we call skeleton dimension. We show that skeleton dimension admits an intuitive definition for both directed and undirected graphs, provides a way of computing labels more efficiently than by using highway dimension, and leads to comparable or stronger theoretical bounds on hub set size.

6.3.2. Sublinear-Space Distance Labeling using Hubs

Continuing work in the previously discussed framework of hub-based distance labeling schemes, in [36], [39], we present a hub labeling which allows us to decode exact distances in sparse graphs using labels of size sublinear in the number of nodes. For graphs with at most n nodes and average degree Δ, the tradeoff between label bit size L and query decoding time T for our approach is given by $L = O(n \log \log_\Delta T / \log_\Delta T)$, for any $T \leq n$. Our simple approach is thus the first sublinear-space distance labeling for sparse graphs that simultaneously admits small decoding time (for constant Δ, we can achieve any $T = \omega(1)$ while maintaining $L = o(n)$), and it also provides an improvement in terms of label size with respect to previous slower approaches.

By using similar techniques, we then present a 2-additive labeling scheme for general graphs, i.e., one in which the decoder provides a 2-additive-approximation of the distance between any pair of nodes. We achieve almost the same label size-time tradeoff $L = O(n \log^2 \log T / \log T)$, for any $T \leq n$. To our knowledge, this is the first additive scheme with constant absolute error to use labels of sublinear size. The corresponding decoding time is then small (any $T = \omega(1)$ is sufficient).

We believe all of our techniques are of independent value and provide a desirable simplification of previous approaches.

6.3.3. Labeling Schemes for Ancestry Relation

In [17], we solve the ancestry-labeling scheme problem which aims at assigning the shortest possible labels (bit strings) to nodes of rooted trees, so that ancestry queries between any two nodes can be answered by inspecting their assigned labels only. This problem was introduced more than twenty years ago by Kannan et al. [STOC '88], and is among the most well-studied problems in the field of informative labeling schemes. We construct an ancestry-labeling scheme for n-node trees with label size $\log_2 n + O(\log \log n)$ bits, thus matching the $log_2 n + \Omega(\log \log n)$ bits lower bound given by Alstrup et al. [SODA '03]. Our scheme is based on a simplified ancestry scheme that operates extremely well on a restricted set of trees. In particular, for the set of n-node trees with depth at most d, the simplified ancestry scheme enjoys label size of $\log_2 n + 2 \log_2 d + O(1)$ bits. Since the depth of most XML trees is at most some small constant, such an ancestry scheme may be of practical use. In addition, we also obtain an adjacency-labeling scheme that labels n-node trees of depth d with labels of size $\log_2 n + 3 \log_2 d + O(1)$ bits. All our schemes assign the labels in linear time, and guarantee that any query can be answered in constant time. Finally, our ancestry scheme finds applications to the construction of small universal partially ordered sets (posets). Specifically, for any fixed integer k, it enables the construction of a universal poset of size $O(n^k)$ for the family of n-element posets with tree-dimension at most k. Up to lower order terms, this bound is tight thanks to a lower bound of $n^{k-o(1)}$ due to Alon and Scheinerman [Order '88].

6.3.4. Independent Lazy Better-Response Dynamics on Network Games

In [43], 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.

6.3.5. Forwarding Tables Verification through Representative Header Sets

Forwarding table verification consists in checking the distributed data-structure resulting from the forwarding tables of a network. A classical concern is the detection of loops. We study in [42] this problem in the context of software-defined networking (SDN) where forwarding rules can be arbitrary bitmasks (generalizing prefix matching) and where tables are updated by a centralized controller. Basic verification problems such as loop detection are NP-hard and most previous work solves them with heuristics or SAT solvers. We follow a different approach based on computing a representation of the header classes, i.e. the sets of headers that match the same rules. This representation consists in a collection of representative header sets, at least one for each class, and can be computed centrally in time which is polynomial in the number of classes. Classical verification tasks can then be trivially solved by checking each representative header set. In general, the number of header classes can increase exponentially with header length, but it remains polynomial in the number of rules in the practical case where rules are constituted with predefined fields where exact, prefix matching or range matching is applied in each field (e.g., IP/MAC addresses, TCP/UDP ports). We propose general techniques that work in polynomial time as long as the number of classes of headers is polynomial and that do not make specific assumptions about the structure of the sets associated to rules. The efficiency of our method rely on the fact that the data-structure representing rules allows efficient computation of intersection, cardinal and inclusion. Finally, we propose an algorithm to maintain such representation in presence of updates (i.e., rule insert/update/removal). We also provide a local distributed algorithm for checking the absence of black-holes and a proof labeling scheme for locally checking the absence of loops.

6.3.6. A Locally-Blazed Ant Trail Achieves Efficient Collective Navigation Despite Limited Information

This work fits into the framework of computationally-inspired analysis of biological systems. Any organism faces sensory and cognitive limitations which may result in maladaptive decisions. Such limitations are prominent in the context of groups where the relevant information at the individual level may not coincide with collective requirements. In [14], we study the navigational decisions exhibited by *Paratrechina longicornis* ants as they cooperatively transport a large food item. These decisions hinge on the perception of individuals which often restricts them from providing the group with reliable directional information. We find that, to achieve efficient navigation despite partial and even misleading information, these ants employ a locally-blazed trail. This trail significantly deviates from the classical notion of an ant trail: First, instead of systematically marking the full path, ants mark short segments originating at the load. Second, the carrying team constantly loses the guiding trail. We experimentally and theoretically show that the locally-blazed trail optimally and robustly exploits useful knowledge while avoiding the pitfalls of misleading information.

6.3.7. Parallel Exhaustive Search without Coordination

In [31], we analyze parallel algorithms in the context of *exhaustive search* over totally ordered sets. Imagine an infinite list of “boxes”, with a “treasure” hidden in one of them, where the boxes’ order reflects the importance of finding the treasure in a given box. At each time step, a search protocol executed by a searcher has the ability to peek into one box, and see whether the treasure is present or not. Clearly, the best strategy of a single searcher would be to open the boxes one by one, in increasing order. Moreover, by equally dividing the workload between them, k searchers can trivially find the treasure k times faster than one searcher. However, this straightforward strategy is very sensitive to failures (*e.g.,* crashes of processors), and overcoming this issue seems to require a large amount of communication. We therefore address the question of designing parallel search algorithms maximizing their *speed-up* and maintaining high levels of *robustness*, while minimizing the amount of resources for coordination. Based on the observation that algorithms that avoid communication are inherently robust, we focus our attention on identifying the best running time performance of *non-coordinating* algorithms. Specifically, we devise non-coordinating algorithms that achieve a speed-up of $9/8$ for two searchers, a speed-up of $4/3$ for three searchers, and in general, a speed-up of $\frac{k}{4}(1+1/k)^2$ for any $k \geq 1$ searchers. Thus, asymptotically, the speed-up is only four times worse compared to the case of full-coordination. Moreover, these bounds are tight in a strong sense as no non-coordinating search algorithm can achieve better speed-ups. Furthermore, our algorithms are surprisingly simple and hence applicable. Overall, we highlight that, in faulty contexts in which coordination between the searchers is technically difficult to implement, intrusive with respect to privacy, and/or costly in term of resources, it might well be worth giving up on coordination, and simply run our non-coordinating exhaustive search algorithms.

6.3.8. Rumor Spreading in Random Evolving Graphs

Randomized gossip is one of the most popular way of disseminating information in large scale networks. This method is appreciated for its simplicity, robustness, and efficiency. In the Push protocol, every informed node selects, at every time step (a.k.a. round), one of its neighboring node uniformly at random and forwards the information to this node. This protocol is known to complete information spreading in $O(\log n)$ time steps with high probability (w.h.p.) in several families of n-node *static* networks. The Push protocol has also been empirically shown to perform well in practice, and, specifically, toe robust against dynamic topological changes. In [15], we aim at analyzing the Push protocol in *dynamic* networks. We consider the *edge-Markovian* evolving graph model which captures natural temporal dependencies between the structure of the network at time t, and the one at time $t+1$. Precisely, a non-edge appears with probability p, while an existing edge dies with probability q. In order to fit with real-world traces, we mostly concentrate our study on the case where $p = \Omega(\frac{1}{n})$ and q is constant. We prove that, in this realistic scenario, the Push protocol does perform well, completing information spreading in $O(\log n)$ time steps w.h.p. Note that this performance holds even when the network is, w.h.p., disconnected at every time step (e.g., when $p \ll \frac{\log n}{n}$). Our result provides the first formal argument demonstrating the robustness of the Push protocol against network changes. We also address another range of parameters p and q, namely $p + q = 1$ with arbitrary p and q. Although this latter range does not precisely fit with the measures performed on real-world traces, they can be of independent interest for other settings. The result in this case confirms the positive impact of dynamism.

6.3.9. Sparsifying Congested Cliques and Core-Periphery Networks

The *core-periphery* network architecture proposed by Avin et al. [ICALP 2014] was shown to support fast computation for many distributed algorithms, while being much sparser than the *congested clique*. For being efficient, the core-periphery architecture is however bounded to satisfy three axioms, among which is the capability of the core to emulate the clique, i.e., to implement the all-to-all communication pattern, in $O(1)$ rounds in the CONGEST model. In [26], we show that implementing all-to-all communication in k rounds can be done in n-node networks with roughly n^2/k edges, and this bound is tight. Hence, sparsifying the core beyond just saving a fraction of the edges requires to relax the constraint on the time to simulate the congested clique. We show that, for $p \gg \sqrt{\log n/n}$, a random graph in $\mathcal{G}_{n,p}$ can, w.h.p., perform the all-to-all communication pattern in $O(\min\{\frac{1}{p^2}, np\})$ rounds. Finally, we show that if the core can emulate the congested clique in t rounds, then there exists a distributed MST construction algorithm performing in $O(t \log n)$ rounds. Hence, for $t = O(1)$, our (deterministic) algorithm improves the best known (randomized) algorithm for constructing MST in core-periphery networks by a factor $\Theta(\log n)$.

6.3.10. Core-periphery Clustering and Collaboration Networks

In [28], we analyse the core-periphery clustering properties of collaboration networks, where the core of a network is formed by the nodes with highest degree. In particular, we first observe that, even for random graph models aiming at matching the degree-distribution and/or the clustering coefficient of real networks, these models produce synthetic graphs which have a spatial distribution of the triangles with respect to the core and to the periphery which does not match the spatial distribution of the triangles in the real networks. We therefore propose a new model, called CPCL, whose aim is to distribute the triangles in a way fitting with their real core-periphery distribution, and thus producing graphs matching the core-periphery clustering of real networks.

7. Bilateral Contracts and Grants with Industry

7.1. Collaboration with Nokia Bell Labs

Gang has a strong collaboration with Bell Labs (Nokia). We notably collaborate with Fabien Mathieu who is a former member of GANG and Nidhi Hegde. An ADR (joint research action) is dedicated to content centric networks and forwarding information verification. The PhD thesis of Leonardo Linguaglossa was funded by this contract.

This collaboration is developed inside the Alcatel-Lucent and Inria joint research lab.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. Laboratory of Information, Networking and Communication Sciences (LINCS)

Gang is participating to the LINCS, a research centre co-founded by Inria, Institut Mines-Télécom, UPMC and Alcatel-Lucent Bell Labs, dedicated to research and innovation in the domains of future information and communication networks, systems and services. Gang contributes to work on online social networks, content centric networking and forwarding information verification.

8.2. National Initiatives

8.2.1. ANR Displexity

Participants: Carole Gallet Delporte, Hugues Fauconnier, Pierre Fraigniaud, Amos Korman, Adrian Kosowski, Laurent Viennot.

Managed by University Paris Diderot, C. Delporte and H. Fauconnier lead this project that grants 1 Post-Doc.

Distributed computation keep raising new questions concerning computability and complexity. For instance, as far as fault-tolerant distributed computing is concerned, impossibility results do not depend on the computational power of the processes, demonstrating a form of undecidability which is significantly different from the one encountered in sequential computing. In the same way, as far as network computing is concerned, the impossibility of solving certain tasks locally does not depend on the computational power of the individual processes.

The main goal of DISPLEXITY (for DIStributed computing: computability and ComPLEXITY) is to establish the scientific foundations for building up a consistent theory of computability and complexity for distributed computing.

One difficulty to be faced by DISPLEXITY is to reconcile the different sub-communities corresponding to a variety of classes of distributed computing models. The current distributed computing community may indeed be viewed as two not necessarily disjoint sub-communities, one focusing on the impact of temporal issues, while the other focusing on the impact of spatial issues. The different working frameworks tackled by these two communities induce different objectives: computability is the main concern of the former, while complexity is the main concern of the latter.

Within DISPLEXITY, the reconciliation between the two communities will be achieved by focusing on the same class of problems, those for which the distributed outputs are interpreted as a single binary output: yes or no. Those are known as the yes/no-problems. The strength of DISPLEXITY is to gather specialists of the two main streams of distributed computing. Hence, DISPLEXITY will take advantage of the experience gained over the last decade by both communities concerning the challenges to be faced when building up a complexity theory encompassing more than a fragment of the field.

In order to reach its objectives, DISPLEXITY aims at achieving the following tasks:

- Formalizing yes/no-problems (decision problems) in the context of distributed computing. Such problems are expected to play an analogous role in the field of distributed computing as that played by decision problems in the context of sequential computing.
- Formalizing decision problems (yes/no-problems) in the context of distributed computing. Such problems are expected to play an analogous role in the field of distributed computing as that played by decision problems in the context of sequential computing.

- Revisiting the various explicit (e.g., failure-detectors) or implicit (e.g., a priori information) notions of oracles used in the context of distributed computing allowing us to express them in terms of decidability/complexity classes based on oracles.
- Identifying the impact of non-determinism on complexity in distributed computing. In particular, DISPLEXITY aims at a better understanding of the apparent lack of impact of non-determinism in the context of fault-tolerant computing, to be contrasted with the apparent huge impact of non-determinism in the context of network computing. Also, it is foreseen that non-determinism will enable the comparison of complexity classes defined in the context of fault-tolerance with complexity classes defined in the context of network computing.
- Last but not least, DISPLEXITY will focus on new computational paradigms and frameworks, including, but not limited to distributed quantum computing and algorithmic game theory (e.g., network formation games).

The project will have to face and solve a number of challenging problems. Hence, we have built the DISPLEXITY consortium so as to coordinate the efforts of those worldwide leaders in Distributed Computing who are working in our country. A successful execution of the project will result in a tremendous increase in the current knowledge and understanding of decentralized computing and place us in a unique position in the field.

The project has been extended until June 2016.

8.2.2. 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.

8.3. European Initiatives

8.3.1. 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.

8.3.2. 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.

8.4. International Initiatives

8.4.1. Inria International Partners

8.4.1.1. 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).

Pierluigi Crescenzi (University of Florence, Italy) is a frequent visitor to the team and maintains an active research collaboration with Gang team members (Pierre Fraigniaud).

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).

8.4.2. Participation in Other International Programs

Involvement in the bilateral Franco-Mexican project ECOS NORD (2013-2016) on "Distributed Verification". Pierre Fraigniaud was the project's co-coordinator for the French side. Partners: IRIF and LaBRI (France), UNAM (Mexico).

8.5. International Research Visitors

8.5.1. Visits of International Scientists

Eli Gafni (1 month – June 2016)

Zvi Lotker, guest of Amos Korman (2 months – May, June 2016)

Thomas Sauerwald, guest of Adrian Kosowski (1 month – November 2016)

8.5.2. Visits to International Teams

Sergio Rasjbaum's Team (UNAM), C. Delporte and H. Fauconnier, 10 days (March 2016)

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Pierre Fraigniaud has organized the 1st conference on Highlights of Algorithms (HALG 2016), Paris, June 2016. This conference has gathered more than 200 participants to attend talks presenting the most significant results on algorithms produced in the academic year 2015-2016. The 2nd issue of this conference, Highlights of Algorithms 2017, will take place in Berlin.
http://highlightsofalgorithms.org

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

- Carole Delporte was PC Co-chair of NETYS 2016 — the 4th International Conference on Networked Systems, Morocco, May 18-20, 2016.
- Pierre Fraigniaud is PC chair for the Track Algorithms of the 31st IEEE International Parallel & Distributed Processing Symposium (IPDPS), to be held in Orlando, Florida, USA, May 29 - June 2, 2017. http://www.ipdps.org

9.1.2.2. Member of Conference Program Committees

- Hugues Fauconnier: NETYS 2016, PODC 2016.
- Carole Delporte: ICDCN 2016, ICDCS 2016, DISC 2016, OPODIS 2016.
- Adrian Kosowski: PODC 2016, ALGOSENSORS 2016, ICALP 2017.
- Pierre Fraigniaud: WWW 2016, ESA 2016, IPDPS 2016, OPODIS 2016, SIROCCO 2016, ALGOSENSORS 2016, FUN 2016, WWW 2017, SPAA 2017, DISC 2017.
- Amos Korman: ICALP 2017

9.1.3. Journal

9.1.3.1. Member of 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).
- Pierre Fraigniaud is a member of the Editorial Board of Fundamenta Informaticae (FI).

9.1.4. Invited Talks

Adrian Kosowski: SIROCCO 2016

9.1.5. Scientific Expertise

Pierre Fraigniaud is member of the evaluation committee for the ERC Starting Grants (Panel 6).

9.1.6. Research Administration

Pierre Fraigniaud is director 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 programmation répartie, 33h, M2, Univ. Paris Diderot

Master: Carole Delporte, Cours et TP Protocoles des services internet, 44h, M2, Univ. Paris Diderot

Master: Carole Delporte, Cours Algorithme réparti, 33h, M2, Univ. Paris Diderot

Master: Carole Delporte and Hugues Fauconnier, Protocoles Réseaux, 72h, M1, Université Paris Diderot

Licence: Carole Delporte and Hugues Fauconnier, Sécurité informatique, 36h, L3, Univ. Paris Diderot

Licence: Hugues Fauconnier, Programmation objet et interfaces graphiques, 48h, L2-L3, EIDD

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

Master: Pierre Fraigniaud, Algorithmique avancée, 24h, Ecole Centrale Supélec Paris, M2

Master: Pierre Fraigniaud, Algorithmique parallèle et distribuée, 24h, Ecole Centrale Supélec Paris, M2

Master: Adrian Kosowski, Randomization in Computer Science: Games, Networks, Epidemic and Evolutionary Algorithms, 18h, 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 and Michel Habib, 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

Master : Laurent Viennot, Graph Mining, 3h, M2 MPRI, Univ. Paris Diderot

Licence: Michel Habib, Algorithmique, 45h, L, ENS Cachan

Master: Michel Habib, Algorithmique avancée, 24h, M1, Univ. Paris Diderot

Master: Michel Habib, Mobilité, 33h, M2, Univ. Paris Diderot

Master: Michel Habib, Méthodes et algorithmes pour l'accès à l'information numérique, 16h, M2, Univ. Paris Diderot

Master: Michel Habib, Algorithmique de graphes, 12h, M2, Univ. Paris Diderot

Licence: Pierre Charbit, Introduction a la Programmation, 30h, L1, Université Paris Diderot, France

Licence: Pierre Charbit, Automates finis, 52h, L2, Université Paris Diderot, France

Licence: Pierre Charbit, Types de Données et Objet, 52h, L1, Université Paris Diderot, France

Master: Pierre Charbit, Programmation, 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: Leonardo Linguaglossa (co-advised by Laurent Viennot, Fabien Mathieu and Diego Perino, both from Nokia Bell Labs) was a PhD hired by Inria through the ADR CCN contract. He obained his PhD last September [1] at Paris Diderot University.

PhD in progress: Simon Collet (co-advised by Amos Korman and Pierre Fraigniaud). Title of thesis is: "Algorithmic Game Theory Applied to Biology". Started September 2015.

PhD in progress: Lucas Boczkowski (co-advised by Amos Korman and Iordanis Kerenidis). Title of thesis is: "Computing with Limited Resources in Uncertain Environments". Started September 2015.

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: Laurent Feuilloley (advised by Pierre Fraigniaud). Title of thesis is: "Synchronous Distributed Computing". Started September 2015.

PhD in progress: Mengchuan Zou (co-advised by Adrian Kosowski and Michel Habib). Title of thesis is: "Local and Adaptive Algorithms for Optimization Problems in Large Networks". Started October 2016.

PhD in progress: Finn Volkel (advised by Michel Habib). Title of Thesis: "Convexity in graphs", started september 2016.

PhD in progress: Léo Planche (co-advised by Étienne Birmelé and Fabien de Montgolfier). Title if thesis is : "Classification de collections de graphes". Started October 2015.

PhD in progress: Alkida Balliu and Dennis Olivetti (PhD students from L'Aquilla University and Gran Sasso Science Institute) are supervised by Pierre Fraigniaud.

PhD in progress: Lucas Hosseini (co-advised by Pierre Charbit, Patrice Ossona de Mendez and Jaroslav Nešetřil since Sept. 2014. Title : Limits of Structures.

9.2.3. Juries

Laurent Viennot was president of the jury in the PhD defense of Claudio Imbrenda on "Analysing Traffic Cacheability in the Access Network at Line Rate" at Telecom ParisTech (November 2016).

Michel Habib was reviewer for the PhD thesis of Matteo Seminaroti, "Combinatorial Algorithms for the seriation problem", Tilburg University, Holland, december 2016.

Laurent Viennot was reviewer for the PhD thesis of Guillaume Ducoffe on "Proprie´te´s me´triques des grands graphes" at Côte d'Azur Univ. (December 2016).

Laurent Viennot was co-advisor for the PhD thesis of Leonardo Linguaglossa on "Two challenges of Software Networking: Name-based Forwarding and Table Verification" at Paris Diderot Univ. (September 2016).

Carole Delporte was reviewer for the PhD thesis of Claire Capdevielle on "Étude de la complexité des implémentations d'objets concurrents et sans attente, abandonnables ou solo-rapides" at Bordeaux Univ (November 2016).

9.3. Popularization

- Laurent Viennot is "commissaire d'exposition" for the permanent exposition on "Informatique et sciences du numérique" at Palais de la déécouverte in Paris (opening in October 2017).
- An article centered around the eLife 2016 paper "A locally-blazed ant trail achieves efficient collective navigation despite limited information" [14], whose co-authors include Amos Korman (co-corresponding author), Lucas Boczkowski, and Adrian Kosowski, appeared on the Israeli daily newspaper "Haaretz", Dec 2016.
- The team has made contributions to the Encyclopedia of Algorithms [44], [41].

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] L. LINGUAGLOSSA.*Two challenges of Software Networking: Name-based Forwarding and Table Verification*, Université Paris Diderot (Paris 7) Sorbonne Paris Cité, September 2016, https://tel.archives-ouvertes.fr/tel-01386788.

Articles in International Peer-Reviewed Journal

[2] E. BAMPAS, L. GĄSIENIEC, N. HANUSSE, D. ILCINKAS, R. KLASING, A. KOSOWSKI, T. RADZIK.*Robustness of the Rotor–Router Mechanism*, in "Algorithmica", June 2016 [*DOI :* 10.1007/S00453-016-0179-Y], https://hal.inria.fr/hal-01416012.

[3] M. BORASSI, P. CRESCENZI, M. HABIB.*Into the Square: On the Complexity of Some Quadratic-time Solvable Problems*, in "Electronic Notes in Theoretical Computer Science", 2016, vol. 322, p. 51-67, https://hal.inria.fr/hal-01422030.

[4] P. CHARBIT, I. PENEV, S. THOMASSÉ, N. TROTIGNON.*Perfect graphs of arbitrarily large clique-chromatic number*, in "Journal of Combinatorial Theory, Series B", January 2016, vol. 116, p. 456-464 [*DOI :* 10.1016/J.JCTB.2015.09.008], https://hal.archives-ouvertes.fr/hal-01324052.

[5] V. COHEN-ADDAD, M. HABIB, F. DE MONTGOLFIER.*Algorithmic aspects of switch cographs*, in "Discrete Applied Mathematics", 2016, vol. 200, p. 23 - 42 [*DOI :* 10.1016/J.DAM.2015.07.008], https://hal.inria.fr/hal-01423053.

[6] D. G. CORNEIL, J. DUSART, M. HABIB, E. KÖHLER.*On the power of graph searching for cocomparability graphs*, in "Siam Journal on Discrete Mathematics", 2016, https://hal.inria.fr/hal-01273687.

[7] D. G. CORNEIL, J. DUSART, M. HABIB, A. MAMCARZ, F. DE MONTGOLFIER.*A tie-break model for graph search*, in "Discrete Applied Mathematics", January 2016, vol. 199, p. 89–100 [*DOI :* 10.1016/J.DAM.2015.06.011], https://hal.inria.fr/hal-01255055.

[8] P. CRESCENZI, P. FRAIGNIAUD, M. M. HALLDORSSON, H. HARUTYUNYAN, C. PIERUCCI, A. PIETRACAPRINA, G. PUCCI.*On the complexity of the shortest-path broadcast problem*, in "Discrete Applied Mathematics", 2016, vol. 199, p. 101-109 [*DOI :* 10.1016/J.DAM.2015.05.004], https://hal.inria.fr/hal-01255198.

[9] J. CZYZOWICZ, D. DERENIOWSKI, L. GĄSIENIEC, R. KLASING, A. KOSOWSKI, D. PAJĄK.*Collision-Free Network Exploration*, in "Journal of Computer and System Sciences", December 2016 [*DOI :* 10.1016/J.JCSS.2016.11.008], https://hal.inria.fr/hal-01416026.

[10] J. CZYZOWICZ, L. GĄSIENIEC, A. KOSOWSKI, E. KRANAKIS, D. KRIZANC, N. TALEB.*When Patrolmen Become Corrupted: Monitoring a Graph Using Faulty Mobile Robots*, in "Algorithmica", October 2016 [*DOI :* 10.1007/S00453-016-0233-9], https://hal.inria.fr/hal-01416010.

[11] C. DELPORTE-GALLET, H. FAUCONNIER.*Asynchronous Consensus with Bounded Memory*, in "Lecture notes in computer science", May 2016, vol. 9944, 15 [*DOI :* 10.1007/978-3-319-46140-3_12], https://hal.inria.fr/hal-01416509.

[12] D. DERENIOWSKI, A. KOSOWSKI, D. PAJĄK, P. UZNANSKI.*Bounds on the cover time of parallel rotor walks*, in "Journal of Computer and System Sciences", August 2016, vol. 82, n^{o} 5, p. 802 - 816 [*DOI :* 10.1016/J.JCSS.2016.01.004], https://hal.archives-ouvertes.fr/hal-01415102.

[13] J. DUSART, M. HABIB.*A new LBFS-based algorithm for cocomparability graph recognition*, in "Discrete Mathematics", 2016 [*DOI :* 10.1016/J.DAM.2015.07.016], https://hal.inria.fr/hal-01274023.

[14] E. FONIO, Y. HEYMAN, L. BOCZKOWSKI, A. GELBLUM, A. KOSOWSKI, A. KORMAN, O. FEINERMAN.*A locally-blazed ant trail achieves efficient collective navigation despite limited information*, in "eLife", November 2016, vol. 2016;5:e20185 [*DOI :* 10.7554/ELIFE.20185], https://hal.inria.fr/hal-01413748.

[15] P. FRAIGNIAUD, A. CLEMENTI, P. CRESCENZI, C. DOERR, F. PASQUALE, R. SILVESTRI.*Rumor spreading in random evolving graphs*, in "Random Structures and Algorithms", 2016, vol. 48, n^{o} 2, p. 290-312, https://hal.inria.fr/hal-01423643.

[16] P. FRAIGNIAUD, M. M. HALLDORSSON, B. PATT-SHAMIR, D. RAWITZ, A. ROSÉN.*Shrinking Maxima, Decreasing Costs: New Online Packing and Covering Problems*, in "Algorithmica", 2016, vol. 74, n^{o} 4, p. 1205-1223, https://hal.inria.fr/hal-01423640.

[17] P. FRAIGNIAUD, A. KORMAN.*An Optimal Ancestry Labeling Scheme with Applications to XML Trees and Universal Posets*, in "Journal of the ACM", 2016, vol. 63, p. 1 - 31 [*DOI :* 10.1145/2794076], https://hal.inria.fr/hal-01393643.

[18] R. KLASING, A. KOSOWSKI, D. PAJĄK, T. SAUERWALD.*The multi-agent rotor-router on the ring: a deterministic alternative to parallel random walks*, in "Distributed Computing", September 2016 [*DOI :* 10.1007/S00446-016-0282-Y], https://hal.inria.fr/hal-01416011.

Invited Conferences

[19] A. KOSOWSKI.*What Makes a Distributed Problem Truly Local?*, in "SIROCCO 2016 - 23rd International Colloquium on Structural Information and Communication Complexity", Helsinki, Finland, J. SUOMELA (editor), Lecture Notes in Computer Science, Springer, July 2016, vol. 9988, 3 [*DOI :* 10.1007/978-3-319-48314-6], https://hal.inria.fr/hal-01415090.

International Conferences with Proceedings

[20] É. BIRMELÉ, F. DE MONTGOLFIER, L. PLANCHE.*Minimum Eccentricity Shortest Path Problem: An Approximation Algorithm and Relation with the k-Laminarity Problem*, in "COCOA 2016, Combinatorial Optimization and Applications - 10th International Conference", Hong Kong, China, December 2016, p. 216 - 229 [*DOI :* 10.1007/978-3-319-48749-6_16], https://hal.inria.fr/hal-01424469.

[21] A. CASTAÑEDA, C. DELPORTE-GALLET, H. FAUCONNIER, S. RAJSBAUM, M. RAYNAL.*Making Local Algorithms Wait-Free: The Case of Ring Coloring*, in "SSS", Lyon, France, Stabilization, Safety, and Security of Distributed Systems, November 2016, vol. 10083, 16, https://hal.inria.fr/hal-01416522.

[22] A. CASTAÑEDA, P. FRAIGNIAUD, E. GAFNI, S. RAJSBAUM, M. ROY.*Asynchronous Coordination Under Preferences and Constraints*, in "23rd International Colloquium on Structural Information and Communication Complexity", Helsinki, Finland, SIROCCO 2016 pre-proceedings, July 2016, https://hal.archives-ouvertes.fr/hal-01341710.

[23] C. DELPORTE-GALLET, H. FAUCONNIER, E. GAFNI, P. KUZNETSOV.*Set-Consensus Collections are Decidable*, in "OPODIS", MADRID, Spain, LNCS, December 2016, https://hal.inria.fr/hal-01416981.

[24] C. DELPORTE-GALLET, H. FAUCONNIER, M. RAYNAL, S. RAJSBAUM.*Implementing Snapshot Objects on Top of Crash-Prone Asynchronous Message-Passing Systems*, in "Algorithms and Architectures for Parallel Processing (ICA3PP)", Granada, France, Algorithms and Architectures for Parallel Processing Volume 10048 of the series Lecture Notes in Computer Science pp 341-355, Springer, December 2016, vol. 10048, 15, https://hal.inria.fr/hal-01416498.

[25] C. DELPORTE-GALLET, H. FAUCONNIER, M. RAYNAL, S. RAJSBAUM.*t-Resilient Immediate Snapshot Is Impossible*, in "SIROCCO", Helsiinki, France, Structural Information and Communication Complexity, July 2016, vol. 9988, 15, https://hal.inria.fr/hal-01416515.

[26] P. FRAIGNIAUD, B. ALKIDA, Z. LOTKER, O. DENNIS.*Sparsifying Congested Cliques and Core-Periphery Networks*, in "23rd International Colloquium on Structural Information and Communication Complexity (SIROCCO)", Helsinki, Finland, 2016, https://hal.inria.fr/hal-01423688.

[27] P. FRAIGNIAUD, B. BONAKDARPOUR, S. RAJSBAUM, D. ROSENBLUETH, C. TRAVERS.*Decentralized Asynchronous Crash-Resilient Runtime Verification*, in "27th International Conference on Concurrency Theory (CONCUR)", Québec, Canada, 2016, https://hal.inria.fr/hal-01423646.

[28] P. FRAIGNIAUD, P. CRESCENZI, Z. LOTKER, P. PENNA.*Core-periphery clustering and collaboration networks*, in "IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)", San Francisco, United States, 2016, https://hal.inria.fr/hal-01423645.

[29] P. FRAIGNIAUD, L. FEUILLOLEY, J. HIRVONEN.*A Hierarchy of Local Decision*, in "43rd International Colloquium on Automata, Languages, and Programming (ICALP)", Roma, Italy, 2016, https://hal.inria.fr/hal-01423644.

[30] P. FRAIGNIAUD, M. HEINRICH, A. KOSOWSKI.*Local Conflict Coloring*, in "FOCS 2016 - 57th Annual IEEE Symposium on Foundations of Computer Science", New Brunswick, NJ, United States, October 2016, https://hal.inria.fr/hal-01223494.

[31] P. FRAIGNIAUD, A. KORMAN, R. YOAV.*Parallel exhaustive search without coordination*, in "48th ACM Symposium on Theory of Computing (STOC)", Cambridge, MA, USA, France, 2016, https://hal.inria.fr/hal-01423634.

[32] P. FRAIGNIAUD, E. NATALE.*Noisy Rumor Spreading and Plurality Consensus*, in "ACM Symposium on Principles of Distributed Computing (PODC)", Chicago, United States, 2016, https://hal.inria.fr/hal-01423686.

[33] P. FRAIGNIAUD, S. RAJSBAUM, C. TRAVERS.*Minimizing the Number of Opinions for Fault-Tolerant Distributed Decision Using Well-Quasi Orderings*, in "12th Latin American Symposium on Theoretical Informatics (LATIN)", Ensenada, Mexico, 2016, https://hal.inria.fr/hal-01423679.

[34] P. FRAIGNIAUD, S. RAJSBAUM, C. TRAVERS, P. KUZNETSOV, T. RIEUTORD.*Perfect Failure Detection with Very Few Bits*, in "18th International Symposium on Stabilization, Safety, and Security of Distributed Systems (SSS)", Lyon, France, 2016, https://hal.inria.fr/hal-01423637.

[35] P. FRAIGNIAUD, I. RAPAPORT, V. SALO, I. TODINCA.*Distributed Testing of Excluded Subgraphs*, in "30th International Symposium on Distributed Computing (DISC 2016)", Paris, France, C. GAVOILLE, D. ILCINKAS (editors), Lecture Notes in Computer Science, Springer, September 2016, vol. LNCS 9888, p. 342 - 356 [*DOI :* 10.1007/978-3-662-53426-7_25], https://hal.inria.fr/hal-01423633.

[36] P. GAWRYCHOWSKI, A. KOSOWSKI, P. UZNANSKI.*Sublinear-Space Distance Labeling Using Hubs*, in "DISC 2016 - 30th International Symposium on DIStributed Computing", Paris, France, C. GAVOILLE, D. ILCINKAS (editors), Distributed Computing, Springer, September 2016, vol. 9888, p. 230-242 [*DOI :* 10.1007/978-3-662-53426-7_17], https://hal.inria.fr/hal-01415064.

[37] R. KLASING, A. KOSOWSKI, D. PAJĄK.*Setting Ports in an Anonymous Network: How to Reduce the Level of Symmetry?*, in "SIROCCO 2016 - 23rd International Colloquium on Structural Information and Communication Complexity", Helsinki, Finland, J. SUOMELA (editor), Structural Information and Communication Complexity, Springer, July 2016, vol. 9988, p. 35-48 [*DOI :* 10.1007/978-3-319-48314-6_3], https://hal.inria.fr/hal-01415079.

[38] A. KOSOWSKI, L. VIENNOT.*Beyond Highway Dimension: Small Distance Labels Using Tree Skeletons*, in "SODA 2017 - 28th ACM-SIAM Symposium on Discrete Algorithms", Barcelona, Spain, January 2017, https://hal.inria.fr/hal-01359084.

Conferences without Proceedings

[39] P. GAWRYCHOWSKI, A. KOSOWSKI, P. UZNANSKI.*Brief Announcement: Sublinear-Space Distance Labeling Using Hubs*, in "PODC 2016 - 35th ACM Symposium on Principles of Distributed Computing", Chicago, IL, United States, July 2016, p. 43-45 [*DOI :* 10.1145/2933057.2933077], https://hal.inria.fr/hal-01415072.

Scientific Books (or Scientific Book chapters)

[40] C. DELPORTE-GALLET, P. A. ABDULLA (editors). *Networked Systems - 4th International Conference, NETYS 2016, Marrakech, Morocco, May 18-20, 2016, Revised Selected Papers*, LNCS, May 2016, vol. 9944, https://hal.inria.fr/hal-01416966.

[41] P. FRAIGNIAUD.*Locality in Distributed Graph Algorithms*, in "Encyclopedia of Algorithms", Springer, 2016, p. 1143-1148 [*DOI :* 10.1007/978-1-4939-2864-4_608], https://hal.inria.fr/hal-01423632.

Research Reports

[42] Y. BOUFKHAD, R. DE LA PAZ, L. LINGUAGLOSSA, F. MATHIEU, D. PERINO, L. VIENNOT.*Forwarding Tables Verification through Representative Header Sets*, Inria Paris Rocquencourt, January 2016, https://hal.inria.fr/hal-01262130.

[43] P. PENNA, L. VIENNOT.*Independent lazy better-response dynamics on network games*, ETH Zurich ; Inria ; Universite Paris Diderot-Paris VII, September 2016, https://hal.inria.fr/hal-01373411.

Scientific Popularization

[44] D. G. CORNEIL, M. HABIB.*Unified View of Graph Searching and LDFS-Based Certifying Algorithms*, in "Encyclopedia of Algorithms", Springer-verlag, 2016 [*DOI :* 10.1007/978-3-642-27848-8_685-1], https://hal.inria.fr/hal-01255083.

Other Publications

[45] A. CASTAÑEDA, C. DELPORTE, H. FAUCONNIER, S. RAJSBAUM, M. RAYNAL.*Wait-freedom and Locality are not Incompatible (with Distributed Ring Coloring as an Example)*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01265958.

[46] C. DELPORTE, H. FAUCONNIER, S. RAJSBAUM, M. RAYNAL.*Implementing Snapshot Objects on Top of Crash-Prone Asynchronous Message-Passing Systems*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01313584.

[47] C. DELPORTE, H. FAUCONNIER, S. RAJSBAUM, M. RAYNAL.*t-Resilient Immediate Snapshot is Impossible*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01315342.

[48] J. DUSART, M. HABIB, D. G. CORNEIL.*Maximal cliques structure for cocomparability graphs and applications*, November 2016, Il s'agit d'une recherche sur les relations entre les graphes d'intervalles et les graphes de cocomparabilité, https://hal.inria.fr/hal-01393927.

[49] L. FEUILLOLEY, P. FRAIGNIAUD.*Survey of Distributed Decision*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01331880.

[50] F. VÖLKEL, E. BAPTESTE, M. HABIB, P. LOPEZ, C. VIGLIOTTI.*Read networks and k-laminar graphs*, March 2016, working paper or preprint, https://hal.inria.fr/hal-01282715.

Project-Team MAMBA

Modelling and Analysis for Medical and Biological Applications

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

IN PARTNERSHIP WITH:
CNRS

Université Pierre et Marie Curie (Paris 6)

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:

3. - Data and knowledge
3.1. - Data
3.1.1. - Modeling, representation
3.4. - Machine learning and statistics
3.4.6. - Neural networks
3.4.7. - Kernel methods
6. - Modeling, simulation and control
6.1. - Mathematical Modeling
6.1.1. - Continuous Modeling (PDE, ODE)
6.1.2. - Stochastic Modeling (SPDE, SDE)
6.1.3. - Discrete Modeling (multi-agent, people centered)
6.1.4. - Multiscale modeling
6.1.5. - Multiphysics modeling
6.2. - Scientific Computing, Numerical Analysis & Optimization
6.2.1. - Numerical analysis of PDE and ODE
6.2.2. - Numerical probability
6.2.3. - Probabilistic methods
6.2.4. - Statistical methods
6.2.6. - Optimization
6.3. - Computation-data interaction
6.3.1. - Inverse problems
6.3.2. - Data assimilation
6.4. - Automatic control
6.4.1. - Deterministic control

Other Research Topics and Application Domains:

1. - Life sciences
1.1. - Biology
1.1.2. - Molecular biology
1.1.3. - Cellular biology
1.1.7. - Immunology
1.1.8. - Evolutionnary biology
1.1.9. - Bioinformatics
1.1.10. - Mathematical biology
1.2. - Ecology
1.3. - Neuroscience and cognitive science
1.4. - Pathologies
2. - Health

2.2. - Physiology and diseases
2.2.3. - Cancer
2.2.4. - Infectious diseases, Virology
2.2.6. - Neurodegenerative diseases
2.3. - Epidemiology
2.4. - Therapies
2.4.1. - Pharmaco kinetics and dynamics
2.4.2. - Drug resistance
2.6.3. - Biological Imaging
9.5.4. - Management science

1. Members

Research Scientists

Marie Doumic [Team leader, Inria, Senior Researcher, HDR]
Luis Lopes Neves de Almeida [CNRS, Senior Researcher, HDR]
Pierre-Alexandre Bliman [Inria, Senior Researcher, HDR]
Jean Clairambault [Inria, Senior Researcher, HDR]
Dirk Drasdo [Inria, Senior Researcher, HDR]

Faculty Members

Alexander Lorz [Univ. Paris VI, Associate Professor]
Benoît Perthame [Univ. Paris VI, Professor]
Nicolas Vauchelet [Univ. Paris VI, Associate Professor, until August 2016; now Prof. Univ. Paris XIII, HDR]

Technical Staff

Tim Johann [Univ. Leipzig]
Paul Van Liedekerke [Inria]
Yi Yin [Inria]

PhD Students

Aurora Armiento [Inria]
Noémie Boissier [Inria]
Youssef Bourfia [UPMC, co-tutela with Cadi Ayyad University, Marrakesh]
Geraldine Cellière [until April 2016]
Sarah Eugène [Inria]
Adrian Friebel [Univ. Leipzig, until June 2016]
Ghassen Haddad [UPMC, co-tutela with Tunis University]
Shalla Hanson [Inria, co-tutela UPMC and Duke University]
Hugo Martin [Univ. Paris VI, from Jun 2016]
Mathieu Mézache [Inria, from May 2016]
Johannes Neitsch [Univ. Leipzig, until Nov 2016]
Camille Pouchol [Univ. Paris VI]
Antonin Prunet [Univ. Paris VI]
Andrada Quillas Maran [Univ. Paris VI]
Martin Strugarek [IPEF]
Cécile Taing [Univ. Paris VI]

Administrative Assistant

Nelly Maloisel [Inria]

Others

Bettina d'Avila Barros [Inria, Escola de Matemática Aplicada, Fundação Getulio Vargas, Rio de Janeiro, Sep 2016, Master student]
Andreas Buttenschoen [Inria, Univ. Edmonton, until May 2016, Intern]
Julie Favre [Inria, EPFL, from Oct 2016, Intern]
Hicham Janati [Inria, ENSAE, from Jun 2016 until Sep 2016, Intern]

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 two kinds of problems involving dynamics of Partial Differential Equations (PDEs).

Problems from life sciences (cell motion, early embryonic development, tissue growth and regeneration, cancer modelling, pharmacology,...) have been considered, and still constitute the core of MAMBA. Models for complex fluid flows (shallow water models, flows with a free surface) were studied until December 2012, when the scientists in charge of the "Géophysique" part left BANG to constitute the new Inria team ANGE (https://team.inria.fr/ange/), while the remaining ("Biophysique") part of the BANG team continue their research work within the new Inria team MAMBA, now headed by Marie Doumic.

The dynamics of complex physical or biophysical phenomena involving many agents, including proteins or cells - seen as *active* agents - can be represented efficiently either by explicitly considering the behaviour of each particle individually (e.g. through branching trees and piecewise deterministic Markov processes, or stochastic differential equations) or by Partial Differential Equations (PDEs) which, under certain hypotheses, represent local averages over a sufficiently large number of agents.

Biology and medicine currently face the difficulty to make sense out of data newly available by means of recent signal acquisition methods. Modelling through agent-based or continuous models is a unique way to explain (i.e., model) the observations and then compute, control and predict. These are the goals of MAMBA.

3. Research Program

3.1. Introduction

At small spatial scales, or at spatial scales of individual matter components, where heterogeneities in the medium occur, agent-based models are developed ([0], [76], Dirk Drasdo's former associate team QUANTISS). Another approach, that is considered in the project-team MAMBA consists in considering gene expression at the individual level by stochastic processes [0], by ordinary differential equations [0], or by a mixed representation of Markov processes and ordinary differential equations [0], the outputs of which quantify focused aspects of biological variability in a population of individuals (cells) under study.

Both these approaches complement the partial differential equation models considered on scales at which averages over the individual components behave sufficiently smoothly. Investigating the links between these models through scales is also part of our research [0]. Moreover, in order to quantitatively assess the adequacy between the biological phenomena we study and the mathematical models we use, we also develop inverse problem methods.

[0]Drasdo, Hoehme, Block, *J. Stat. Phys.*, 2007

[0]as in M. Sturrock et al., spatial stochastic modelling of the Hes1 gene regulatory network: intrinsic noise can explain heterogeneity in embryonic stem cell differentiation, *Journal of The Royal Society Interface*, 2013

[0]as in A. Friedman et al, Asymptotic limit in a cell differentiation model with consideration of transcription, *J. Diff. Eq.*, 2012

[0]as in R. Yvinec et al., Adiabatic reduction of stochastic gene expression with jump Markov processes, *J. Math. Biol.*, 2013.

[0]H. Byrne and D. Drasdo, Individual-based and continuum models of growing cell populations: a comparison, *J. Math. Biol*, 2009

3.2. PDE analysis and simulation

PDEs arise at several levels of our models. Parabolic equations [0] can be used for large cell populations and also for intracellular spatio-temporal dynamics of proteins and their messenger RNAs in gene regulatory networks, transport equations [0] are used for protein aggregation / fragmentation models and for the cell division cycle in age-structured models of proliferating cell populations. Existence, uniqueness and asymptotic behaviour of solutions have been studied [65], [62]. Other equations, of the integro-differential type, dedicated to describing the Darwinian evolution of a cell population according to a phenotypic trait, allowing exchanges with the environment, genetic mutations and reversible epigenetic modifications, are also used [81], [80], [79], [82], possibly enriched to classical PDEs by the adjunction of diffusion and advection terms [63]. Through multiscale analysis, they can be related to stochastic and free boundary models used in cancer modelling.

3.3. Inverse problems

When studying biological populations (usually cells or big molecules) using PDE models, identification of the functions and parameters that govern the dynamics of a model may be achieved to a certain extent by statistics performed on individuals to reconstruct the probability distribution of their relevant characteristics in the population they constitute, but quantitative observations at the individual level (e.g., fluorescence in single cells [60] or size/age tracking [87]) require sophisticated techniques and are most often difficult to obtain. Relying on the accuracy of a PDE model to describe the population dynamics, inverse problem methods offer a tractable alternative in model identification, and they are presently an active theme of research in MAMBA. Following previous studies [68], [69], some combining statistical and deterministic approaches [67] with application to raw experimental data [66], we plan to develop our methods to new structured-population models (or stochastic fragmentation processes as in [66]), useful for other types of data or populations (e.g. size/age tracking, polymer length distribution, fluorescence in single cells).

3.4. Stochastic and agent-based models

The link between stochastic processes and kinetic equations is a domain already present in our research [0] [67] and that we plan to develop further. They can be viewed either as complementary approaches, useful to take into account different scales (smaller scales for stochastic models, larger scales for mean-field limits), or even as two different viewpoints on the same problem [66], enriching each other. Neuroscience is a domain where this is particularly true because noise contributes significantly to the activity of neurons; this is the case of networks where mean field limits are derived from stochastic individual-based models and lead to fundamental questions on the well-posedness and behaviours of the system [0]. One strength and originality of our project is our close connection and collaboration not only with probability theorists but also with statisticians, who provide us with efficient help in the identification of our model parameters.

Agent-based systems consider each component individually. For example, in multi-cellular system modelling, the basic unit is the cell, and each cell is considered [70], [89]. This approach has advantages if the population of cells reveals heterogeneities on small spatial scales as it occurs if organ architecture is represented [76], or if the number of cells in a particular state is small. Different approaches have been used to model cellular agents in multi-cellular systems in space, roughly divided in lattice models (e.g. [85]) and in lattice-free (or off-lattice) models, in which the position [70], [73] or even the shape (e.g. [89]) of the cell can change gradually.

[0] B. Perthame, Parabolic equations in biology, Springer, 2015
[0] B. Perthame, Transport equations in biology, Springer, 2007
[0] H. Byrne and D. Drasdo, Individual-based and continuum models of growing cell populations: a comparison, *J. Math. Biol*, 2009
[0] Cáceres, Carrillo, Perthame *J. Math. Neurosci.* 2011; Pakdaman, Perthame, Salort *Nonlinearity* 2010

The dynamics of cells in lattice-based models is usually described by rules chosen to mimic the behaviour of a cell including its physical behavior. The advantage of this approach is that it is simpler and that simulation times for a given number of cells are shorter than in lattice-free models. In contrast, most lattice-free models attempt to parameterise cells by measurable values with a direct physical or biological meaning, hence allowing identification of physiologically meaningful parameter ranges. This improves model simulation feasibility, since parameter sensitivity analyses in simulations shows significant improvements when a high dimensional parameter space can be reduced. It also facilitates the development of systematic systems biology and systems medicine strategies to identify mechanisms underlying complex tissue organisation processes ([89], [71]).
Moreover, it is straightforward to include relevant signal transduction and metabolic pathways in each cell within the framework of agent-based models, which is a key advantage in the present times, as the interplay of components at many levels is more and more precisely studied [91].

3.5. Multi-level modelling

Multi-level modelling addresses models spanning many spatial scales composed of functional connected modules on each of these scales [64]. Typical representatives of multilevel systems are organs, that are composed of cells of different types coordinated in space, extracellular matrix, etc. Development, parameterisation, verification and validation of such models is challenging as it is usually not possible to simultaneously perform experimental measurements on each level simultaneously.

The fundamental strategy is composed of a multi-step strategy, parameterising sub-models individually before connecting them [71]. For this, models shall be parameterised by measurable quantities for which parameter ranges can be reliably estimated. Then simulated parameter sensitivity simulations are run, comparing results with experiments. If the best agreement between model and experiment is insufficient, the model is wrong or incomplete. If several models are able to explain the data, settings should be run with these models that lead to experimentally testable distinguishable outcomes.

4. Application Domains

4.1. Cancer modelling

Evolution of healthy or cancer cell populations under environmental pressure; drug resistance. Considering cancer as an *evolutionary disease*, evolution meaning here Darwinian evolution, but also Lamarckian instruction, of populations structured according to phenotypes relevant to describe their heterogeneity at stake in studies led in collaboration with our biologist partners within the Institut Universitaire de Cancérologie (IUC) of UPMC, we tackle the problem of understanding and limiting: a) the evolution from pre-malignancy to malignancy in cell populations (in particular we study early leukaemogenesis, leading to acute myeloid leukaemia), and b) in established cancer cell populations, the evolution towards drug-induced drug resistance. The environmental pressure guiding evolution has many sources, including signalling molecules induced by the peritumoral stroma (e.g., between a breast tumour and its adipocytic stroma), and anticancer drugs and their effects on both the tumour and its stromal environment. The models we use [63], [79], [80], [81] are close to models used in ecology for adaptive dynamics.

Drugs: pharmacokinetics-pharmacodynamics, therapy optimisation. We focus on multi-drug multi-targeted anticancer therapies aiming at finding combinations of drugs that theoretically minimise cancer cell population growth with the constraint of limiting unwanted toxic side effects under an absolute threshold (this is not L^2 nor L^1, but L^∞ optimisation, i.e. the constraints as well as the objective function are L^∞) in healthy cell populations and avoiding the emergence of resistant cell clones in cancer cell populations [59], [80], [60], [79]. Prior to using optimisation methods, we design models of the targeted cell populations (healthy and tumour, including molecular or functional drug targets [58]) by PDEs or agent-based models [56], and molecular pharmacological (pharmacokinetic-pharmacodynamic, PK-PD) models of the fate and effects in the organism of the drugs used, usually by ODE models. A special aspect of such modelling is the representation of multi-cellular spatio-temporal patterns emerging from therapies.

Multi-scale modelling of cancer invasion. The major step from a benign tumour to an invasive cancer is the development step at which cells detach from the tumour mass and invade individually the surrounding tissue [0]. We performed *in vitro* simulations of cancer cell invasion for breast cancer evaluating under which conditions the observed migration pattern occurs. (In collaboration with our biologist partners within the Institut Curie)

4.2. Modelling and control in epidemiology

The spread of certain strains of the intracellular parasitic bacterium *Wolbachia* in populations of mosquitoes *Aedes aegypti* drastically reduces their competence as vector of dengue and other severe mosquito-borne viral diseases known as arboviral infections (chikungunya, Zika, yellow fever...). In absence of vaccine, or of preventive or curative treatment, the release of mosquitoes infected by this bacterium has been recently considered a promising tool to control these diseases.

Technically the situation can be described by a bistable model, and the issue consists in moving from a Wolbachia-free equilibrium to a fully contaminated equilibrium. Therefore mathematical modeling is of great interest for the study of the feasibility of the control of dengue fever using this strategy.

Key questions about this method concern the efficacy of the strategies used to release Wolbachia-infected mosquitoes in the field that can be applied successfully and with limited cost.

4.3. Protein polymerisation

Self-assembly of proteins into amyloid aggregates is an important biological phenomenon associated with various human neurodegenerative diseases such as Alzheimer's, Parkinson's, Prion (in particular variant Creutzfeldt-Jakob disease, epidemically linked to bovine spongiform encephalopathy, or so-called "mad cow", disease), Huntington's disease. Amyloid fibrils also have potential applications in nano-engineering of biomaterials.

However, the mechanisms of polymerisation are far from being quantitatively understood by biologists. They can be modelled with the help of coagulation-fragmentation equations, a field of expertise of MAMBA [16], [36], or with stochastic models [20]. One difficulty of this application is that the reactions imply both very small and very large scales for the sizes of polymers [7], experimental data giving only access to the time evolution of size-averaged quantities [6]. Moreover, there exists an intrinsic variability among experiments, which has to be distinguished from a lack of reproducibility [20].

The European starting grant SKIPPER^{AD} involves a long-term collaboration with Human Rezaei's team, a biologist expert group in amyloid diseases at INRA Jouy-en-Josas. It allowed us to further develop new collaborations, in particular with Wei-Feng Xue's team in Canterbury, who is one of the rare biophysicists in this area who is able to measure not only size-averaged quantities, as for instance the time-evolution of the total polymerised mass, but also size distribution of polymers (at least over a certain threshold). Such measurements allow us to use much more powerful inverse problems and data assimilation methods [6].

Moreover, this field of applications to human neurogenerative diseases brings us new questions [17], which is a stimulation for our mathematical research and at the same time allows us to provide biologists with a new and efficient tool.

4.4. Cell motion

Several processes are employed by cells to communicate, regulate and control their movements, and generate collective motion. Among them, chemotaxis is the phenomenon by which cells direct their active motion in response to an external chemical (or physical) agent. In chemotaxis, cells not only respond but can also produce the chemical agent, leading to a feedback loop. Understanding this phenomenon is a major challenge for describing the collective behaviour of cells. Many mathematical models have been proposed at different scales, yielding a good description of cell aggregation. In particular, mathematical models at macroscopic scale may be derived departing from kinetic description at mesoscopic scale. An interesting study at the

[0] Weinberg, The biology of cancer, Garland, 2007

numerical level is to provide numerical schemes able to treat both scales. Then in [27], we have proposed an asymptotic preserving scheme for a model describing the formation of networks of cells in tissues. In collaboration with biophysicists at Institut Curie in Paris, we develop and study [0] mathematical models based on kinetic equations for bacterial travelling waves in a microchannel. These models have shown a remarkable quantitative agreement with experimental observations. In [18], we extend this approach to study the behavior of the interaction between two populations of E. Coli. We show that in certain cases populations that travel with its own speed in the channel when separated, may synchronise their movements when put together.

Cell motion arises also in the growth of solid tumours, which can be described through cell population models or multiphase flows [0]. This is a very active subject because several bio-chemico-physical mechanisms are at work; for instance motion can arise from pressure forces resulting from cell divisions and from active cell motility. At the smaller scale stochastic agent-based models of tumour cells invading the tumour environment or blood vessels are considered [0], and allow to represent detailed behaviours and interactions. At a larger scale, free boundary problems are widely used, e.g., for image-based prediction because of the reduced number of parameters [0]. Asymptotic analysis makes a link between these different mechanistic models [88]. One other setting where we will study cell motion is 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 [86], [90] 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 settings, indicating the existence of conserved mechanisms that should be widespread across living beings. In the 01365414 thesis of Telmo Pereira, some differences between the two settings are also studied.

4.5. Physics of tissue organisation

Many new insights in the last years indicate that migration, growth and division of cells are largely impacted by cell and tissue mechanics [0], [0], [0]. Centre-based growth models already account for many of the observed phenomena [0], [0]. They furthermore allow calculation of the stress tensor in the tissue. A critical shortcoming of centre-based models is that forces between cells are calculated based on pairwise interactions hence multi-cellular interactions leading to true cell compression cannot be taken into account.

In order to scope with this shortcoming we (1.) developed a strategy in which forces are calibrated with a high resolution agent based model (so called deformable cell model), so that stress in tissue can then be calculated also at high cell density [54]; (2.) integrated cell division in deformable cell models to permit direct simulations of phenomena with this model type; (3.) developed hybrid models permitting to simulate centre-based and deformable cell models in the same simulations to be able to reach sufficiently high cell numbers.

Deformable cell models [0] resolve cell surface at reasonable resolution, and allow to calculate cell deformation as function of stress emerging in the tissue, hence the stress tensor cannot only be resolved at the position of the cell centre, as in the case of centre-based models, but in this case at any point on the cell surface or inside the cell. The higher resolution causes much longer simulation times which is why currently simulation of large multi-cellular systems with deformable cell models on standard computers is not feasible.

---

[0] N. Bournaveas, V. Calvez, S. Gutiérrez and B. Perthame, Global existence for a kinetic model of chemotaxis via dispersion and Strichartz estimates, *Comm. PDE*, 2008

[0] J. Ranft et al, Fluidisation of tissues by cell division and apoptosis, *PNAS*, 2010 and L. Preziosi and A. Tosin, Multiphase modelling of tumour growth and extracellular matrix interaction: mathematical tools and applications, *J. Math. Biol.*, 2009.

[0] I. Ramis-Conde et al., J. Phys. Biol., 2009

[0] Works by O. Saut, T. Colin, A. Iollo, N. Ayache, J. Lowengrub

[0] Ingber, Proc. Natl. Acad. Sci (USA), 2005

[0] Trepat et. al., Nat. Phys. 2009

[0] Alessandri et. al., Proc. Natl. Acad. Sci. (USA) 2013

[0] Drasdo and Hoehme, Phys. Biol. 2005

[0] Drasdo and Hoehme, New Journal of Physics 2012

[0] Odenthal, Smeets, van Liedekerke, et. al., PloS Comput Biol. 2013

4.6. Liver modelling

Liver is the main detoxifying organ of the human body and can regenerate up to about 70% of its mass. It performs its task by using a complex tissue architecture, with hepatocytes aligning along micro-capillaries and forming a dense network. The incidence rate of liver diseases is steadily increasing, liver cancer ranking 6th among all cancers. About one person in 12, otherwise said 500 million people worldwide, will suffer from viral hepatitis. Hepatitis B and C as well as misuse of drugs or alcohol are major causes of liver cancer. Notwithstanding the importance of this public health problem, disease pathogenesis and regeneration in liver are still not well understood.

So far systems biology approaches addressing the tissue scale are rare. Most of those which do so base on compartment models (e.g. [0]); only recently are approaches addressing the tissue scale being developed [76] [0], [0], [0], [0]. We have developed a multi-scale model of liver regeneration representing the tissue architecture, the different cell types, the flow systems, hepatocyte metabolism and signal transduction controlling cell cycle entrance in the regeneration processes, taking into account extrahepatic compartments when relevant. Applications are regeneration after drug-induced damage and after partial hepatectomy, drug pharmacodynamics and pharmacokinetics in liver and liver cancer, and model-based prediction of in-vivo drug toxicity from in-vitro measurements [0].

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Personnel

Marie Doumic has moved in September 2015 for a 1-year sabbatical to the Wolfgang Pauli Institute in Vienna. Stefan Hoehme left in July 2015 to start a prestigious “Emmy Noether” junior research group at University of Leipzig, faculty for computer sciences. Of note, this is the first Emmy Noether research group in Leipzig, and he was the only one accepted this year (out of 20 presented).

Nicolas Vauchelet left the team in September 2015, becoming a full professor at University Paris XIII.

5.1.2. THE ITMO Cancer national call.

The team has been successful in simultaneously participating in 2 different funded projects of the ITMO Cancer THE (“Tumour Heterogeneity in its Ecosystem”, a programme managed by INSERM) national call for 2016: one, EcoAML (4 teams), on early leukaemogenesis in Acute Myelogenous Leukaemia (AML), headed by François Delhommeau (CDR St Antoine, Paris), with whom we have a long-lasting collaboration, and the other, MoGlImaging (8 teams), on treatment-induced treatment resistance and heterogeneity in glioblastoma, headed by Elizabeth Moyal (INSERM, Toulouse), a project inside which we have recently developed a work collaboration with the team of François Vallette (INSERM, Nantes) on the in-vitro resistance of glioblastoma to temozolomide. In both these collaborative projects, begun in November 2016 and to be integrated in 2017 in the future THE consortium (gathering the 6 projects laureates to the national call), we propose to develop our phenotype-structured models for both the cancer and the supporting stromal cell populations, with representation of mutualistic interactions between them.

[0] Diaz-Ochoa et. al. Frontiers in Pharmacology, 2013
[0] Ricken, Dahmen, Dirsch, Biomech. Model. Mechanobiol. 2010
[0] Debbaut et. al., J. Biomech. Eng. 2014
[0] Siggers, Leungchavphongse, Ho, Repetto, Biomech. Model. Mechanobiol. 2014
[0] Schwen et. al., PLoS Comput. Biol. 2014
[0] Godoy et al., *Arch Toxicol. 2013* Aug;87(8):1315-1530

6. New Software and Platforms

6.1. TiQuant

Tissue Quantifier
KEYWORDS: Systems Biology - Bioinformatics - Biology - Physiology
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 recognised. 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 [72]. It is implemented in portable object-oriented ANSI C++. The GUI is based on QT and supports real-time visualisation using OpenGL. TiQuant is embedded in the tissue modelling framework CellSys and thus is tightly linked with TiSim, a versatile and efficient simulation environment for tissue models. TiQuant provides an interface to VolView and further complements its functionality by linking to the open-source libraries ITK and VTK (itk/vtk.org). The image/volume processing chains currently implemented in TiQuant for example include techniques to segment conduit and cell segmentation from 3D confocal micrographs of liver tissue based on the Adaptive Otsu Thresholding method and a number of morphological operators [75]. TiQuant is currently extended by a machine learning component.

FUNCTIONAL DESCRIPTION

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.

- Contact: Dirk Drasdo
- URL: http://www.msysbio.com

6.2. TiSim

Tissue Simulator
KEYWORDS: Systems Biology - Bioinformatics - Biology - Physiology
FUNCTIONAL DESCRIPTION

TiSim (Tissue Simulator) is a software for agent-based models of multicellular systems. It permits model development with centre-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. A non-interactive mode to use the software also exists, accepting a combination of XML and HDF5 (hierarchical data format v5) as input, which produces output data in VTP (VTK) and HDF5 format. SBML, SBML_ODESolver and sundials are deployed for the creation and solution of the differential equations of metabolic networks and signalling pathways presented in SBML data format. TiSim permits agent-based simulations of multicellular systems and can be directly fed by processed image data from TiQuant.

- Contact: Dirk Drasdo
- URL: (No url yet)

7. New Results

7.1. Cancer

Participants: Luis Lopes Neves de Almeida, Group Emmanuel Barillot [Institut Curie], Catherine Bonnet [DISCO team, Saclay], Thibault Bourgeron, Group Kai Breuhahn [Hospital of University of Heidelberg, Pathology], Rebecca Chisholm, Jean Clairambault, François Delhommeau [Haematology department, St Antoine Hospital, Paris], Marie Doumic, Dirk Drasdo, Walid Djema [DISCO team, Saclay], Julie Favre [EPFL, Lausanne], Olivier Fercoq [Télécom ParisTech], Ghassen Haddad [ENIT, Tunis], Shalla Hanson [Department of mathematics, Duke University, Durham, NC], Pierre Hirsch [Haematology department, St Antoine Hospital, Paris], Groups Invade, Lungsysii, Hicham Janati [ENSAE, Paris], Tim Johann, Group Klingmueller [German Cancer Center, Heidelberg], Michal Kowalczyk [Univ. Santiago de Chile], Annette Larsen [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Tommaso Lorenzi [University of St Andrews, Scotland], Alexander Lorz, Frédéric Mazenc [DISCO team, Saclay], Benoît Perthame, Camille Pouchol, Andrada Quillas Maran, Fernando Quirós [Univ. Autónoma de Madrid], Michèle Sabbah [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Min Tang [Jiaotong University, Shanghai], Teresa Teixeira [IBCP], Emmanuel Trélat [LJLL, UPMC], Paul Van Liedekerke, Oliver Sedlaczek [German Cancer Center (DKFZ) and Hospital of University of Heidelberg, Germany], François Vallette [INSERM, CRCNA, Nantes], Nicolas Vauchelet, Irène Vignon-Clementel [REO], Zhou Xu [IBCP], Yi Yin.

7.1.1. Senescence and telomere shortening

In many animals, aging tissues accumulate senescent cells, a process which underlies the loss of regeneration capacity of organs and is ultimately detrimental to the organism. Senescence is also required to protect organisms from unlimited proliferation that may arise from numerous stimuli or deregulations. Due to these opposing effects in aging and cancer, senescence is considered antagonistic pleiotropic; senescence is beneficial to protect from cancer in the young organism, but becomes detrimental late in life. Therefore, understanding the mechanisms of cellular senescence may lead to the development of global therapies to debilities specific for the aged, as well age-associated diseases and cancer. These are major public health issues in France, and other western aging countries.

Replicative senescence, induced by telomere shortening, exhibits considerable asynchrony and heterogeneity, the origins of which remain unclear. In [19], following [61], we formally study how telomere shortening mechanisms impact on senescence kinetics and define two regimes of senescence, depending on the initial telomere length variance. We provide analytical solutions to the model, highlighting a non-linear relationship between senescence onset and initial telomere length distribution. This study reveals the complexity of the collective behaviour of telomeres as they shorten, leading to senescence heterogeneity.

7.1.2. Stability analysis of a delay differential model of healthy and leukaemic haematopoiesis

The collaboration with the DISCO team (Inria Saclay, C. Bonnet, F. Mazenc and their PhD student W. Djema), supported by the collaboration with the team of haematologists led by F. Delhommeau at St. Antoine Hospital in Paris, has been continued, with common research work underway. A new model describing the coexistence between ordinary and mutated haematopoietic stem cells was introduced and analysed in [32]. Interpreting theoretical conditions found to guarantee the survival of healthy cells while eradicating unhealthy ones leads us to propose possibly innovative therapies obtained by combining the infusion of different drugs (Flt-3 inhibitors such as quizartinib, cytosine arabinoside, anthracyclines).

7.1.3. Interactions between tumour cell populations and their cellular micro-environments

This is the main object of study, together with the consideration of phenotype and genotype heterogeneity in cancer cell populations (see *Highlights of the Year*), of the *THE ITMO Cancer* national call 2016, to which two (out of three) of the submitted projects involving our team, which themselves were two out of the six laureate projects at the national level, have been successfully funded. The two projects, EcoAML and MoGlImaging have been launched in November 2016.

7.1.4. Evolution and cancer; drug resistance in cancer cell populations

We have continued to develop our phenotypically based models of drug-induced drug resistance in cancer cell populations, representing their Darwinian or Lamarckian evolution under drug pressure by integro-differential equations. The properties of phenotype-structured PDEs are explored in theoretical articles with examples [25], [78]. We will also use them in the two projects laureates to the THE (*Tumour Heterogeneity in its Ecosystem*) ITMO Cancer call of 2016 (see *Highlights of the Year*), EcoAML and MoGlImaging, to help predict early evolution towards leukaemogenesis (EcoAML, leader F. Delhommeau, St. Antoine Hospital, Paris) and emergence of resistance to temozolomide in glioblastoma cell populations (MoGlImaging, leader E. Moyal, Toulouse, F. Vallette, Nantes, being our main work correspondent). In this version, mutualistic exchanges between the cancer cell population and its supporting stroma will be represented as impinging on the phenotypic variables that describe the relevant heterogeneity at stake in the two cell populations.

With F. Vallette, we have co-supervised Hicham Janati's ENSAE 2nd year (M1) internship on the investigation of cancer resistance in a Glioblastoma cell line [46] with gene expression data coming from F. Vallette's lab in Nantes. This internship represents for us a first step in the quest for relevant (most likely multidimensional) phenotypes, based on bioinformatic and biostatistic methods to process experimental dynamic gene expression data, to interactively identify our physiologically structured models of heterogeneity and its evolution in cancer cell populations. The task ahead is immense, but our commitment in the THE consortium (see *Highlights of the Year*) with biologists providing us with such data (F. Vallette, F. Delhommeau) gives us good expectations to be successful with it in a close future. Following Hicham Janati's internship [46], Julie Favre (M1 student at EPFL) has been hired in a new internship to set the practical grounds for the interactive collaboration (begun with the THE program in November 2016) between our team and F. Delhommeau's team on model-based processing of gene expression data produced by a heterogeneous leukaemic cell population and by its surrounding stromal cell population.

The evolution towards drug-induced drug resistance in cancer cell populations may be described by methods of adaptive dynamics for continuous phenotype-structured populations, as such cell populations are fundamentally phenotypically, if not genetically, heterogeneous. In [11], [40], we review the bases of heterogeneity and drug resistance in cancer, its assessment by biological experiments and by mathematical modelling and methods of optimal control that may be applied to represent and optimise combined delivery of cytotoxic and cytostatic drugs, see below "Optimal control and drug resistance" [52].

7.1.5. Therapy optimisation

PK-PD: optimisation with respect to unwanted side effects. A previous pharmacokinetics-pharmacodynamics (PK-PD) model for the action of anticancer drugs at the molecular level, coupled with an age-structured linear model of the cell division cycle, has been updated in [12] (introduced in a special issue on PK-PD [9]) with optimisation of the combined delivery of 3 different drugs (5-fluorouracil, oxaliplatin, leucovorin). This is joint work with Olivier Fercoq, Télécom ParisTech. It represents the coalescence of two distinct types of models, both studied in previous years in our team: molecular ODE-based models of the action of anticancer drugs, and optimisation (using a Uzawa-like algorithm applied to the first eigenvalues of the two growing populations, minimising the cancer eigenvalue - objective - while maintaining the healthy eigenvalue above a reference threshold - constraint -, supposed to be linked to the state of health of the patient) of the control of linear growth models based on age-structured transport equations for the cell division cycle in the two populations separately.

Optimal control and drug resistance. In the framework of Camille Pouchol's PhD thesis, co-supervised at LJLL by E. Trélat and J. Clairambault, analysing the behaviour of healthy and cancer cell populations structured in a continuous resistance phenotype to a cytotoxic drug, and exposed to cytostatic and cytotoxic chemotherapies, we have firstly established, in an asymptotic analysis using a Lyapunov functional inspired from works by P.-E. Jabin and G. Raoul [77], results of convergence and concentration for constant drug concentrations [52] (following [84]). In a second part of this work, we have derived from them analytical conditions of optimality for the delivery of the drugs in a general class of controls. A numerical example of the optimal strategy is illustrated on Figure 1, where the phenotype x continuously ranges from totally sensitive ($x = 0$) to totally resistant ($x = 1$), and healthy and cancer cells are represented by densities of cells $n_H(t,x)$, $n_C(t,x)$. The simulations confirm that the optimal strategy consists of letting the cancer cell population become more and more homogeneous around a sensitive phenotype, and then to use the maximal amount of drugs. This proposed strategy may be related with the "drug holiday" practiced in the clinic of cancers. We also show *en passant* the clearly detrimental effect of delivering cytotoxic drugs at high *constant* doses, as they inevitably induce the emergence of a thriving resistant subpopulation, which is illustrated on Figure 2.

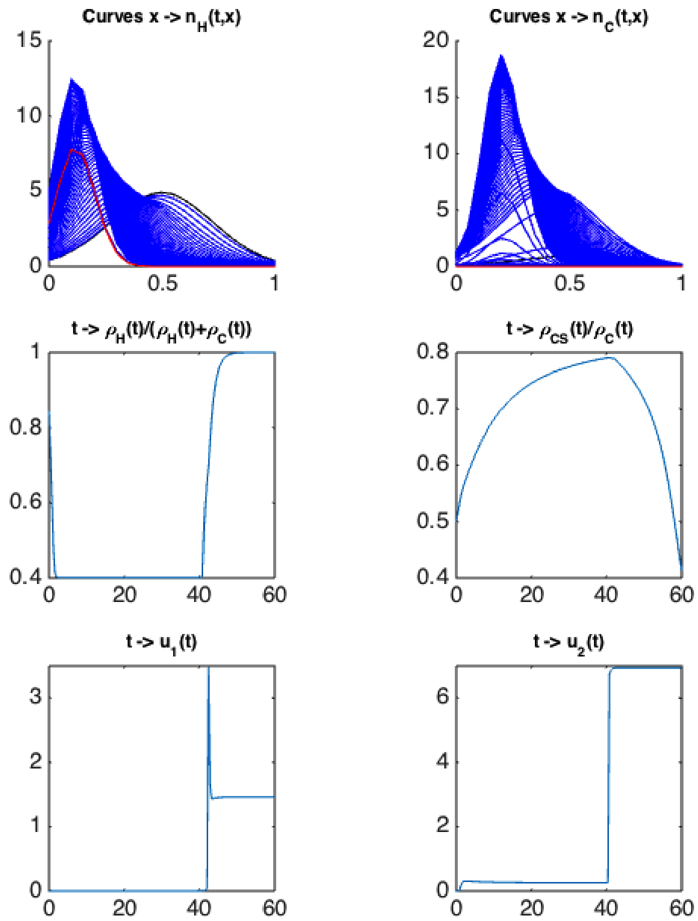

Figure 1. Simulation of the optimal control problem in time horizon $T = 60$. Top, left and middle: time evolution of $x \mapsto n_H(t,x)$, number of healthy cells with drug resistance expression phenotype x, and of $x \mapsto n_C(t,x)$, number of cancer cells with the same phenotype x. The initial conditions are in black, the final ones in red. Top right (resp., centre right): evolution with time of the total number of cancer cells $\rho_C(t) = \int_0^1 n_C(t,x)dx$ (resp., of healthy cells $\rho_H(t) = \int_0^1 n_H(t,x)dx$). Centre left (resp., centre middle), evolution with time of the ratio of healthy cells to total cells (resp., of sensitive cancer cells defined by the weighted integral $\rho_{CS}(t) = \int_0^1 (1-x)n_C(t,x)dx$ to the total cancer cell population). Bottom, left and middle: evolution with time of the optimal drug infusions of cytotoxic (u_1) and cytostatic (u_2) drugs. One can check on this simulation the quasi-bang-bang character of the optimal control.

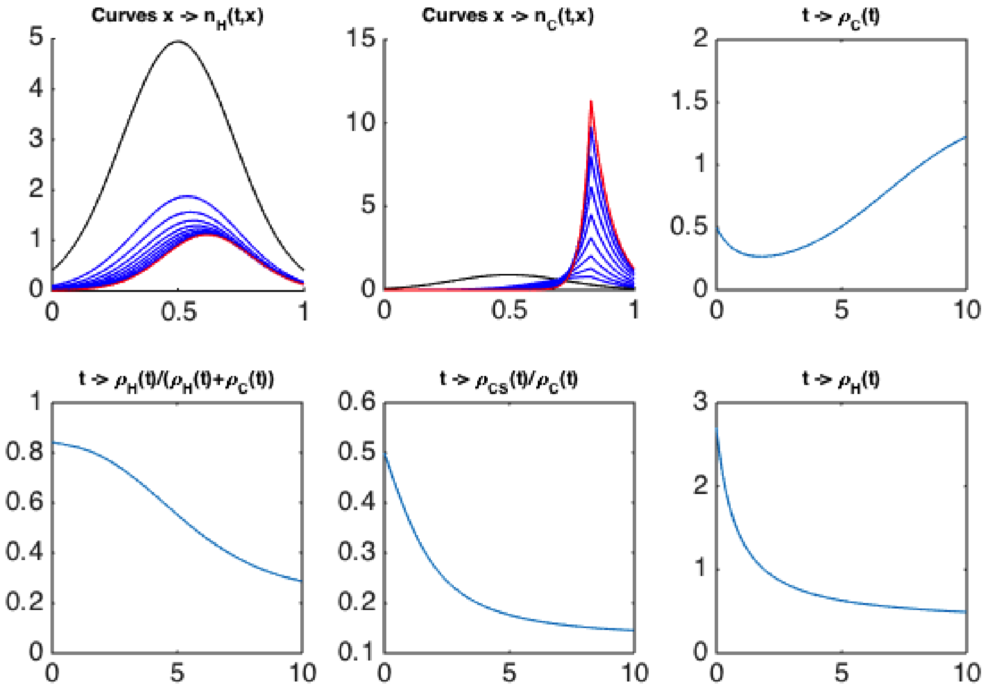

Figure 2. Comparative evolution with constant high drug doses. Catastrophic deleterious effects of the treatment on the concentration of the drug resistance phenotype x in the cancer cell population (top middle), and on the cell population numbers $\rho_C, \rho_H, \rho_{CS}$. Simulation with $u_1(t) = \mathrm{Cst} = 3.5$, $u_2(t) = \mathrm{Cst} = 2$, in time horizon $T = 10$.

7.1.6. Lung and breast cancer

Diffusion-weighted magnetic resonance imaging (DWI) is a key non-invasive imaging technique for cancer diagnosis and tumour treatment assessment, reflecting Brownian movement of water molecules in tissues. Since densely packed cells restrict molecule mobility, tumour tissues produce less attenuated DWI signals than 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 have developed an image processing and analysis chain, which was used to study the correlation between the diffusion coefficient (D value) estimated from DWI and tumour cellularity from serial histological slides of a resected non-small cell lung cancer (NSCLC) tumour. Colour deconvolution followed by cell nuclei segmentation was performed on digitised histological images to determine local and cell-type specific 2D (two-dimensional) densities. From these the 3D (three-dimensional) cell density was inferred by a model-based sampling technique, which is necessary for the calculation of local and global 3D tumour 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 for data. The integration of cell numbers information and DWI data derived from different tumour areas revealed a clear negative correlation between cell density and D value. Importantly, spatial tumour density can be quantitatively calculated based on DWI data to estimate tumour heterogeneity [55].

In a followup we currently study to what extent the relation between cellularity and DWI - diffusion coefficient can be inferred from biopsies instead of tumour serial sections. Moreover, we are studying the relation between DWI and tumour microvasculature [33].

7.1.7. Biomechanically mediated growth control of cancer cells

Mechanical feedback has been identified as a key regulator of tissue growth, by which external signals are transduced into a complex intracellular molecular machinery. Using multiscale computational modeling of multicellular growth in two largely different experimental settings with the same tumour cell line we demonstrated that the cellular growth response on external mechanical stress may nevertheless be surprisingly quantitatively predictable. Our computational model represents each cell as an individual unit capable of migration, growth, division, and death and is parameterised by measurable biophysical and bio-kinetic parameters. A cell cycle progression function depending on cell compression was established by comparisons of computer simulations with experiments of spheroids growing in an alginate elastic capsule. After a calibration step with free growing spheroids growing in a liquid suspension to capture the different growth conditions, the model using the same cell cycle progression function can predict the mechanical stress response of spheroid growth in a completely different experimental technique using Dextran, where stress is exerted by osmotic pressure. Our findings suggest that the stress response of cell growth may be highly reproducible even in otherwise different environments. This encourages the idea that robust functional modules may be identified, thus helping us to understand complex cell behaviours such as cell growth and division in relation to mechanical stress. The model analysis further elucidates the relation between applied pressure, cell compressibility and cell density. Moreover, the model developments within this paper points a way of how to handle the so far open issue of high compression within the popular so-called Centre-Based Models, in which force between cells is modelled as forces between cell centres [54].

7.2. Epidemiology

Participants: Luis Lopes Neves de Almeida, M. Soledad Aronna [FGV, Rio de Janeiro], Pierre-Alexandre Bliman, Flávio C. Coelho [FGV, Rio de Janeiro], Martin Strugarek, Nicolas Vauchelet, Jorge Zubelli [IMPA, Rio de Janeiro].

7.2.1. Establishing Wolbachia by feedback

The releases of *Wolbachia*-positive mosquitoes are usually completed on an open-loop approach, that is, with a schedule computed once for all before the beginning of the experiment. Using the fact that measurements are achieved and available during the whole release process, we applied feedback control technique to devise an introduction protocol which is proved to guarantee that the population converges to a stable equilibrium where the totality of mosquitoes carry *Wolbachia*. A major advantage of feedback compared to open-loop approaches is its ability to cope with the uncertainties in the model dynamics (typically in the modelling of the life stages and the population structure), in the parameters (population size, mortality, reproductive rates, etc.), and in the size of the population to be treated.

7.2.2. Travelling waves in the problem of infestation by Wolbachia

As described above, a new method of control of dengue fever consists in releasing Wolbachia-infected mosquitos in the field, in the aim to replace the whole existing population by a population unable to transmit Dengue fever. In the study of the feasibility of such a strategy, an important issue concerns the spacial propagation of the mosquitoes. More precisely, releasing infected mosquitoes in a given domain (which can be a part of the city of Rio de Janeiro), the hope is to invade the whole area. The study of this propagation phenomena falls into the study of existence of traveling wave. In a recent paper [30], the authors have proposed a mathematical model to study such phenomena and they have simplified it to recover a well-know simple bistable system for which existence of traveling wave is known. The study of the probability of success of spacial invasiveness has been performed in [53].

7.3. Aggregation Kinetics

Participants: Aurora Armiento, Tom Banks [CRSC, NCSU, Raleigh, USA], Etienne Bernard, Thibault Bourgeron, José Antonio Carrillo [Imperial College, London, United Kingdom], Marie Doumic, Dirk Drasdo, Miguel Escobedo [Universidad del País Vasco, Bilbao, Spain], Sarah Eugène, Pierre Gabriel [Université Paris-Dauphine], Marc Hoffmann [Ceremade, Université Paris-Dauphine], François James [MAPMO, Université d'Orléans], Nathalie Krell [Université de Rennes 1], Frédéric Lagoutière [Département de mathématiques d'Orsay], Philippe Moireau [Inria Paris Saclay, M3DISIM project-team], Benoît Perthame, Stéphanie Prigent, Human Rezaei [VIM, INRA Jouy-en-Josas], Lydia Robert [Laboratoire Jean Perrin, UPMC], Philippe Robert [Inria Paris, RAP project-team], Maria Teresa Teixeira [IBCP, Paris], Joan Torrent [INRA, Jouy-en-josas], Magali Tournus [Ecole Centrale de Marseille], Nicolas Vauchelet, Min Tang [Jiaotong University, Shanghai], Zhou Xu [IBCP, Paris], Wei-Feng Xue [University of Kent, United Kingdom], Yi Yin.

7.3.1. Heterogeneity as an intrinsic feature in biological dynamics

Variability in nucleated polymerisation The kinetics of amyloid assembly show an exponential growth phase preceded by a lag phase, variable in duration as seen in bulk experiments and experiments that mimic the small volumes of cells. Sarah Eugène's Ph.D, defended in September 2016, was devoted to the study of the origins and the properties of the observed variability in the lag phase of amyloid assembly currently not accounted for by deterministic nucleation dependent mechanisms. In [20], we formulated a new stochastic minimal model that is capable of describing the characteristics of amyloid growth curves despite its simplicity. We then solve the stochastic differential equations of our model and give mathematical proof of a central limit theorem for the sample growth trajectories of the nucleated aggregation process. These results give an asymptotic description for our simple model, from which closed form analytical results capable of describing and predicting the variability of nucleated amyloid assembly were derived. We also demonstrate the application of our results to inform experiments in a conceptually friendly and clear fashion. Our model offers a new perspective and paves the way for a new and efficient approach on extracting vital information regarding the key initial events of amyloid formation.

However, this first model does not explain completely the variability observed in the experiments. In [17], we thus investigated extensions to take into account other mechanisms of the polymerisation process that may have an impact on fluctuations.The first variant consists in introducing a preliminary conformation step to take into account the biological fact that, before being polymerised, a monomer has two states, regular or misfolded. Only misfolded monomers can be polymerised so that the fluctuations of the number of misfolded monomers can be also a source of variability of the number of polymerised monomers. The second variant represents the reaction rate of spontaneous formation of a polymer as of the order of α, with α some positive constant. First and second order results for the starting instant of nucleation are derived from these limit theorems. The proofs of the results rely on a study of a stochastic averaging principle for a model related to an Ehrenfest urn model, and also on a scaling analysis of a population model.

Image and statistical analysis of protein fibrils Protein fibrils present an important structural diversity, not only their length, but also their width, whether they present branches or not, etc. These structures may reveal the presence of different types of aggregates, possibly formed out of different polymerisation pathways. To analyse this diversity of shapes and structures, we developed an image analysis software, based on the expertise acquired by Y. Yin during her PhD for the image analysis of vessels. This software is able to track fibrils and measure their length, number of branches, and variable widths, even with poor quality images and crossing fibrils. This done, it allows us to perform a statistical analysis of the fibrils, to elucidate the main structuring features (Figure 3).

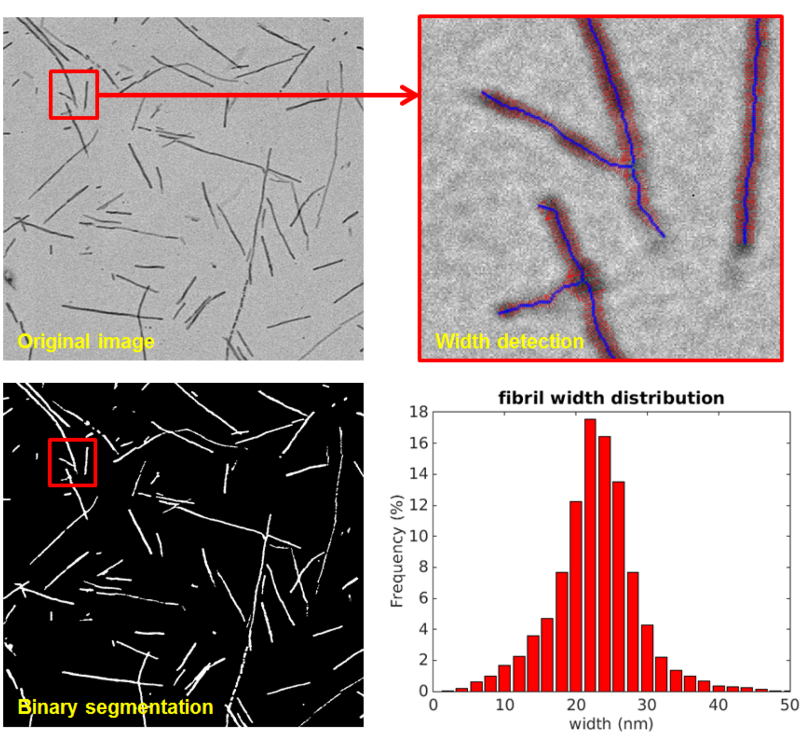

Figure 3. Fibril analysis yielding from original data via segmentation and analysis to the fibril length distribution.

7.3.2. Inverse Problems and Data Assimilation Applied to Protein Aggregation and other settings

Estimating reaction rates and size distributions of protein polymers is an important step for understanding the mechanisms of protein misfolding and aggregation, a key feature for amyloid diseases. A. Armiento's Ph.D was devoted to the question of adapting data assimilation strategies to the specific context and difficulties of protein aggregation. In [6], we settled a framework problem when the experimental measurements consist in the time-dynamics of a moment of the population (*i.e.*, for instance the total polymerised mass, as in Thioflavine T measurements, or the second moment measured by Static Light Scattering). We propose a general methodology, and we solve the problem theoretically and numerically in the case of a depolymerising system. We then apply our method to experimental data of degrading oligomers, and conclude that smaller aggregates of ovPrP protein should be more stable than larger ones. This has an important biological implication, since it is commonly admitted that small oligomers constitute the most cytotoxic species during prion misfolding process.

7.3.3. Time asymptotics for growth-fragmentation equations

The long-term dynamics of fragmentation and growth-fragmentation equations has constantly been for the MAMBA (and for the ex-BANG) team an important research field. Thanks to these common efforts, these equations are now well understood. However, there remain some interesting open questions. In particular, if the generic long-time behaviour for the linear equation is known - given by a (generally exponential) trend towards a steady exponential growth described by the positive eigenvector linked to the dominant eigenvalue, see [83] for most recent results - critical cases are not yet fully understood.

With Miguel Escobedo, we focused on an important critical case, when the fragmentation is constant and the growth rate is either null or linear [16]. Using the Mellin transform of the equation, we determine the long time behaviour of the solutions and the speed of convergence, which may be either exponential or at most polynomial according to the subdomain of $(t, x) \in \mathbb{R}_+^2$ which is considered. Our results show in particular the strong dependence of this asymptotic behaviour with respect to the initial data, in contrast to the generic results. Following our study, J. Bertoin and A. Watson proposed a complementary probabilistic analysis of related models [57]. These results exemplify the continuing need for further analysis of these interesting equations.

With E. Bernard and P. Gabriel, in [36], we investigated the "idealised" mitotic case, when the growth is exponential and the division results in two exactly equal parts. This case exhibits a lack of dissipativity, and the solutions appear to have a periodic limit cycle. We were nonetheless able to prove an entropy inequality, and to express the limit as an explicit oscillatory function, analytically given by the projection of the initial state on the space generated by the countable set of the dominant eigenvectors of the operator.

7.3.4. Cell aggregation by chemotaxis

Bacterial chemotaxis is now a well-known phenomenon. In particular, it has been established that the motion of bacteria is due to the alternation of straight swims in a given direction with tumble phases. More precisely, when bacteria notice that they do not go in a favorable direction, they may change their direction. Well established models are now available. In particular, the use of such systems allows to recover successfully the behaviour observed in biological experiments (see e.g. [18]). The bacterial response to changes in their environment can be described by an internal variable. In a recent work [29], it has been established that a well-known kinetic model can be obtained from such a model incorporating an internal variable.

When, the frequency of tumbling is high, the motion is mainly driven by tumbling and models reduce to describe aggregation phenomena. From a mathematical point of view, the study of such model is challenging since classical solution may not exist for any time. Then a notion of weak measure solution should be introduced [10]. Numerical investigation of such solutions has been performed in [21], [48].

7.4. Modelling of the liver

Participants: François Bertaux [Imperial College, London], Noémie Boissier, Dirk Drasdo, Géraldine Cellière, Adrian Friebel, Group Heinzle [Univ. Saarbruecken, Germany], Group Hengstler [IfADo, Germany], Stefan Hoehme, Tim Johann, Irène Reo [Vignon-Clementel], Paul Van Liedekerke, Eric Vibert [Hopital Paul Brousse], Group Zerial [Max-Planck Inst. for Molecular Genetics, Dresden, Germany], Groups Iflow, Notox, Vln.

7.4.1. Ammonia detoxification after drug-induced damage

Overdosing acetaminophen (APAP) is the main reason for acute liver failure in the US and UK. Overdose of APAP destroys the hepatocytes localised in the center of each liver lobule (pericentral damage), the repetitive functional and anatomical tissue units of liver. The Human has about 1 million of such lobules. As a consequence, the blood is not sufficiently detoxified from ammonia, which is toxic to the body and can lead to encephalopathy. In France about 1000 cases occur with ammonia toxicity each year. In recent papers we demonstrated by an integrated model that the widely accepted scheme of key reactions for ammonia detoxification is insufficient to explain ammonia detoxification after pericentral lobule damage and predicts a missing ammonia sink [71]. The integrated model couples ODEs representing the consensus reactions in the

spatial temporal liver lobule regeneration model. This finding has triggered new experiments leading to the identification of a widely ignored but fundamentally important ammonia sink mechanism. We could show by testing a number of different mechanisms within novel models that this sink mechanism was the only one able to explain the data [74] (and Geraldine Cellière's PhD thesis [3], 2016). The reaction turned out to have the potential to be used in therapeutics by injection of a molecular cocktail triggering it.
In a follow-up work, the ammonia detoxifying reactions have been integrated into each hepatocyte of the previously established tissue-level liver lobule model of regeneration. The final multi-level model simulates blood flow, transport of metabolites and detoxification of ammonia in every hepatocyte of a regenerating lobule. This multi-level model could validate the missing ammonia sink found in the integrated model in ref. [74] but yields differences to the integrated model if the ammonia sink mechanism is integrated. Still by reparameterisation, adding the ammonia sink mechanism, the model is able to explain the data but the results clearly show that spatio-temporal modelling can give results different from pure compartment modelling. In the case of quantitative modelling in pharmacology or toxicology this can be fundamental. We were able to analyse and generalise these findings.

7.4.2. Predicting in vivo drug toxicity from in vitro data

In vitro experiments on APAP (aka paracetamol, acetaminophen) have been used to calibrate a model of APAP drug toxicity using in vitro data, modifying this model to predict in vivo toxicity. This procedure is aimed at as a general pathway from cosmetic and pharmaceutical companies to eliminate or at least reduce animal experiments and, in perspective, permit a better prediction of drug toxicity in the Human. Three critical differences between in vitro and in vivo were stepwise integrated in the model calibrated with in vitro toxicity data to study their impact on in vivo toxicity predictions. (1) The temporal drug exposure profile, (2) the temporal concentration profile of a class of key enzymes, CYP enzymes. Only in hepatocytes in which CYP enzymes are present is APAP metabolised and can downstream cell death occur. (3) The liver architecture represents critical differences in the spatial distribution of the drug. The results are in preparation for publication (Géraldine Cellière's PhD thesis 2016, Cellière et. al., in preparation).

7.4.3. Liver cancer

The aggressiveness of a tumour may be reflected by its micro-architecture. To gain a deeper understanding of the mechanisms controlling the spatial organisation of tumors at early stages after tumour initiation, we used an agent-based spatio-temporal model previously established to simulate features of liver regeneration [76]. This model was further developed to simulate scenarios in early tumour development, when individual initiated hepatocytes gain increased proliferation capacity [37]. The model simulations were performed in realistic liver microarchitectures obtained from 3D reconstruction of confocal laser scanning micrographs. Interestingly, the here established model predicted that initially initiated hepatocytes arrange in elongated patterns. Only when the tumour progresses to cell numbers of approximately $4,000$ does it adopt spherical structures. This model prediction was validated by the analysis of initiated cells in a rat liver tumour initiation study using single doses of 250 mg/kg of the genotoxic carcinogen N-nitrosomorpholine (NNM). Indeed, small clusters of GST-P positive cells induced by NNM were elongated, almost columnar, while larger GDT-P positive foci of approximately the size of liver lobules, adopted spherical shapes. Simulation of numerous possible mechanisms demonstrated that only hepatocyte-sinusoidal-alignment (HSA), a previously discovered order mechanism involved in the coordination of liver tissue architecture, could explain the experimentally observed initial deviation from spherical shape. The present study demonstrates that the architecture of small hepatocellular tumour cell clusters early after initiation is still controlled by physiological control mechanisms. However, this coordinating influence is lost when the tumour grows to approximately $4,000$ cells, leading to further growth in spherical shape (Figure 4). Our findings stress the potential importance of organ micro-architecture in understanding tumour phenotypes.

7.5. Miscellaneous

Participants: M. Soledad Aronna [FGV, RIo de Janeiro], Bettina d'Avila Barros [FGV, Rio de Janeiro], Pierre-Alexandre Bliman, Noémie Boissier, Géraldine Cellière, Flávio C. Coelho [FGV, Rio de Janeiro],

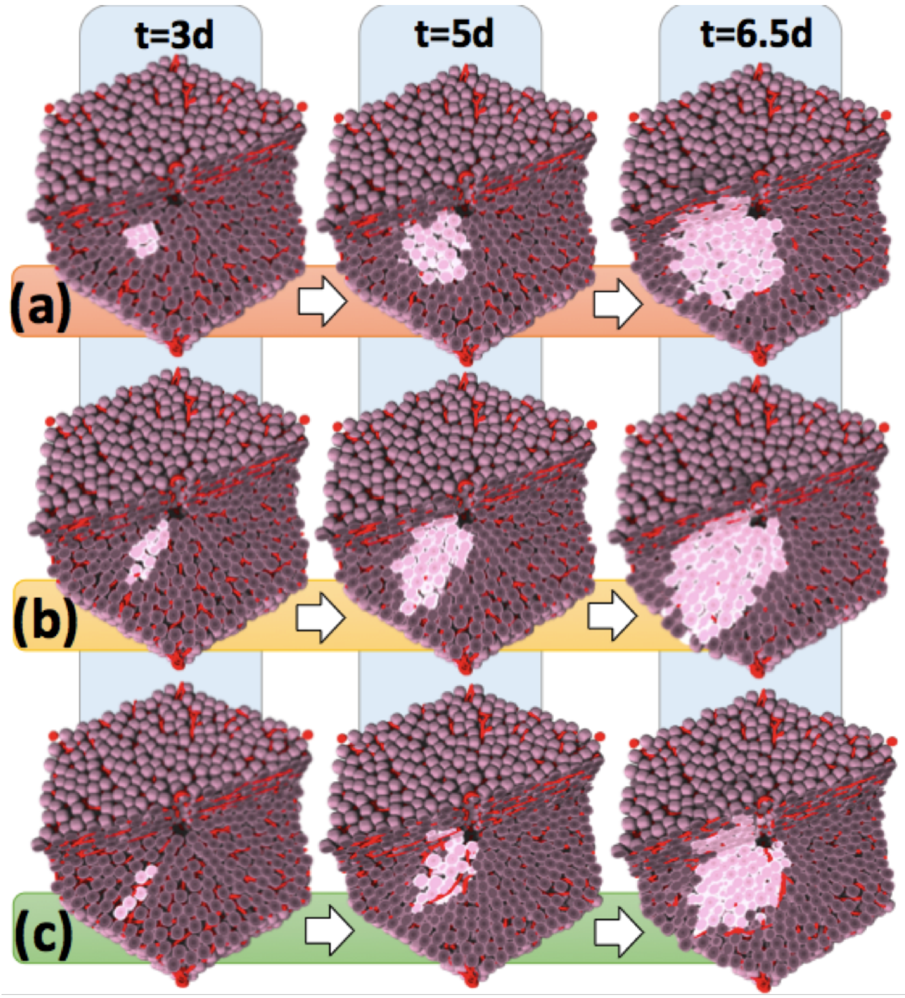

Figure 4. Scenarios of tumour growth in a single liver lobule in (a) absence of hepatocyte-sinusoidal-alignment (HSA), (b) presence of HSA, and (c) presence of HSA with elevated tangential friction impeding hepatocyte movement perpendicular to the columns formed along the sinusoids [35]. The images represent snapshots 3, 5 and 6.5 days after initiation, defined as the time point when a transformed hepatocyte adopts an increased proliferation rate. Notice that HSA (b, c) clearly causes early asymmetry of tumour cell assemblies (leftmost image column at 3 days) while with increasing tumour size this asymmetry is increasingly lost (right panel at 6.5 days). A one-cell thick column could be found if the movement perpendicular to the sinusoids was impeded by elevated shear forces, e.g., from tight junctions. This predicted evolutionary scenario reproduces the experimentally observed scenario.

Marie Doumic, Benoît Perthame, Tales Rands Amazonas [FGV, Rio de Janeiro], Group Reo [Inria Paris - Rocquencourt], Edouard Ribes [SANOFI], Martin Strugarek, Nathan Toubiana [École Polytechnique], Paul Van Liedekerke, Nicolas Vauchelet, Jorge Zubelli [IMPA, Rio de Janeiro].

7.5.1. *Diffusive waves generated by a travelling wave*

Observations in developmental biology show that calcic waves, generated after fertilisation within the egg cell endoplasmic reticulum, propagate within the egg cell. This motivates to explore in which circumstances a travelling wave solution of a reaction-diffusion equation can generate a travelling wave for the diffusion equation. For this purpose, we construct analytical solutions for a system composed of a reaction-diffusion equation coupled with a purely diffusive equation. We consider both the monostable (of the Fisher-KPP type) and bistable cases. We use a piecewise linear reaction term so as to build explicit solutions, which leads us to compute exponential tails, the exponents of which are roots of second, third or fourth-order polynomials. These rise conditions on the coefficients for existence of a travelling wave of the diffusion equation. The question of positivity and monotonicity is only partially answered. See [49].

7.5.2. *Dealing with uncertainty in modelling*

Interval observers for time-varying uncertain epidemiological models. SIR models constitute an elementary class of deterministic models of evolution of epidemics. We examine here the issue of state estimation for such models, subjected to seasonal variations and uncertainties in the transmission rates. Direct or indirect (through a vector) transmission is considered. In both cases, the measurement is assumed to consist of the number of new infectives per unit time, that is the information usually provided by the public health systems. We construct classes of interval observers with estimate-dependent gain, and provide corresponding asymptotic error bounds.

7.5.3. *Modelling strategic workforce planning with structured population equations*

We initiated a promising collaboration with the human resource department of Sanofi (E. Ribes), aiming at proposing a unified modelling of workforce planning based on structured population equations. Strategic Workforce Planning is a company process providing best in class, economically sound, workforce management policies and goals. Despite the abundance of literature on the subject, this is a notorious challenge in terms of implementation. Reasons span from the youth of the field itself to broader data integration concerns that arise from gathering information from financial, human resource and business excellence systems. In [43], we set the first stones to a simple yet robust quantitative framework for Strategic Workforce Planning exercises. Firstly, a method based on structured equations is detailed. It is then used to answer two main workforce-related questions: how to optimally hire to keep labour costs flat? How to build an experience-constrained workforce at a minimal cost? Further developments are in progress.

8. Partnerships and Cooperations

8.1. National Initiatives

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 2014-2017 IFLOW*

Eric Vibert, Hopital Paul Brousse (coordinator). Partners: Inria REO, Hopital Toulouse, Dirk Drasdo. Objectives are simulation of liver perfusion after partial hepatectomy (PHx) with and without therapeutic manipulations to improve patients survival after PHx.

8.1.1.3. 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 REO team, Inria Paris. The pursued objective is the bioengineering design of an artifical liver intended for liver replacement.

8.1.2. ITMO Cancer

8.1.2.1. ITMO Cancer 2014 - 2016, INVADE.

Emmanuel Barillot, Institut Curie (coordinator). Partners: Groups from Institut Curie, Dirk Drasdo. Objective is a model for a better understanding of breast cancer invasion.

8.1.2.2. ITMO Cancer 2016 THE call

See above “Highlights of the year"

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. ERC Starting Grant SKIPPER^{AD}, 2012-2017, Principal Investigator: Marie Doumic.

This grant allowed to fund Sarah Eugène’s and Mathieu Mézache’s Ph.Ds, as well as to develop new collaborations as with Wei-Feng Xue in Canterbury, Piotr Gwiazda in Poland, Teresa Teixeira and Zhou Xu in IBCP.

8.3. International Initiatives

8.3.1. Participation in Other International Programs

8.3.1.1. International Initiatives

CAPES-COFECUB Modelling innovative control methods for dengue fever

- Brazilian part headed by Claudio Struchiner
- French part headed by Benoît Perthame

MOSTICAW MOdelling the Spread and (opTImal) Control of Arboviroses by Wolbachia
International Partners (Institution - Laboratory - Researcher):

- Universidad de Buenos Aires (Argentina) - Hernán G. Solari
- Universidad de Chile (Chile) - Carlos Conca
- Universidade Federal Fluminense (Brazil) - Max Souza
- Universidad Tecnica Federico Santa Maria (Chile) - Pablo Aguirre
- EMAp (Brazil) - Pierre-Alexandre Bliman
- CIRAD (France) - Yves Dumont
- Duration: 2016 - 2017
- Start year: 2016
- The spread of certain strains of the intracellular parasitic bacterium Wolbachia in populations of mosquitoes Aedes aegypti drastically reduces their competence as vector of dengue and other severe mosquito-borne viral diseases known as arboviral infections. In absence of vaccine, or of preventive or curative treatment, the release of mosquitoes infected by the bacterium has been recently considered a promising tool to control these diseases, and experimental introductions in wild populations are currently under way in Brazil and Colombia. A key question about this method concerns the effective strategies of release of the infected mosquitoes in the field that can be applied with limited cost to reach the desired state of complete exclusion of Wolbachia-free mosquitoes. The mathematical study of central topics is the core of this project. The scientific questions to be addressed during this project are related to the study of the dynamic and control of the key invasion mechanism on finite-dimensional compartments.

- Pierre-Alexandre Bliman is International and Brazilian coordinator of the STIC Am-Sud project *MOdeling the Spread and (opTImal) Control of Arboviroses by Wolbachia* (MOSTICAW), 2016-2017. Partners: UBA (Argentina); FGV, Fiocruz, UFF (Brazil); UC, UTFSM (Chile), Universidad de Quindio, Universidad Autónoma de Occidente (Colombia), EPI MAMBA, INRA-Montpellier, CIRAD-Montpellier(France); UNA (Paraguay); Universidad Nacional Mayor de San Marcos (Peru).
- Pierre-Alexandre Bliman is also French coordinator of the ECOS-NORD project *New methods for the control of epidemics of dengue and arboviroses*, 2017-2019. Partner: Universidad del Valle, Cali, Colombia.

LiSym Liver Systems Medicine, BMBF funded project.

- Duration: 2016 - 2020
- Start year: 2016
- LiSym addresses liver diseases and regeneration, namely, steatosis, fibrosis and cirrhoses, and acutisation of chronic liver disease. It is composed of three subprojects and three junior research groups. Dirk Drasdo is co-coordinator of one of these three projects and participates in one of the others. He is also member of the leadership board.

8.4. International Research Visitors

8.4.1. Internships

- Andreas Buttenschoen (PhD student of Thomas Hillen, Univ. Edmonton, Alberta, Canada) has been welcomed in the MAMBA team, under Dirk Drasdo's supervision, for a 6-month internship within the framework of the Inria-MITACS programme. Program of the stay: Training on agent-based modeling of growth and cell migration; training on the software tool TiSim.
- Shalla Hanson (Duke University, Durham, NC) has been welcomed in the MAMBA team for a 6-month internship within the framework of the Chateaubriand programme. She is since October 2015 in a PhD thesis in co-tutela under the supervision of Michael Reed (Duke) and Jean Clairambault (MAMBA & UPMC).

8.4.2. Visits to International Teams

8.4.2.1. Sabbatical programme

BLIMAN Pierre-Alexandre

Date: Jun 2014 - Jul 2016

Escola de Matemática Aplicada

Institution: Fundação Getulio Vargas, Rio de Janeiro, Brazil

Chargé de mission at Direction des Partenariats Européens et Internationaux (DPEI), Inria

DOUMIC-JAUFFRET Marie

Date: Jun 2016 - Jul 2017

Institution: Wolfgang Pauli Institute, University of Vienna (Austria)

Sabbatical

8.4.2.2. Research Stays Abroad

STRUGAREK Martin

Date: Oct 2016

Institution: Fundação Oswaldo Cruz, Rio de Janeiro

Programme CAPES-COFECUB "Modelling innovative control methods for dengue fever"

VAUCHELET Nicolas

Date: Jan-Feb 2016

Institution: IMPA, Rio de Janeiro

Teaching collaboration between IMPA, Rio and UPMC, Paris

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Pierre-Alexandre Bliman: Chairman of the Conference on Mathematical Modeling and Control of Communicable Diseases, Fundação Getulio Vargas, Rio de Janeiro (RJ), Brazil, January 11-15, 2016

Marie Doumic: Chair at the Summer School PDE and Probability for Life Sciences, CIRM, Luminy, July 4-8, 2016.

9.1.1.2. Member of the Organising Committees

Luis Almeida, Benoît Perthame and Nicolas Vauchelet: Co-organisers of the “Second meeting on mathematical modeling and control in epidemic spread”, Laboratoire Jacques-Louis Lions, UPMC, Paris, May 23, 2016

Jean Clairambault: co-organiser of the mini-symposium “Heterogeneity, evolution and drug resistance in cancer”, ECMTB, Nottingham, England, July 12, 2016

Marie Doumic: co-organiser of the workshop on “Models in Cancer Therapy", WPI, Vienna, 1-2 july, 2016; mini-symposium organisation at the ECMTB, Nottingham, England, July 2016; mini-course and chair at CIMPA school in Moka, Mauritius, December 2016

Dirk Drasdo: co-coordinator of the modelling workpackage in ANR RHU project iLITE (coordinator: Jean-Charles Duclos-Vallée, Paul Brousse Hospital, Villejuif)

Dirk Drasdo: member of the scientific leadership team of the Liver Systems Medicine grant composed of three subprojects and three junior research groups, each subproject composed of about 10 PIs

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

Pierre-Alexandre Bliman: Program chairman of the 1st meeting of the STIC AmSud project MOSTICAW, Asunción, Paraguay, October 5–9 2016

9.1.2.2. Member of the Conference Program Committees

Pierre-Alexandre Bliman: Member of the Conference Editorial Board of European Control Association (EUCA), actuating for 15th European Control Conference, Aalborg, Denmark, June-July 2016

9.1.2.3. Reviewer

Pierre-Alexandre Bliman: Reviewer for the 55th IEEE Conference on Decision and Control, Las Vegas, USA, December 2016.

Dirk Drasdo: member of the reviewing committee for foundation of “Einstein Center for Regeneration in Compromised Patients Medicine”, Berlin, April 2016

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Dirk Drasdo is member of the boards of *TheScientificWorldJOURNAL* and *Royal Society open science* (UK) and was guest editor for *PloS Comput. Biol.* (2016)

Benoît Perthame is member of the boards of *Communications in PDEs, M3AS, NoDEA, Mathematical Medicine and Biology*

9.1.3.2. Reviewer - Reviewing Activities

Pierre-Alexandre Bliman: Reviewer for the journals *Automatica, IET Control Theory & Applications, Memórias do Instituto Oswaldo Cruz*
Jean Clairambault: Reviewer for the journals *Evolutionary Applications, Bulletin of Mathematical Biology, Mathematical Modelling of Natural Phenomena, Journal of Inorganic and Organometallic Polymers and Materials, Journal of Theoretical Biology, British Journal of Cancer, PLoS Computational Biology, BMC Cancer*
Marie Doumic: Reviewer for the journals *Inverse Problems, Analytical Biochemistry, European Journal of Applied Mathematics, Bull. of Math. Biology, Comm. in Math. Sciences*
Dirk Drasdo: Reviewer for *Nature, Scientific Reports* and other journals
Nicolas Vauchelet: Reviewer for Transaction AMS, SIAM J. Numer. Anal., M3AS, J. Optim. Theory Appl., Math. Reviews

9.1.4. Invited Talks and Courses

Luis Almeida: EMS (European Mathematical Society) Diderot Mathematical Forum, Paris, March 15, 2016
Luis Almeida: IWorkshop “Models in cancer therapy”, WPI, Vienna, Austria, July 1-2, 2016
Luis Almeida: Master course on Reaction-Diffusion Equations Arising in the Mathematical Modelling of Population Dynamics, Univ. Verona, October 2016
Pierre-Alexandre Bliman: Keynote speaker at the International conference on Digital Sciences and Technologies for Health, Paris, France, June 10, 2016
Pierre-Alexandre Bliman: Seminars at UMR ESPACE-DEV, Université de Guyane, Cayenne, France, December 2016
Pierre-Alexandre Bliman: Seminars at Laboratoire de Mathématiques et Dynamique de Populations, Université Cadi Ayyad, Marrakesh, Morocco, December 2016
Jean Clairambault: 2h course at the Winter School and Workshop “Nonlocal aspects in mathematical biology”, Bedlewo, Poland, January 27, 2016
Jean Clairambault: French-Serbian Novi Sad Oncology Congress, Novi Sad, Serbia, March 18-19, 2016
Jean Clairambault: Journées du département ONCO, Nantes, May 3-4, 2016
Jean Clairambault: Workshop “Le cancer en équations”, Rabat, Morocco, May 5-6, 2016
Jean Clairambault: First Waterloo University - Sorbonne Universités Seminar, Waterloo, Ontario, Canada, May 9-11, 2016
Jean Clairambault: 3h course at the BIOMAT Summer school “Cell dynamics and polymerization”, Granada, Spain, June 1-3, 2016
Jean Clairambault: 4.5h course at the CIMPA Summer school “Mathematical modeling in Biology and Medicine”, Santiago de Cuba, June 14-15-16, 2016
Jean Clairambault: Workshop “Models in cancer therapy”, WPI, Vienna, Austria, July 1-2, 2016
Jean Clairambault: International conference “Mathematical models in biology and medicine”, Moscow, October 31-November 3, 2016
Jean Clairambault: 4h course at the Winter school “Mathematical Models of Tumour and Disease”, Jiaotong University, Shanghai, December 5-6-7-8, 2016
Jean Clairambault: Workshop on mathematical biology, Jiaotong University, Shanghai, December 10, 2016
Jean Clairambault: Workshop on Mathematical Modelling and Computation in Medicine/Biology, Tsinghua International Mathematics Forum (TSIMF), Sanya, Hainan, China, December 12-16, 2016
Marie Doumic: Plenary Speaker at the Diderot Mathematical Forum, March 15; seminar at the Polish Academy of Sciences (Warsaw), March 2016
Marie Doumic: 3h course at the BioMat2016 Conference in Granada, Spain, 1-3 June 2016; seminar in Orsay, June 16, 2016
Marie Doumic: Workshop on fragmentation processes, November 17, Villetaneuse; workshop on “Recent contributions of women in PDEs”, Vienna, November 28-30, 2016
Marie Doumic: 4.5h course at the CIMPA Winter school, Moka, Mauritius, December 4-16, 2016
Dirk Drasdo: Workshop OPENTOX Basel, March 2016

Dirk Drasdo: Workshop Biomath/Bioinfo/BioStat of Cancer, Lyon, June 2016
Dirk Drasdo: CMBBE (14th international Symposium on Computational Methods in Biomechanics and Biomedical Engineering), Tel Aviv, September 2016
Benoît Perthame: Seminar, University of Chicago, January 2016
Benoît Perthame: Distinguished lecture, Hong Kong Polytechnic University, February 2016
Benoît Perthame: Seminar, Basel, Switzerland, March 2016
Benoît Perthame: Seminar, Padova, April 2016
Benoît Perthame: Course in mathematical biology, Edmonton, Alberta, Canada, May 2016
Benoît Perthame: Course on "Kinetic equations for cell motility", Porto Ercole, Italy, June 2016
Benoît Perthame: Conference in honour of Peter Markowich's 60th birthday, Beijing, July 2016
Benoît Perthame: Conference on "Kinetics and quantum dynamics", Shanghai, July 2016
Benoît Perthame: Course on "Adaptive evolution", Valparaiso, September 27-30, 2016
Benoît Perthame: Conference on "SCL with rough fluxes", Mittag-Leffler Institute, Stockholm, September 12-15, 2016
Benoît Perthame: Seminar, ETH Zürich, October 18-19, 2016
Benoît Perthame: Conference in honour of Peter Markowich's 60th birthday, KAUST, Saudi Arabia, October 31-November 3, 2016
Benoît Perthame: Conference on "Networks and collective behaviours", Seoul, November 7-10, 2016
Benoît Perthame: Courses, Analysis school, Cotonou, Benin, December 5-9, 2016
Camille Pouchol: Winter School and Workshop "Nonlocal aspects in mathematical biology", Bedlewo, Poland, January 27, 2016
Camille Pouchol: International conference "Mathematical models in biology and medicine", Moscow, October 31-November 3, 2016
Nicolas Vauchelet: 4h course at Imperial College, London, October 2016, UK
Nicolas Vauchelet: 3h course at the CIMPA Summer school "Mathematical modeling in Biology and Medicine", Santiago de Cuba, June 2016, Cuba
Nicolas Vauchelet: INdAM Workshop "Interactions between Analysis and Innovative Algorithmics", Rome, May 2016, Italy

9.1.5. Scientific Expertise

Pierre-Alexandre Bliman: Member of the Scientific committee of the ANR program "Environnement, pathogènes et maladies émergentes ou ré-émergentes - One health"
Pierre-Alexandre Bliman: Expert for the Belgium agency FNRS, for the Dutch agency NWO
Pierre-Alexandre Bliman: Reviewer for the European PhD Award on Control for Complex and Heterogeneous Systems
Pierre-Alexandre Bliman: Member of the National network of specialists of Zika and related diseases (Rede Nacional de Especialistas em Zika e doenças correlatas, RENEZIKA), Health Ministry of Brazil
Pierre-Alexandre Bliman: Member of the Brazilian National Institute for Science and Technology (INCT)
Jean Clairambault: Expert for Belgian FNRS, for the Moffitt Center (Tampa, FL), for the University of Yaoundé (Cameroon), for the BBSRC (UK), for the ERC (Consolidator Grant 2016), for the Royal Society of Edinburgh
Dirk Drasdo: Member of the program committee for SBMC 2016 (Conference on Systems Biology of Mammalian Cells) in Munich

9.1.6. Research Administration

Luis Almeida: Member of the bureau of CID 51 of the Comité National de la Recherche Scientifique
Luis Ameida: in charge of the Major MathBio of the speciality "Mathematics of modelling", M2 level, UPMC
Jean Clairambault: member of the bureau of the IPV (Interfaces pour le Vivant) doctoral funding programme of UPMC, representative of ED 386 (since 2014)
Jean Clairambault: member of the expert group of ITMO Cancer, representative of Inria (since 2008)
Marie Doumic: member of the selection committee for an assistant professor position in Grenoble

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Pierre-Alexandre Bliman: *Análise*, Escola de Matemática Aplicada, Fundação Getulio Vargas, Rio de Janeiro, Brazil (60 h)
Pierre-Alexandre Bliman: *Introdução à Teoria do controle*, Escola de Matemática Aplicada, Fundação Getulio Vargas, Rio de Janeiro, Brazil (60 h)
Marie Doumic: Master course on inverse problems and applications in population dynamics (24 h)
Dirk Drasdo: Master M2, Mathematical biology, on "Agent-based models of tissue organization", from January 2016 to April 2016 at Paris VI (Mathematics department, in 10 units per semester, total 24h)

9.2.2. Supervision

9.2.2.1. PhD defences in 2016

- PhD defence: François Bertaux, " Cell-based multi-scale modeling for systems and synthetic biology: from stochastic gene expression in single cells to spatially organized cell populations" [1], UPMC, June 2016, supervision by G. Batt (Lifeware, Inria Saclay) and D. Drasdo
- PhD defence: Youssef Bourfia, "Modélisation et Analyse de Modèles en Dynamique Cellulaire avec Applications à des Problèmes Liés aux Cancers" [2], Cadi Ayyad University, Marrakesh, December 28, 2016, supervision by M. Adimy (Lyon), J. Clairambault and H. Hbid (Marrakesh)
- PhD defence: Géraldine Cellière: "Multi-scale modeling of drug and detox-metabolism in liver" [3], Ecole du Vivant, Univ. Paris-Diderot, July 2016, supervision by D. Drasdo
- PhD defence: Casimir Emako, "Study of two-species chemotaxis model" [4], March 17, 2016, supervision by L. Almeida and N. Vauchelet
- PhD defence: Sarah Eugène, "Stochastic modelling in molecular biology: a probabilistic analysis of protein polymerisation and telomere shortening" [5], September 30, 2016, supervision by M. Doumic and Ph. Robert (Inria Paris, RAP team)

9.2.2.2. Ongoing PhD theses

- PhD in progress: Aurora Armiento, "Inverse problems for aggregation kinetics", UPMC, begun September 2013, supervision by M. Doumic and Ph. Moireau (Inria Saclay, M3DISIM team)
- PhD in progress: Noémie Boissier (since November 2013), supervision by D. Drasdo and I. Vignon-Clementel
- PhD in progress: Walid Djema, "Analysis of an AML model enabling evaluation of polychemotherapies delivered in the case of AML which have a high level of Flt-3 duplication (Flt-3-ITD)", supervision by C. Bonnet (DISCO, Saclay), J. Clairambault, and F. Mazenc (DISCO, Saclay)
- PhD in progress: Adrian Friebel, "Software of image processing and analysis of liver tissue at histological scales.", supervision by D. Drasdo and S. Hoehme
- PhD in progress: Ghassen Haddad, "Optimisation of cancer treatments: application to bladder cancer", UPMC in co-tutela with ENIT Tunis, begun October 2015, supervision by J. Clairambault and S. Ben Miled (Tunis)
- PhD in progress: Shalla Hanson, "Modelling evolution of interactions between cancer and immune cells in solid tumours", UPMC in co-tutela with Duke University, begun October 2015, supervision by J. Clairambault and M. Reed (Duke)
- PhD in progress: Hugo Martin, "New structured population models for bacterial growth", begun October 2016, supervision by M. Doumic in co-tutela with Pierre Gabriel (Versailles)
- PhD in progress: Mathieu Mézache, begun October 2016, , "Oscillatory dynamics in protein aggregation", supervision by M. Doumic in co-tutela with Human Rezaei (INRA)
- PhD in progress: Johannes Neitsch, "Growth and regeneration modeling based on an agent-based model with deformable cells", (since June 2011) supervision by D. Drasdo and P. Van Liedekerke

- PhD in progress: Pastor Pérez-Estigarribia, Universidad Nacional de Asunción, Paraguay, supervision by C. Schaerer and P.-A. Bliman
- PhD in progress: Camille Pouchol, "Modelling interactions between tumour cells and adipocytes in breast cancer", UPMC, begun September 2015, supervision by J. Clairambault, M. Sabbah, and E. Trélat
- PhD in progress: Antonin Prunet, UPMC, begun October 2014, supervision by L. Almeida and M. Sabbah
- PhD in progress: Andrada Quillas Maran, "Modelling early leukaemogenesis", UPMC, begun March 2014, supervision by J. Clairambault, F. Delhommeau and B. Perthame
- PhD in progress: Martin Strugarek, "Structured population dynamics for transmissible diseases", UPMC, begun October 2015, supervision by N. Vauchelet and B. Perthame
- PhD in progress: Cécile Taing, UPMC, begun October 2014, supervision by A. Lorz and B. Perthame

9.2.2.3. Graduate thesis defences in 2016

- Graduate thesis defence: Bettina D'Avila Barros, Escola de Matemática Aplicada, Fundação Getulio Vargas, Brazil, advisor P.-A. Bliman
- Graduate thesis (ENSAE 2nd year internship) defence: Hicham Janati [46], Malakoff, France, December 2016, supervision by J. Clairambault and M. Doumic
- Graduate thesis defence: Tales Amazonas Rands, Escola de Matemática Aplicada, Fundação Getulio Vargas, Brazil, advisor P.-A. Bliman

9.2.3. Juries

- Luis Almeida: Casimir Emako, UPMC 17/03/2016
- Luis Almeida (reviewer): Perrine Berment, Univ. Bordeaux 06/07/2016
- Luis Almeida: Thibault Liard, UPMC 04/11/2016
- Pierre-Alexandre Bliman: Hafiz Ahmed, Université de Lille, 22/09/2016
- Pierre-Alexandre Bliman: Youssef Bourfia, Université Cadi Ayyad, Maroc & UPMC, 28/12/2016
- Jean Clairambault (reviewer): Douglas Friesen, University of Edmonton, remote defence committee member, 23/02/2016
- Jean Clairambault: Tiphaine Obara, Nancy, 07/10/2016
- Jean Clairambault: Youssef Bourfia, Université Cadi Ayyad, Maroc & UPMC, 28/12/2016
- Marie Doumic (reviewer): Etienne Baratchart, Université de Bordeaux, 2016
- Marie Doumic: Casimir Emako, UPMC, 2016
- Marie Doumic: Sarah Eugène, UPMC, 2016
- Dirk Drasdo: François Bertaux, UPMC, 2016
- Dirk Drasdo: Geraldine Cellière, Ecole du Vivant, Univ. Paris-Diderot, 2016
- Benoît Perthame (reviewer): Thierry Pichard, Université de Bordeaux
- Benoît Perthame: Vincent Renault, UPMC
- Benoît Perthame: Julien Chevalier, Université de Nice-Sophia Antipolis
- Benoît Perthame: Thibault Balois, LPS-ENS
- Nicolas Vauchelet: Casimir Emako, UPMC, March 2016
- Nicolas Vauchelet: Pierre-Louis Colin, Université Lille 1, June 2016

9.3. Popularisation

Marie Doumic: Invited talk in the "Science et société" dissemination conference (http://www.iecl.univ-lorraine.fr/Cycle-Conferences-Sciences-et-Societe/lanceur.php?action=accueil) in Nancy, May 26, 2016

Nicolas Vauchelet : talks for Animath (http://www.animath.fr/) in two high schools : Lycée Racine and Lycée Notre-Dame de Bourg-la-Reine, March 2016

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] F. BERTAUX.*Cell-based multi-scale modeling for systems and synthetic biology: from stochastic gene expression in single cells to spatially organized cell populations*, Université Pierre & Marie Curie - Paris 6, June 2016, https://tel.archives-ouvertes.fr/tel-01405430.

[2] Y. BOURFIA.*Mathematical Modeling in Cellular Dynamics: Applications to Cancer Research*, Université Paris 6 - Pierre et Marie Curie ; Université Cadi Ayyad, December 2016, https://hal.archives-ouvertes.fr/tel-01426296.

[3] G. CELLIERE.*Multi-scale modeling of hepatic drug toxicity and its consequences on ammonia detoxification*, MAMBA - Modelling and Analysis for Medical and Biological Applications, LJLL - Laboratoire Jacques-Louis Lions, Inria Paris-Rocquencourt, Universite Paris 7 Diderot, 2016, 221.

[4] C. EMAKO KAZIANOU.*Study of two-species chemotaxis models*, Université Pierre et Marie Curie - Paris VI, March 2016, https://tel.archives-ouvertes.fr/tel-01365414.

[5] 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.

Articles in International Peer-Reviewed Journal

[6] 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.

[7] H. T. BANKS, M. DOUMIC, C. KRUSE.*Efficient Numerical Schemes for Nucleation-Aggregation Models: Early Steps*, in "Journal of Mathematical Biology", 2016 [*DOI :* 10.1007/S00285-016-1026-0], https://hal.inria.fr/hal-00954437.

[8] A. BHAYA, P.-A. BLIMAN, F. PAZOS.*Cooperative concurrent asynchronous computation of the solution of symmetric linear systems*, in "Numerical Algorithms", 2016 [*DOI :* 10.1007/S11075-016-0213-9], https://hal.inria.fr/hal-01395756.

[9] G. BOCHAROV, A. BOUCHNITA, J. CLAIRAMBAULT, V. VOLPERT.*Mathematics of Pharmacokinetics and Pharmacodynamics: Diversity of Topics, Models and Methods*, in "Mathematical Modelling of Natural Phenomena", 2016, https://hal.inria.fr/hal-01413795.

[10] 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.

[11] 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 "Biochimica et Biophysica Acta (BBA) - General Subjects", June 2016, vol. 1860, p. 2627 - 2645 [*DOI :* 10.1016/J.BBAGEN.2016.06.009], https://hal.inria.fr/hal-01321535.

[12] J. CLAIRAMBAULT, O. FERCOQ.*Physiologically structured cell population dynamic models with applications to combined drug delivery optimisation in oncology*, in "Mathematical Modelling of Natural Phenomena", 2016, 22, V2 d'un dépôt précédemment effectué sous la référence clairambault:hal-01321536v1, https://hal.inria.fr/hal-01413791.

[13] F. CODEÇO COELHO, C. TORRES CODEÇO, O. GONÇALVES CRUZ, S. CAMARGO, P.-A. BLIMAN.*Epidemiological data accessibility in Brazil: Current challenges for an adequate response to emergencies*, in "Lancet Infectious Diseases", 2016, vol. 16, p. 524 - 525 [*DOI :* 10.1016/S1473-3099(16)30007-X], https://hal.inria.fr/hal-01395754.

[14] J. DE FIGUEIRÓ SANTOS, F. CODEÇO COELHO, P.-A. BLIMAN.*Behavioral modulation of the coexistence between Apis melifera and Varroa destructor : A defense against colony colapse disorder?*, in "PLoS ONE", September 2016 [*DOI :* 10.1371/JOURNAL.PONE.0160465], https://hal.inria.fr/hal-01395823.

[15] B. DESPRÉS, B. PERTHAME.*Uncertainty propagation;intrusive kinetic formulations of scalar conservation laws*, in "SIAM/ASA Journal on Uncertainty Quantification", 2016, vol. 4, n^o 1, p. 980-1013, http://hal.upmc.fr/hal-01146188.

[16] 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.

[17] 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.

[18] C. EMAKO, C. GAYRARD, A. BUGUIN, L. NEVES DE ALMEIDA, N. VAUCHELET.*Traveling Pulses for a Two-Species Chemotaxis Model*, in "PLoS Computational Biology", April 2016, vol. 12, n^o 4, e1004843 [*DOI :* 10.1371/JOURNAL.PCBI.1004843], https://hal.archives-ouvertes.fr/hal-01302632.

[19] S. EUGENE, T. BOURGERON, Z. XU.*Effects of initial telomere length distribution on senescence onset and heterogeneity*, in "Journal of Theoretical Biology", January 2017, vol. 413, 8, https://hal.inria.fr/hal-01378596.

[20] S. EUGENE, W.-F. XUE, P. ROBERT, M. DOUMIC.*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.

[21] 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.

[22] J. HASKOVEC, P. MARKOWICH, B. PERTHAME, M. SCHLOTTBOM.*Notes on a PDE System for Biological Network Formation*, in "Nonlinear Analysis: Real World Applications", June 2016, vol. 138, p. 127-155 [*DOI :* 10.1016/J.NA.2015.12.018], https://hal.archives-ouvertes.fr/hal-01232080.

[23] 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° 2, e1004412 [*DOI :* 10.1371/JOURNAL.PCBI.1004412], http://hal.upmc.fr/hal-01244593.

[24] I. C. KIM, B. PERTHAME, P. E. SOUGANIDIS.*Free boundary problems for tumor growth: a viscosity solutions approach*, in "Nonlinear Analysis Series A: theory, methods and applications", 2016, vol. 138, p. 207-228 [*DOI :* 10.1016/J.NA.2016.01.019], http://hal.upmc.fr/hal-01155696.

[25] 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° 1, 43 [*DOI :* 10.1186/S13062-016-0143-4], https://hal.inria.fr/hal-01355314.

[26] T. LORENZI, A. LORZ, B. PERTHAME.*On interfaces between cell populations with different mobilities*, in "Kinetic and Related Models ", March 2017, vol. 10, n° 1, p. 299-311, https://hal.inria.fr/hal-01257180.

[27] N. OUTADA, N. VAUCHELET, T. AKRID, M. KHALADI.*From Kinetic Theory of Multicellular Systems to Hyperbolic Tissue Equations: Asymptotic Limits and Computing*, in "Mathematical Models and Methods in Applied Sciences", 2017, https://hal.archives-ouvertes.fr/hal-01378301.

[28] B. PERTHAME, P. E. SOUGANIDIS.*Rare mutations limit of a steady state dispersal evolution model*, in "Mathematical Modelling of Natural Phenomena", 2016, vol. 11, n° 4 [*DOI :* 10.1051/MMNP/201611411], http://hal.upmc.fr/hal-01149029.

[29] B. PERTHAME, M. TANG, N. VAUCHELET.*Derivation of the bacterial run-and-tumble kinetic equation from a model with biochemical pathway*, in "Journal of Mathematical Biology", 2016, http://hal.upmc.fr/hal-01131101.

[30] M. STRUGAREK, N. VAUCHELET.*Reduction to a single closed equation for 2 by 2 reaction-diffusion systems of Lotka-Volterra type*, in "SIAM Journal on Applied Mathematics", 2016, vol. 76, n° 5, p. 2060-2080, https://hal.archives-ouvertes.fr/hal-01264980.

International Conferences with Proceedings

[31] P.-A. BLIMAN, B. D 'AVILA BARROS.*Interval observer for SIR epidemic model subject to uncertain seasonality*, in "Positive Systems - POSTA 2016", Roma, Italy, September 2016, https://hal.inria.fr/hal-01309743.

[32] W. DJEMA, F. MAZENC, C. BONNET, J. CLAIRAMBAULT, P. HIRSCH, F. DELHOMMEAU.*Stability of a Delay System Coupled to a Delay 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.

[33] Y. YIN, O. SEDLACZEK, J. LOTZ, J. OLESCH, K. BREUHAHN, D. DRASDO, I. VIGNON-CLEMENTEL.*Tumor Microvasculature in Lung Cancer and Diffusion-Weighted MRI: Preliminary Results*, in "IEEE Nuclear Science Symposuim & Medical Imaging Conference", Strasbourg, France, October 2016, https://hal.inria.fr/hal-01421152.

[34] Y. YIN, O. SEDLACZEK, A. WARTH, M. GONZÁLEZ-VALLINAS, K. BREUHAHN, I. VIGNON-CLEMENTEL, D. DRASDO.*Quantitative Estimation of Tumor Cellularity Based on Histology Data*, in "IEEE Nuclear Science Symposuim & Medical Imaging Conference, Oct 2016", Strasbourg, France, October 2016, https://hal.inria.fr/hal-01421163.

Scientific Books (or Scientific Book chapters)

[35] S. HOEHME, A. FRIEBEL, S. HAMMAD, D. DRASDO, J. G. HENGSTLER.*Creation of Three-dimensional Liver Tissue Models from Experimental Images for Systems Medicine*, in "Hepatocyte Transplantation", B. C. PEGGY STOCK (editor), Methods in Molecular Biology, Springer, November 2016, vol. 1506, p. 319-362, PubMed-listed [*DOI :* 10.1007/978-1-4939-6506-9_22], https://hal.inria.fr/hal-01426722.

Other Publications

[36] E. BERNARD, M. DOUMIC, P. GABRIEL.*Cyclic asymptotic behaviour of a population reproducing by fission into two equal parts*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01363549.

[37] F. BERTAUX, S. HOEHME, W. WEENS, B. GRASL-KRAUPP, J. G. HENGSTLER, D. DRASDO.*Model prediction and validation of an order mechanism controlling the spatio-temporal phenotype of early hepatocellular carcinoma*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01426629.

[38] S. V. BITSEKI PENDA, A. OLIVIER.*Autoregressive Functions Estimation in Nonlinear Bifurcating Autoregressive Models*, February 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01159255.

[39] M. BURGER, A. LORZ, M.-T. WOLFRAM.*Balanced Growth Path Solutions of a Boltzmann Mean Field Game Model for Knowledge Growth*, February 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01267078.

[40] J. CLAIRAMBAULT, L. ALMEIDA, R. CHISHOLM, T. LORENZI, A. LORZ, B. PERTHAME, C. POUCHOL, E. TRÉLAT.*Tumour growth and drug resistance: an evolutionary view with perspectives in therapeutics*, July 2016, ECMTB 2016 - European Conference on Mathematical and Theoretical Biology, Poster, https://hal.inria.fr/hal-01378290.

[41] J. CLAIRAMBAULT, O. FERCOQ.*Physiologically structured cell population dynamic models with applications to combined drug delivery optimisation in oncology*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01321536.

[42] M. DOUMIC.*Simulation of the growth-fragmentation equation in a periodic case - Non-dissipative numerical scheme*, December 2016, https://hal.inria.fr/medihal-01407826.

[43] M. DOUMIC, B. PERTHAME, E. RIBES, D. SALORT, N. TOUBIANA.*Toward an integrated workforce planning framework using structured equations*, July 2016, working paper or preprint, https://hal.inria.fr/hal-01343368.

[44] C. EMAKO, M. TANG.*Well-balanced and asymptotic preserving schemes for kinetic models* , March 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01265029.

[45] M. HOFFMANN, A. OLIVIER.*Nonparametric estimation of the division rate of an age dependent branching process*, January 2016, working paper or preprint [*DOI :* 10.1016/J.SPA.2015.11.009], https://hal.archives-ouvertes.fr/hal-01254203.

[46] H. JANATI.*Investigating cancer resistance in a Glioblastoma cell line with gene expression data*, Inria, December 2016, https://hal.inria.fr/hal-01412944.

[47] C. JOURDANA, P. PIETRA, N. VAUCHELET.*Hybrid coupling of a one-dimensional Energy-Transport Schrödinger system*, 2016, Accepted for publication in Monatshefte für Mathematik, https://hal.archives-ouvertes.fr/hal-01052415.

[48] F. LAGOUTIÈRE, N. VAUCHELET.*Analysis and simulation of nonlinear and nonlocal transport equations*, August 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01352548.

[49] H. MOUNDOYI, A. MOUSSA, B. PERTHAME, B. SARELS.*Analytical examples of diffusive waves generated by a traveling wave*, November 2016, working paper or preprint, http://hal.upmc.fr/hal-01404972.

[50] A. OLIVIER.*How does variability in cells' aging and growth rates influence the malthus parameter?*, February 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01274529.

[51] B. PENDA, M. HOFFMANN, A. OLIVIER.*Adaptive Estimation for Bifurcating Markov Chains*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01254200.

[52] 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, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01416594.

[53] M. STRUGAREK, N. VAUCHELET, J. ZUBELLI.*Quantifying the Survival Uncertainty of Wolbachia-infected Mosquitoes in a Spatial Model* *, August 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01355118.

[54] 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.

[55] 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*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01421398.

References in notes

[56] 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.

[57] J. BERTOIN, A. R. WATSON.*Probabilistic aspects of critical growth-fragmentation equations*, in "ArXiv e-prints", June 2015.

[58] F. BILLY, J. CLAIRAMBAULT.*Designing proliferating cell population models with functional targets for control by anti-cancer drugs*, in "Discrete and Continuous Dynamical Systems - Series B", 2013, vol. 18, n^{o} 4, p. 865 - 889, 24 pages [*DOI :* 10.3934/DCDSB.2013.18.865], https://hal.archives-ouvertes.fr/hal-00726195.

[59] F. BILLY, J. CLAIRAMBAULT, O. FERCOQ.*Optimisation of cancer drug treatments using cell population dynamics*, in "Mathematical Methods and Models in Biomedicine", U. LEDZEWICZ, H. SCHÄTTLER, A. FRIEDMAN, E. KASHDAN (editors), Lecture Notes on Mathematical Modelling in the Life Sciences, Springer New York, January 2013, 265 [*DOI :* 10.1007/978-1-4614-4178-6_10], https://hal.archives-ouvertes.fr/hal-00770366.

[60] F. BILLY, J. CLAIRAMBAULT, O. FERCOQ, S. GAUBERT, T. LEPOUTRE, T. OUILLON, S. SAITO.*Synchronisation and control of proliferation in cycling cell population models with age structure*, in "Mathematics and Computers in Simulation", February 2014, vol. 96, p. 66-94 [*DOI :* 10.1016/J.MATCOM.2012.03.005], https://hal.archives-ouvertes.fr/hal-00662885.

[61] 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.

[62] V. CALVEZ, M. DOUMIC, P. GABRIEL.*Self-similarity in a General Aggregation-Fragmentation Problem; Application to Fitness Analysis*, in "J. Maths Pures Appl.", November 2010, vol. 98, p. 1-27, https://hal.archives-ouvertes.fr/hal-00539279.

[63] 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.

[64] L. A. D'ALESSANDRO, S. HOEHME, A. HENNEY, D. DRASDO, U. KLINGMÜLLER.*Unraveling liver complexity from molecular to organ level: Challenges and perspectives*, in "Progress in Biophysics and Molecular Biology", January 2015, vol. 117, n^{o} 1, p. 78–86 [*DOI :* 10.1016/J.PBIOMOLBIO.2014.11.005], https://hal.inria.fr/hal-01257160.

[65] M. DOUMIC, P. GABRIEL.*Eigenelements of a General Aggregation-Fragmentation Model*, in "Mathematical Models and Methods in Applied Sciences", May 2010, vol. 20, p. 757-783, https://hal.archives-ouvertes.fr/hal-00408088.

[66] 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.

[67] M. DOUMIC, M. HOFFMANN, P. REYNAUD-BOURET, V. RIVOIRARD.*Nonparametric estimation of the division rate of a size-structured population*, in "SIAM Journal on Numerical Analysis", April 2012, vol. 50, n^{o} 2, p. 925–950 [*DOI :* 10.1137/110828344], https://hal.archives-ouvertes.fr/hal-00578694.

[68] M. DOUMIC, P. MAIA, J. ZUBELLI.*On the Calibration of a Size-Structured Population Model from Experimental Data*, in "Acta Biotheoretica", 2010, vol. 58, n^{o} 4, p. 405–413, http://dx.doi.org/10.1007/s10441-010-9114-9.

[69] M. DOUMIC, L. TINE.*Estimating the Division Rate for the Growth-Fragmentation Equation*, in "Journal of Mathematical Biology", 2013, vol. 67, n^o 1, p. 69–103, http://dx.doi.org/10.1007/s00285-012-0553-6.

[70] D. DRASDO, S. HOEHME, M. BLOCK.*On the Role of Physics in the Growth and Pattern Formation of Multi-Cellular Systems: What can we Learn from Individual-Cell Based Models?*, in "Journal of Statistical Physics", 2007, vol. 128, n^o 1-2, p. 287–345, http://dx.doi.org/10.1007/s10955-007-9289-x.

[71] 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.

[72] A. FRIEBEL, J. NEITSCH, T. JOHANN, S. HAMMAD, D. DRASDO, S. HOEHME.*TiQuant: software for tissue analysis, quantification and surface reconstruction*, in "Bioinformatics", June 2015, vol. 31, n^o 19, p. 3234-3236 [*DOI :* 10.1093/BIOINFORMATICS/BTV346], https://hal.inria.fr/hal-01257137.

[73] J. GALLE, G. AUST, G. SCHALLER, T. BEYER, D. DRASDO.*Individual cell-based models of the spatial-temporal organization of multicellular systems–achievements and limitations*, in "Cytometry. Part A : the journal of the International Society for Analytical Cytology", July 2006, vol. 69, n^o 7, p. 704–10 [*DOI :* 10.1002/CYTO.A.20287], http://www.ncbi.nlm.nih.gov/pubmed/16807896.

[74] 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.

[75] S. HAMMAD, S. HOEHME, A. FRIEBEL, I. VON RECKLINGHAUSEN, A. OTHMAN, B. BEGHER-TIBBE, R. REIF, P. GODOY, T. JOHANN, A. VARTAK, K. GOLKA, P. O. BUCUR, E. VIBERT, R. MARCHAN, B. CHRIST, S. DOOLEY, C. MEYER, I. ILKAVETS, U. DAHMEN, O. DIRSCH, J. BÖTTGER, R. GEBHARDT, D. DRASDO, J. G. HENGSTLER.*Protocols for staining of bile canalicular and sinusoidal networks of human, mouse and pig livers, three-dimensional reconstruction and quantification of tissue microarchitecture by image processing and analysis.*, in "Archives of Toxicology", May 2014, vol. 88, n^o 5, p. 1161-1183 [*DOI :* 10.1007/S00204-014-1243-5], https://hal.inria.fr/hal-01110657.

[76] 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.

[77] P.-E. JABIN, G. RAOUL.*On selection dynamics for competitive interactions*, in "Journal of Mathematical Biology", Oct 2010, vol. 63, n^o 3, p. 493–517, http://dx.doi.org/10.1007/s00285-010-0370-8.

[78] 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.

[79] 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.

[80] 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.

[81] A. LORZ, S. MIRRAHIMI, B. PERTHAME.*Dirac mass dynamics in a multidimensional nonlocal parabolic equation*, in "Communications in Partial Differential Equations", 2011, vol. 36, p. 1071-1098 [*DOI :* 10.1080/03605302.2010.538784], https://hal.archives-ouvertes.fr/hal-00533693.

[82] S. MIRRAHIMI, B. PERTHAME, J. Y. WAKANO.*Evolution of species trait through resource competition*, in "Journal of Mathematical Biology", June 2012, vol. 64, n^o^ 7, p. 1189-1223, 28 pages [*DOI :* 10.1007/s00285-011-0447-z], https://hal.archives-ouvertes.fr/hal-00566888.

[83] S. MISCHLER, J. SCHER.*Spectral analysis of semigroups and growth-fragmentation equations*, October 2013, working paper or preprint, https://hal.archives-ouvertes.fr/hal-00877870.

[84] C. POUCHOL.*Modelling interactions between tumour cells and supporting adipocytes in breast cancer*, UPMC, September 2015, https://hal.inria.fr/hal-01252122.

[85] M. RADSZUWEIT, M. BLOCK, J. G. HENGSTLER, E. SCHÖLL, D. DRASDO.*Comparing the growth kinetics of cell populations in two and three dimensions*, in "Phys. Rev. E", May 2009, vol. 79, 051907, http://link.aps.org/doi/10.1103/PhysRevE.79.051907.

[86] A. RAVASIO, I. CHEDDADI, T. CHEN, T. PEREIRA, H. T. ONG, C. BERTOCCHI, A. BRUGUES, A. JACINTO, A. J. KABLA, Y. TOYAMA, X. TREPAT, N. GOV, L. NEVES DE ALMEIDA, B. LADOUX.*Gap geometry dictates epithelial closure efficiency*, in "Nature Communications", July 2015, vol. 6, 6:7683 [*DOI :* 10.1038/NCOMMS8683], https://hal.archives-ouvertes.fr/hal-01245750.

[87] 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.

[88] M. TANG, N. VAUCHELET, I. CHEDDADI, I. VIGNON-CLEMENTEL, D. DRASDO, B. PERTHAME.*Composite waves for a cell population system modelling tumor growth and invasion*, in "Chinese Annals of Mathematics - Series B", 2013, vol. 34B, n^o^ 2, p. 295-318 [*DOI :* 10.1007/s11401-007-0001-x], https://hal.archives-ouvertes.fr/hal-00685063.

[89] P. VAN LIEDEKERKE, M. M. PALM, N. JAGIELLA, D. DRASDO.*Simulating tissue mechanics with agent-based models: concepts, perspectives and some novel results*, in "Computational Particle Mechanics", Nov 2015, vol. 2, n^o^ 4, p. 401–444, http://dx.doi.org/10.1007/s40571-015-0082-3.

[90] S. R. K. VEDULA, G. PEYRET, I. CHEDDADI, T. CHEN, A. BRUGUÉS, H. HIRATA, H. LOPEZ-MENENDEZ, Y. TOYAMA, L. NEVES DE ALMEIDA, X. TREPAT, C. T. LIM, B. LADOUX.*Mechanics of epithelial closure over non-adherent environments*, in "Nature Communications", 2015, vol. 6, Article number: 6111 [*DOI :* 10.1038/NCOMMS7111], http://hal.upmc.fr/hal-01298859.

[91] O. WOLKENHAUER, C. AUFFRAY, O. BRASS, J. CLAIRAMBAULT, A. DEUTSCH, D. DRASDO, F. GERVASIO, L. PREZIOSI, P. MAINI, A. MARCINIAK-CZOCHRA, C. KOSSOW, L. KUEPFER, K. RATEITSCHAK, I. RAMIS-CONDE, B. RIBBA, A. SCHUPPERT, R. SMALLWOOD, G. STAMATAKOS, F. WINTER, H. BYRNE.*Enabling multiscale modeling in systems medicine: From reactions in cells to organ physiology*, in "Genome Medicine", May 2014, vol. 6, n^{o} 21, 3, https://hal.inria.fr/hal-01109002.

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:

6.1.1. - Continuous Modeling (PDE, ODE)
6.1.2. - Stochastic Modeling (SPDE, SDE)
6.1.4. - Multiscale modeling
6.1.5. - Multiphysics modeling
6.2.1. - Numerical analysis of PDE and ODE
6.2.2. - Numerical probability
6.2.3. - Probabilistic methods
6.2.4. - Statistical methods
6.2.7. - High performance computing
6.3.1. - Inverse problems
6.3.4. - Model reduction
6.4.1. - Deterministic control
7.13. - Quantum algorithms

Other Research Topics and Application Domains:

1.1.2. - Molecular biology
4.3.4. - Solar Energy
5.3. - Nanotechnology
5.5. - Materials
9.4.2. - Mathematics
9.4.3. - Physics
9.4.4. - Chemistry

1. Members

Research Scientists

Claude Le Bris [Team leader, ENPC, HDR]
Antoine Levitt [Inria, Researcher]
Mathias Rousset [Inria, Researcher, until Aug 2016, HDR]
Sébastien Boyaval [ENPC]
Éric Cancès [ENPC, HDR]
Virginie Ehrlacher [ENPC]
Frédéric Legoll [ENPC, HDR]
Tony Lelièvre [ENPC, HDR]
Gabriel Stoltz [ENPC, HDR]

Faculty Members

Ludovic Chamoin [ENS Cachan, Associate Professor, *en délégation*, until Aug 2016, HDR]
Dorian Le Peutrec [Univ. Paris XI, Faculty Member, Associate Professor, *en délégation*, until Jul 2016]
Pierre-André Zitt [Univ. Paris Est, Associate Professor, *en délégation*, until Aug 2016, HDR]

Technical Staff

Cédric Doucet [Inria SED, Part-time, until March 2016]

PhD Students

Grégoire Ferré [ENPC, from Oct 2016]
Marc Josien [Min. Ecologie]
Boris Nectoux [ENPC]
Pierre-Loïk Rothé [ENPC, from Oct 2016]
Julien Roussel [ENPC]
Laura Silva Lopes [ENPC, from Oct 2016]
Pierre Terrier [Min. Ecologie]
François Madiot [ENPC, until Dec 2016]

Post-Doctoral Fellows

Michaël Bertin [Univ. Paris Est, until Feb 2016]
Giacomo Di Gesù [ENPC, until Aug 2016]
Pierre Monmarché [Inria, from Oct 2016]
Thomas Hudson [Inria, until Aug 2016]

Administrative Assistants

David Dinis [Inria, until Apr 2016]
Sarah Le [Inria, from Jul 2016]

Other

Frédérique Noël [Intern Inria, from May to Sep 2016]

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], [5], [6] that other scientists may consult in order to enter the field.

3. Research Program

3.1. Research Program

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. Such problems are described in the following sections.

4. Application Domains

4.1. 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.

4.2. Electronic structure of large systems

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.

4.3. 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 when the proton mass is the reference mass.

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 0.1 μs.

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?"

5. New Software and Platforms

5.1. SIMOL

KEYWORDS: C++ - Statistical physics - Quantum chemistry - Molecular simulation
FUNCTIONAL DESCRIPTION

SIMOL (SIMulation of MOLecular systems) is a software written in C++. It is a research code aimed at testing new ideas and algorithms, and provides a unified development platform for the members of the project-team. It is composed of three parts: a common core of input/output functions, linear algebra, random number generators, etc; and two specific applicative branches: one for computational statistical physics and one for quantum chemistry. The methods implemented for computational statistical physics are based on discretizations of ergodic stochastic differential equations such as the Langevin dynamics and its overdamped limit. The systems that can be simulated range from a single isolated particle to Lennard-Jones fluids. For quantum chemistry, the building block is the Hartree-Fock model, solved via fixed-point iterations; and various refinements including greedy methods.

- Contact: Gabriel Stoltz
- URL: https://gitlab.inria.fr/matherials/simol

6. New Results

6.1. Electronic structure calculations

Participants: Éric Cancès, Virginie Ehrlacher, Claude Le Bris, Antoine Levitt, Gabriel Stoltz.

In electronic structure calculation as in most of our scientific endeavors, we pursue a twofold goal: placing the models on a sound mathematical grounding, and improving the numerical approaches.

6.1.1. Molecular systems

The work of the project-team on molecular systems has focused on advanced approaches for the computation of the electronic state of molecular systems, including the effects of electronic correlation and of the environment.

In [12], E. Cancès, D. Gontier (former PhD student of the project-team, now at Université Paris Dauphine) and G. Stoltz have analyzed the GW method for finite electronic systems. This method enables the computation of excited states. To understand it, a first step is to provide a mathematical framework for the usual one-body operators that appear naturally in many-body perturbation theory. It is then possible to study the GW equations which construct an approximation of the one-body Green's function, and give a rigorous mathematical formulation of these equations. With this framework, results can be established for the well-posedness of the GW_0 equations, a specific instance of the GW model. In particular, the existence of a unique solution to these equations is proved in a perturbative regime.

Implicit solvation models aim at computing the properties of a molecule in solution (most chemical reactions indeed take place in the liquid phase) by replacing all the solvent molecules except the ones strongly interacting with the solute, by an effective continuous medium accounting for long-range electrostatics. E. Cancès, Y. Maday (Paris 6), and B. Stamm (Paris 6) have recently introduced a very efficient domain decomposition method for the simulation of large molecules in the framework of the so-called COSMO implicit solvation models. In collaboration with F. Lipparini (UPMC), B. Mennucci (Department of Chemistry, University of Pisa) and J.-P. Picquemal (Paris 6), they have implemented this algorithm in widely used computational software products (Gaussian and Tinker). E. Cancès, Y. Maday, F. Lipparini and B. Stamm have also extended this approach to the more complex polarizable continuum model (PCM).

C. Le Bris has pursued his collaboration with Pierre Rouchon (Ecole des Mines de Paris) on the study of high dimensional Lindblad type equations at play in the modelling of open quantum systems. In order to complement and better understand the numerical approaches developed in the past couple of years, some theoretical aspects are now under study, in particular regarding the well-posedness of the equations and their convergence in the long time limit.

6.1.2. Crystals and solids

Periodic systems are mathematically treated using Bloch theory, raising specific theoretical and numerical issues.

A. Bakhta (CERMICS) and V. Ehrlacher are working on the design of an efficient numerical method to solve the inverse band structure problem. The aim of this work is the following: given a set of electronic bands partially characterizing the electronic structure of a crystal, is it possible to recover the structure of a material which could achieve similar electronic properties? The main difficulty in this problem relies in the practical resolution of an associated optimization problem with numerous local optima.

As an external collaborator of the MURI project on 2D materials (PI: M. Luskin), E. Cancès has started a collaboration with P. Cazeaux and M. Luskin (University of Minnesota) on the computation of the electronic and optical properties of multilayer 2D materials. Together with E. Kaxiras (Harvard) and members of his group, they have developped a perturbation method for computing the Kohn-Sham density of states of incommensurate bilayer systems. They have also adapted the C*-algebra framework for aperiodic solids introduced by J. Bellissard and collaborators, to the case of tight-binding models of incommensurate (and possibly disordered) multilayer systems [36].

É. Cancès, A. Levitt and G. Stoltz, in collaboration with G. Panati (Rome) have proposed a new method for the computation of Wannier functions, a standard post-processing of density functional theory computations [38]. Compared to previous approaches, it does not require an initial guess for the shape of the Wannier functions, and is therefore more robust.

6.1.3. Numerical analysis

Members of the project-team have worked on the numerical analysis of partial differential equations arising from electronic structure theory.

E. Cancès and N. Mourad (CERMICS) have clarified the mathematical framework underlying the construction of norm-conserving semilocal pseudopotentials for Kohn-Sham models, and have proved the existence of optimal pseudopotentials for a family of optimality criteria.

E. Cancès has pursued his long-term collaboration with Y. Maday (UPMC) on the numerical analysis of electronic structure models. Together with G. Dusson (UMPC), B. Stamm (UMPC), and M. Vohralík (Inria), they have designed a new post processing method for planewave discretizations of nonlinear Schrödinger equations, and used it to compute sharp *a posteriori* error estimators for both the discretization error and the algorithmic error (convergence threshold in the iterations on the nonlinearity). They have then extended this approach to the Kohn-Sham model. In parallel, they have derived a posteriori error estimates for conforming numerical approximations of the Laplace eigenvalue problem with homogeneous Dirichlet boundary conditions [37]. In particular, upper and lower bounds for any simple eigenvalue are established. These bounds are guaranteed, fully computable, and converge with the optimal rate to the exact eigenvalue.

A. Levitt, in collaboration with X. Antoine and Q. Tang (Nancy), has proposed a new numerical method to compute the ground state of rotating Bose-Einstein condensates [31]. This method combines a nonlinear conjugate gradient method with efficient preconditionners. Compared to the state of the art (implicit timestepping on the imaginary-time equation), gains of one to two orders of magnitude are achieved.

6.2. Complex fluids

Participant: Sébastien Boyaval.

The aim of the research performed in the project-team about complex fluids is to guide the mathematical modelling of gravity flows with a free-surface for application to the hydraulic engineering context, and to account for non-Newtonian rheologies in particular (like in mudflows for instance). On the one hand, thin-layer (reduced) models have long been favored, and one current trend aims at incorporating non-Newtonian effects [10]. This has stimulated some research about a new hyperbolic PDE system [35]. On the other hand, there is currently a strong need to perform full 3D numerical simulations using new non-Newtonian models in complex geometries with a view to comparing them with physical observations ; this is an ongoing work, in the framework of the ANR project SEDIFLO with E. Audusse (Paris 13), A. Caboussat (Genève), A. Lemaitre (ENPC), M.Parisot (Inria).

6.3. Homogenization

Participants: Michaël Bertin, Ludovic Chamoin, Virginie Ehrlacher, Thomas Hudson, Marc Josien, Claude Le Bris, Frédéric Legoll, François Madiot, Pierre-Loïk Rothé.

6.3.1. Deterministic non-periodic systems

The homogenization of (deterministic) non-periodic systems is a well-known topic. Although well explored theoretically by many authors, it has been less investigated from the standpoint of numerical approaches (except in the random setting). In collaboration with X. Blanc (Paris 7) and P.-L. Lions (Collège de France), C. Le Bris has introduced a possible theory, giving rise to a numerical approach, for the simulation of multiscale non-periodic systems. In former publications, several theoretical aspects have been considered, for the case of linear elliptic equations in divergence form. In the context of the PhD thesis of M. Josien, new issues are being explored, including the rate of convergence of the approximation, along with the convergence of the Green functions associated to the problems under consideration. The studies are motivated by several practically relevant problems, in particular the problem of defects in periodic structures and the "twin boundaries" problem in materials science. Also, some other equations than linear elliptic equations in divergence form have been considered lately. The case of advection-diffusion equations is currently examined. In addition, one ongoing work, in collaboration with P. Souganidis (University of Chicago) and P. Cardaliaguet (Université Paris-Dauphine), considers the non-periodic setting for Hamilton-Jacobi type equations.

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 both 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. In practice, the space $\mathbb{R}^d$ is truncated to some bounded domain, on which the corrector problem is numerically solved. In turn, this yields a converging approximation of the homogenized tensor, which happens to be a random matrix.

Over the past years, the project-team has proposed several variance reduction techniques, which have been reviewed and compared to one another in [9], [20]. In particular, in [23], C. Le Bris, F. Legoll and W. Minvielle have investigated the possibility to use a variance reduction technique based on computing the corrector equation only for selected environments. These environments are chosen based on the fact that their statistics in the finite supercell matches the statistics of the materials in the infinite supercell. The efficiency of the approach has been demonstrated for various types of random materials, including composite materials with randomly located inclusions.

Besides the (averaged) behavior of the oscillatory solution u_ε on large space scales (which is given by the homogenized limit u_* of u_ε), another question of interest is to understand how much u_ε fluctuates around its coarse approximation u_*. This question will be explored in the PhD thesis of P.-L. Rothé, which started in October 2016.

Still another question investigated in the project-team is to find an alternative to standard homogenization techniques when the latter are difficult to use in practice, because not all the information on the microscopic medium is available. Following an interaction with A. Cohen (Paris 6), C. Le Bris, F. Legoll and S. Lemaire (post-doc in the project-team until 2015), have shown that the constant matrix that “best” (in a sense made precise in [44]) represents the oscillatory matrix describing the medium converges to the homogenized matrix in the limit of infinitely rapidly oscillatory coefficients. Furthermore, the corresponding optimization problem can be efficiently solved using standard algorithms and yield accurate approximation of the homogenized matrix. It has also been shown that it is possible to construct, in a second stage, approximations to the correctors, in order to recover an approximation of the *gradient* of the solution. The details are now available in [44].

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).

The MsFEM has been introduced almost 20 years ago. However, even in simple deterministic cases, there are still some open questions, for instance concerning multiscale advection-diffusion equations. Such problems are possibly advection dominated and a stabilization procedure is therefore required. How stabilization interplays with the multiscale character of the equation is an unsolved mathematical question worth considering for numerical purposes.

In the context of the PhD thesis of F. Madiot, current efforts are focused on the study of an advection-diffusion equation with a dominating convection in a perforated domain. The multiscale character of the problem stems here from the geometry of the domain. On the boundary of the perforations, we set either homogeneous Dirichlet or homogeneous Neumann conditions. In the spirit of the work [21], the purpose of our ongoing work is to investigate, on perforated domains, the behavior of several variants of the Multiscale Finite Element method, specifically designed to address multiscale advection-diffusion problems in the convection-dominated regime. Generally speaking, the idea of the MsFEM is to perform a Galerkin approximation of the problem using specific basis functions that are precomputed (in an offline stage) and adapted to the problem considered. All the variants considered are based upon local functions satisfying weak continuity conditions in the Crouzeix-Raviart sense on the boundary of mesh elements. Several possibilities for the basis functions have been examined (for instance, they may or may not encode the convection field). Depending on how basis functions are defined, stabilization techniques (such as SUPG) may be required. The type of boundary conditions on the perforations (either homogeneous Dirichlet or homogeneous Neumann boundary conditions) drastically affects the nature of the flow, and therefore the conclusions regarding which numerical approach is best to adopt. In short, homogeneous Dirichlet boundary conditions on the perforations damp the effect of advection, making the flow more stable than it would be in the absence of perforations, while this is not the case for homogeneous Neumann boundary conditions. This intuitive fact is investigated thoroughly at the numerical level, and particularly well exemplified, at the theoretical level, by the comparison of the respective homogenization limits.

Advection-diffusion equations that are both non-coercive and advection-dominated have also been considered (in a single-scale framework). Many numerical approaches have been proposed in the literature to address such difficult cases. C. Le Bris, F. Legoll and F. Madiot have proposed an approach based on the invariant measure associated to the original equation. The approach has been summarized in [22], and extensively described, analyzed and numerically tested in [45]. It is shown there that this approach allows for an unconditionally well-posed finite element approximation, and that it can be stable, as accurate as, and more robust than classical stabilization approaches.

Most of the 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 (from ENS Cachan, from September 2014 to August 2016), 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 thus needs to be adapted to that context. The proposed approach yields very interesting results, and provides an accurate and robust estimation of the global error. The approach has also been adapted towards the design of adaptive algorithms for specific quantities of interest (in the so-called “goal-oriented” setting), and towards the design of model reduction approaches (such as the Proper Generalized Decomposition (PGD)) in the specific context of multiscale problems. The work will be reported on in a forthcoming publication in preparation.

6.3.4. Discrete systems and their thermodynamic limit

In collaboration with X. Blanc (Paris 7), M. Josien has studied the macroscopic limit of a chain of atoms governed by Newton's equations. It is known from the works of X. Blanc (Paris 7), C. Le Bris and P.-L. Lions (Collège de France) that this limit is the solution of a nonlinear wave equation, as long as the solution remains smooth. For a large class of interaction potentials, X. Blanc and M. Josien have shown in [34], theoretically and numerically, that, if the distance between particles remains bounded, the above description in terms of a non-linear wave equation equation no longer holds when there are shocks. Indeed, the system of particles produces dispersive waves that are not predicted by the nonlinear wave equation.

6.3.5. Dislocations

Plastic properties of crystals are due to dislocations, which are thus objects of paramount importance in materials science. The geometrical shape of dislocations may be described by (possibly time-dependent) nonlinear integro-differential equations (e.g. Weertman's equation and the dynamical Peierls-Nabarro equation), involving non-local operators. In collaboration with Y.-P. Pellegrini (CEA), M. Josien has first focused on the steady state regime (where the equation of interest is the Weertman equation), and has designed an efficient numerical method for approximating its solution. The approach is based on a splitting strategy between the nonlinear local terms (which are integrated in real space) and the linear nonlocal terms (which are integrated in Fourier space). Current efforts are devoted to the simulation of physically relevant test-cases, with the aim of comparing the obtained numerical results with results of the physics literature. The work will be reported on in a forthcoming publication in preparation.

6.4. Computational Statistical Physics

Participants: Grégoire Ferré, Giacomo Di Gesù, Thomas Hudson, Dorian Le Peutrec, Frédéric Legoll, Tony Lelièvre, Pierre Monmarché, Boris Nectoux, Julien Roussel, Mathias Rousset, Laura Silva Lopes, Gabriel Stoltz, Pierre Terrier, Pierre-André Zitt.

In [24], T. Lelièvre and G. Stoltz have given an overview of state-of-the art mathematical techniques which are useful to analyze and quantify the efficiency of the algorithms used in molecular dynamics, both for sampling thermodynamic quantities (canonical averages and free energies) and dynamical quantities (transition rates, reactive paths and transport coefficients).

6.4.1. Improved sampling methods

This section is devoted to recent methods which have been proposed in order to improve the sampling of the canonical distribution by modifying the Langevin or overdamped Langevin dynamics, or its discretization. Two general strategies have been pursued by the project-team along these lines: (i) constructing dynamics with better convergence rate and hence smaller statistical errors; (ii) the stabilization of discretization schemes by Metropolis procedures in order to allow for larger timesteps while maintaining acceptable rejection rates.

A first approach to obtaining better convergence rates consists in modifying the drift term in the overdamped-Langevin dynamics, in order to improve the rate of converge to equilibrium. This method was considered by T. Lelièvre with A. Duncan and G.A. Pavliotis (Imperial College) in [14]. It is shown that nonreversible dynamics always result in a smaller asymptotic variance (statistical error). The efficiency of the whole algorithm crucially depends on the time discretization, which may induce some bias (deterministic error). It is shown on some examples how to balance the two errors (bias and statistical errors) in order to obtain an efficient algorithm.

The discretization of overdamped Langevin dynamics, using schemes such as the Euler-Maruyama method, may lead to numerical methods that are unstable when the forces are non-globally Lipschitz. One way to stabilize numerical schemes is to superimpose some acceptance/rejection rule, based on a Metropolis-Hastings criterion for instance. However, rejections perturb the dynamical consistency of the resulting numerical method with the reference dynamics. G. Stoltz and M. Fathi (Toulouse) present in [15] some modifications of the standard stabilization of discretizations of overdamped Langevin dynamics by a Metropolis-Hastings procedure, which allow to either improve the strong order of the numerical method, or to reduce the bias in the estimation of transport coefficients characterizing the effective dynamical behavior of the dynamics.

The sampling properties of Langevin dynamics can be improved by considering more general non-quadratic kinetic energies. This was accomplished in [26], where G. Stoltz, with S. Redon and Z. Trstanova (Inria Grenoble), have studied the properties of Langevin dynamics with general, non-quadratic kinetic energies $U(p)$, showing in particular the ergodicity of the dynamics even when the kinetic force ∇U vanishes on open sets and proving linear response results for the variance of the process for kinetic energies which correspond to the so-called adaptively restrained particle simulations. This work has been complemented by [51], where G. Stoltz and Z. Trstanova provide accurate numerical schemes to integrate the modified Langevin dynamics with general kinetic energies, with possibly non globally Lipschitz momenta.

6.4.2. Adaptive methods

When direct sampling methods fail, it is worth considering importance sampling strategies, where the slowest direction is described by a reaction coordinate ξ, and the invariant measure is biased by (a fraction of) the free energy associated with ξ.

The first group of results along these lines concerns the study of adaptive biasing methods to compute free energy differences:

- The result obtained by H. Al Rachid (CERMICS) in collaboration with T. Lelièvre and R. Talhouk (Beirut) on the existence of a solution to the non linear Fokker Planck equation associated to the ABF process has been published, see [7].
- T. Lelièvre and G. Stoltz, together with G. Fort (Toulouse) and B. Jourdain (CERMICS), have studied the well-tempered metadynamics and many variants of this method in [41]. This dynamics can be seen as some extension of the so-called self-healing umbrella sampling method, with a partial biasing of the dynamics only. In particular, the authors propose a version which leads to much shorter exit times from metastable states (accelerated well-tempered metadynamics).

The project-team also works on adaptive splitting techniques, which forces the exploration in the direction of increasing values of the reaction coordinate. In [29], T. Lelièvre, together with C. Mayne, K. Schulten and I. Teo (Univ. Illinois), has reported on the calculation of the unbinding rate of the benzamidine-trypsin system using the Adaptive Multilevel Splitting algorithm. This is the first "real-life" test case for the adaptive multilevel splitting. In [11], T. Lelièvre and M. Rousset, in collaboration with C.E. Bréhier (Lyon), M. Gazeau (Créteil) and L. Goudenège (Centrale), propose a generalization of the Adaptive Multilevel Splitting method for discrete-in-time processes. It is shown how to make the estimator unbiased. Numerical experiments illustrate the performance of the method.

6.4.3. Coarse-graining and reduced descriptions

A fully atomistic description of physical systems leads to problems with a very large of unknowns, which raises challenges both on the simulation of the system and the interpretation of the results. Coarse-grained approaches, where complex molecular systems are described by a simplified model, offer an appealing alternative.

F. Legoll and T. Lelièvre, together with S. Olla (Dauphine), have proposed an analysis of the error introduced when deriving an effective dynamics for a stochastic process in large dimension on a few degrees of freedom using a projection approach à la Zwanzig [48]. More precisely, a pathwise error estimate is obtained, which is an improvement compared to a previous result by F. Legoll and T. Lelièvre where only the marginal in times were considered.

Another line of research concerns dissipative particle dynamics, where a complex molecule is replaced by an effective mesoparticle. The work [17] by G. Stoltz, together with A.-A. Homman and J.-B. Maillet (CEA), on new parallelizable numerical schemes for the integration of Dissipative Particle Dynamics with Energy conservation, has been published. Together with G. Faure and J.-B. Maillet, G. Stoltz has proposed in [16] a new formulation of smoothed dissipative particle dynamics, which can be seen as some meshless discretization of the Navier–Stokes equation perturbed by some random forcing arising from finite size effects of the underlying mesoparticles. The reformulation, in terms of internal energies rather than internal entropies, allows for a simpler and more efficient simulation, and also opens the way for a coupling with standard dissipative particle dynamics models.

G. Stoltz also suggested in [50] a new numerical integrator for DPDE which is more stable than all the previous integrators. The key point is to reduce the stochastic part of the dynamics to elementary one-dimensional dynamics, for which some Metropolis procedure can be used to prevent the appearance of negative energies at the origin of the instability of the numerical methods.

During the post-doctoral stay of I.G. Tejada (ENPC), G. Stoltz, F. Legoll and E. Cancès studied in collaboration with L. Brochard (ENPC) the derivation of a concurrent coupling technique to model fractures at the atomistic level by combining a reactive potential with a reduced harmonic approximation. The results have appeared in [28].

G. Stoltz and P. Terrier, in a joint work with M. Athènes, T. Jourdan (CEA) and G. Adjanor (EDF), have presented a coupling algorithm for cluster dynamics [52]. Rate equation cluster dynamics (RECD) is a mean field technique where only defect concentrations are considered. It consists in solving a large set of ODEs (one equation per cluster type) governing the evolution of the concentrations. Since clusters might contain up to million of atoms or defects, the number of equations becomes very large. Therefore solving such a system of ODEs becomes computationally prohibitive as the cluster sizes increase. Efficient deterministic simulations propose an approximation of the equations for large clusters by a single Fokker-Planck equation. The proposed coupling algorithm is based on a splitting of the dynamics and combines deterministic and stochastic approaches. In addition, F. Legoll and G. Stoltz have proposed in [19], with T. Jourdan (CEA) and L. Monasse (CERMICS), a new method for numerically integrating the Fokker–Planck approximation of large cluster dynamics.

6.4.4. Eyring–Kramers formula and quasi-stationary distributions

G. Di Gesù, T. Lelièvre and B. Nectoux, together with D. Le Peutrec, have explored the interest of using the quasi-stationary distribution approach in order to justify kinetic Monte Carlo models, and more precisely their parameterizations using the Eyring-Kramers formulas, which provide a simple rule to compute transition rates from one state to another [13]. The paper is essentially a summary of the results which have been obtained during the first two years of the PhD of B. Nectoux. A preprint with detailed proofs of these results is in preparation.

In [33], G. Di Gesù has studied with N. Berglund (Orléans) and H. Weber (Warwick) the spectral Galerkin approximations of an Allen-Cahn equation over the two-dimensional torus perturbed by weak space-time white noise. They show sharp upper and lower bounds on the transition times from a neighborhood of the stable configuration -1 to the stable configuration 1 in the small noise regime. These estimates are uniform in the discretization parameter, suggesting an Eyring-Kramers formula for the limiting renormalized stochastic PDE.

6.4.5. Functional inequalities and theoretical aspects

The interplay between probability theory and analysis in statistical physics is best exemplified by the functional analysis study of the semigroups associated with the generator of the stochastic processes under consideration. These generators are elliptic or hyperbolic operators. Several functional-analytic results were obtained by the team on problems of statistical physics.

D. Le Peutrec has derived Brascamp-Lieb type inequalities for general differential forms on compact Riemannian manifolds with boundary from the supersymmetry of the semiclassical Witten Laplacian [47]. These results imply the usual Brascamp-Lieb inequality and its generalization to compact Riemannian manifolds without boundary.

T. Hudson has considered with C. Hall (Oxford) and P. van Meurs (Univ. Kanazawa, Japan) the minimization of the potential energy of N particles mutually interacting under a repulsive interaction potential with a certain algebraic decay assumption [42]. A major novelty of the approach is that it does not assume a finite range of interaction. The main focus of the work is on characterizing the boundary behavior of minimizers in the limit where the number of particles N tends to infinity with a constant density of particles per unit volume.

G. Di Gesù has studied with M. Mariani (Rome) the small temperature limit of the Fisher information of a given probability measure with respect to the canonical measure with density proportional to $\exp(-\beta V)$ [39]. The expansion reveals a hierarchy of multiple scales reflecting the metastable behavior of the underlying overdamped Langevin dynamics: distinct scales emerge and become relevant depending on whether one considers probability measures concentrated on local minima of V, probability measures concentrated on critical points of V, or generic probability measures on $\mathbb{R}^d$.

6.5. Various topics

Participants: Virginie Ehrlacher, Tony Lelièvre, Antoine Levitt.

In [18], T. Lelièvre has explored with J. Infante Acevedo (CERMICS) the interest of using the greedy algorithm (also known as the Proper Generalized Decomposition) for the pricing of basket options.

V. Ehrlacher and D. Lombardi have developped a new tensor-based numerical method for the resolution of kinetic equations [40] in a fully Eulerian framework.This theory enables to describe a large system of particles by a distribution function $f(x, v, t)$ that encodes the probability of finding a particle at time t, position $x \in \mathbb{R}^3$ and velocity $v \in \mathbb{R}^3$. These systems are used to model the behavior of plasma of the transport of electrons in semiconductors for instance. However, simulating such systems involves the resolutions of problems defined on $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}_+$, which leads to very high-dimensional systems. The new approach developped in [40] circumvents the curse of dimensionality for these systems, by efficiently adapting the rank of the decomposition of the solution through time. Encouraging preliminary numerical results have been obtained on $3D \times 3D$ systems.

A system of cross-diffusion equations has been proposed in [32] by A. Bakhta and V. Ehrlacher for the modelling of a Physical Vapor Deposition (PVD) process used for the manufacturing of thin film solar cells. This process works as follows: a substrate wafer is introduced in a hot chamber where different chemical species are injected under gaseous form. These different species deposit on the surface of the substrate, so that a thin film layer grows upon the surface of the substrate. Two phenomena have to be taken into account in the modelling: 1) the evolution of the thickness of the thin film layer; 2) the diffusion of the various species inside the bulk. The existence of a weak solution to the system proposed in [32] has been proved, along with the existence of optimal fluxes to be injected in the chamber in order to obtain target concentration profiles at the end of the process. The long-time behavior of solutions has been studied in the case when the injected fluxes are constant. Moreover, numerical results on the simulation of this system have been compared with experimental data given by IRDEP on CIGS (Copper, Indium, Gallium, Selenium) solar cells. The project is a collaboration with IRDEP.

A. Levitt, in collaboration with F. Aviat, L. Lagardère, Y. Maday, J.-P. Piquemal (UPMC), B. Stamm (Aachen), P. Ren (Texas) and J. Ponder (Saint Louis), has proposed a new method for the solution of the equations of polarizable force fields [8]. Previous methods had to solve a linear system to high accuracy in order for the energy to be preserved in simulations. The method presented, based on an explicit differentiation of the energy produced by the truncated iterative method, is able to conserve the energy even with loose convergence criteria, thus allowing stable and fast simulations at degraded accuracy.

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, IRDEP, EDF, IFPEN. 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's SEDIFLO project, funded by ANR as a JCJC (Jeunes Chercheuses Jeunes Chercheurs) grant, has started investigating new numerical models of solid transport in rivers that include new non-Newtonian terms.
- E. Cancès is involved in the ANR BECASIM, which is concerned with the numerical simulation of Bose-Einstein condensates. This ANR has been accepted in June 2012, and is coordinated by I. Danaila (Université de Rouen).
- T. Lelièvre is member of the ANR-project "STAB" (PI: I. Gentil, Université de Lyon).
- F. Legoll is a member of the ANR project CINE-PARA (PI: Y. Maday, Paris 6)
- The ANR COSMOS (PI: G. Stoltz) 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.

In addition, the project-team is participating in

- the GdR CORREL (correlated methods in electronic structure computations),
- the GdR EGRIN (gravity flows),
- the GdR MASCOT-NUM (stochastic methods for the analysis of numerical codes),
- the GdR Maths-entreprise (math/industry collaboration),
- the GdR DYNQUA (time evolution of quantum systems, with applications to transport problems, nonequilibrium systems, etc.),
- the GdR REST (theoretical spectroscopy),
- the GdR CHOCOLAS (experimental and numerical study of shock waves).

The project-team is involved in two Labex, namely the Labex Bezout (started in 2011) and the Labex MMCD (started in 2012).

8.2. European Initiatives

The ERC consolidator Grant MSMATH (ERC Grant Agreement number 614492, PI T. Lelièvre) is running (it started in June 2014).

8.3. International Initiatives

The *Germaine de Staël* grant to S. Boyaval (from CampusFrance Hubert-Curien program) has been renewed for 2017 to pursue the collaboration with A. Caboussat (Lausanne) about 3D numerical simulations of free-surface flows.

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, Université de Lyon and Inria Rennes.

9. Dissemination

9.1. Promoting Scientific Activities

E. Cancès

- is a member of the editorial boards of Mathematical Modelling and Numerical Analysis (2006-), SIAM Journal of Scientific Computing (2008-), Communications in Mathematical Sciences

(2011-), and SIAM MMS (2012-),

- is a member of the executive committee of the CEA-EDF-Inria schools in applied mathematics and computer science.

He has organized or co-organized:

- a CECAM workshop on the mathematical and numerical analysis of electronic structure models, Roscoff, Jul. 8-12, 2016,
- the MMM 2016 conference (Multiscale Modelling of Materials), Dijon, Oct. 10-14, 2016,
- an IPAM workshop on collective variables in quantum mechanics, Los Angeles, Nov. 14-18, 2016.

L. Chamoin has organized the mini-symposium "Verification of reduced models in computational mechanics" within the ECCOMAS 2016 conference, Greece, June 2016.

V. Ehrlacher has co-organized with A. Cousin the semester on "Uncertainty Quantification" in the framework of the IHP thematic semester on "Monte-Carlo methods" organized by B. Bouchard, E. Gobet and B. Jourdain.

C. Le Bris is editor-in-chief of Applied Mathematics Research Express (2003-). He 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-), COCV (Control, Optimization and Calculus of Variations) (2003-), Mathematics in Action (2008-), Nonlinearity (2005-) and Journal de Mathématiques Pures et Appliquées (2009-).

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 Cabinet of the High Commissioner for Atomic Energy,
- the “Comité d’experts” for the Fondation de Recherche pour l’Aéronautique et l’Espace,
- the “International Scientific Advisory Committee” of the Centre de Recherche Mathématique, Université de Montréal,
- the “Advisory Board” of the DFG Cluster of Excellence Engineering of Advanced Materials, Erlangen,
- the “International Scientific Advisory Board” of the DFG research center Matheon, Berlin,
- the “Conseil scientifique de la SMAI” (Scientific Council of the French Applied Maths Society),
- the International Mathematical Union Circle.

He has held a regular position of Visiting Professor at the University of Chicago.

F. Legoll

- is a member of the editorial board of SIAM MMS (2012-) and of ESAIM Proc (2012-),
- has co-organized the mini-symposium "Mathematical theory and computational techniques for multiscale materials modelling" within the MMM 2016 conference, Dijon, October 10-14, 2016 (with W. Curtin, C. Garcia-Cervera, J. Kermode, X. Li, A. Lozinski, M. Luskin and C. Ortner).

T. Lelièvre

- is editor-in-chief of ESAIM: Proceedings (with D. Chafai, P. Lafitte and C. Mouhot),
- is a member of the "Conseil d’Administration" of SMAI and Ecole des Ponts,
- has been a member of the ANR committee CES-40 "mathématiques et informatique",
- has co-organized the Journées EDP-Probas at Institut Henri Poincaré (with F. Malrieu),
- has co-organized the workshop "COmputational Statistics and MOlecular Simulation" in Paris, February 2-5th, 2016 (with A. Guyader and G. Stoltz),

- has co-organized the IHP conference on "Recent developments in numerical methods for model reduction", November 7-10th 2016. (with S. Perotto and G. Rozza),
- will co-organize the IPAM Long Program on "Complex High-Dimensional Energy Landscapes", September 11th - December 15th 2017 (with C. Clementi, G. Henkelman, R. Hennig, M. Luskin, N. Marom, P. Plechac and C. Schuette),
- will co-organize the ICTS program on "Large deviation theory in statistical physics: Recent advances and future challenges", August 14th - October 13th 2017 (with A. Ayyer, F. den Hollander, A. Dhar, J.P. Garrahan, C. Jarzynski, M. Krishnapur, S. Sabhapandit and H. Touchette).

G. Stoltz

- is a member of the scientific council of UNIT (Université Numérique Ingénierie et Technologie),
- will co-organize the IHP trimester "Stochastic Dynamics Out of Equilibrium", Spring 2017 (with G. Giacomin, S. Olla, E. Saada and H. Spohn).

9.2. Teaching - Supervision - Juries

The members of the project-team have taught the following courses:

- Licence: Outils mathématiques pour l'ingénieur, 15h, L3, École des Ponts (E. Cancès, V. Ehrlacher, M. Josien, F. Legoll, T. Lelièvre),
- Licence: Équations aux dérivées partielles et éléments finis, 15h, L3, École des Ponts (T. Hudson, F. Legoll, A. Levitt),
- Licence: Hydrodynamique numérique, 15h, L3, École des Ponts (S. Boyaval),
- Licence: Maths 1 et 2, 9h, L3, École des Mines (G. Stoltz),
- Licence: Mathématiques pour l'économie, 36h, L1, Dauphine (J. Roussel),
- Licence: Analyse et calcul scientifique, 30h, L3, Ecole des Ponts (T. Hudson, M. Josien, B. Nectoux, G. Stoltz),
- Master: Mécanique des milieux continus - partie solides, 14h, M1, ENS Cachan (L. Chamoin),
- Master: Ondes et chocs dans les structures, 8h, M1, ENS Cachan (L. Chamoin),
- Master: Mathématiques des modèles multiéchelles, 39h, M1, École des Ponts (F. Legoll),
- Master: Analyse et équations aux dérivées partielles, 36 h, M1, École des Ponts (T. Lelièvre),
- Master: Projet de département IMI, 12h, M1, École des Ponts (J. Roussel),
- Master: Projets de physique, 10h, M1, École des Ponts, France (A. Levitt, G. Stoltz),
- Master: Modélisation mathématique des vagues, 3h, École des Ponts (S. Boyaval),
- Master: Analyse de Fourier et applications, 16h, M1, École des Ponts (V. Ehrlacher, A. Levitt, G. Stoltz),
- Master: Approximation numérique et optimisation, 32h, École Polytechnique (E. Cancès, T. Lelièvre) ,
- Master: Analyse variationnelle des équations aux dérivées partielles, 32h, École Polytechnique (E. Cancès),
- Master: Contrôle des modèles et dualité, 24h, M2 Mathématiques et Applications, ENS Cachan (L. Chamoin),
- Master: Problèmes multi-échelles, 24h, M2 Mathématiques et Applications, Paris 6 (F. Legoll),
- Master: Méthodes variationnelles et théorie spectrale, 10h, M2 Mathématiques et Applications, Paris 6 (E. Cancès),
- Master: Méthodes numériques probabilistes, 24 h, M2 Mathématiques et Applications, Paris 6 (T. Lelièvre),
- Master: Introduction to computational statistical physics, 20h, M2 Mathématiques et Applications, Paris 6 (G. Stoltz).

The following PhD theses have been defended in the group at École des Ponts:

- Eddy Bernard, Université Paris-Est, defended on Nov. 25, 2016, supervised by G. Chambaud (Université Paris-Est) and E. Cancès,
- Jean-Léopold Vié, Université Paris-Est, École des Ponts and Ecole Polytechnique, defended on Dec. 16, 2016, supervised by G. Allaire (Polytechnique) and E. Cancès,
- Ahmed-Amine Homman, Multiscale methods for the simulation of shock and detonation waves, Université Paris-Est, École des Ponts and CEA/DAM, defended on June 16, 2016, supervised by G. Stoltz and J.-B. Maillet (CEA),
- François Madiot, Multiscale finite element methods for advection diffusion problems, Université Paris-Est, Ecole des Ponts ParisTech, defended on December 8, 2016, supervised by C. Le Bris and F. Legoll,
- Rémi Sainct, Study of instabilities in traffic models, defended on September 22, 2016, supervised by T. Lelièvre and X. Louis (IFSTTAR),

The following PhD theses are ongoing in the group at École des Ponts:

- Athmane Bakhta, Modélisation and simulation for photovoltaic applications, Université Paris-Est, École des Ponts, started October 1st, 2014, supervised by E. Cancès and T. Lelièvre, co-supervised by V. Ehrlacher,
- Amina Benaceur, Thèse CIFRE EDF, started January 1st, 2016, supervised by A. Ern, co-supervised by V. Ehrlacher, in collaboration with G. Blatman (EDF) and S. Meunier (EDF),
- Lingling Cao, Mathematical analysis of models of thermo-electronic transport, Université Paris-Est, École des Ponts, started November 1st, 2016, supervised by E. Cancès and G. Stoltz,
- Qiming Du, Mathematical analysis of splitting methods, École Doctorale Sciences Mathématiques de Paris Centre, started September 1st, 2016, supervised by A. Guyader (UPMC) and T. Lelièvre,
- Gérôme Faure, Multiscale methods for the simulation of shock and detonation waves, Université Paris-Est, École des Ponts and CEA/DAM, started November 1st 2014, supervised by G. Stoltz and J.-B. Maillet (CEA),
- Grégoire Ferré, Efficient sampling methods for nonequilibrium systems, Université Paris-Est, École des Ponts, started October 1st, 2016, supervised by G. Stoltz,
- Marc Josien, Multiscale approaches for materials science, started September 1st, 2015, supervised by C. Le Bris,
- Henri Louvin, Splitting methods and radioprotection, Ecole Doctorale PHENIICS, started September 1st, 2014, supervised by Check Diop (CEA) and T. Lelièvre,
- Boris Nectoux, Metastability and quasi stationary distribution, started November 1st, 2014, supervised by T. Lelièvre and E. Cancès,
- Julien Roussel, Variance reduction techniques for nonequilibrium systems, Université Paris-Est, École des Ponts, started September 1st, 2015, supervised by G. Stoltz,
- Pierre-Loik Rothé, Numerical methods for the estimation of fluctuations in multi-scale materials and related problems, started October 1st, 2016, supervised by F. Legoll,
- Laura Silva Lopes, Rare event simulation and applications to biological systems, started October 1st, 2016, supervised by J. Hénin (IBPC) and T. Lelièvre,
- Pierre Terrier, Reduced models for defect migration in metals, Université Paris-Est, École des Ponts and CEA Saclay, started September 1st, 2015, supervised by G. Stoltz and M. Athènes (CEA),

Project-team members have participated in the following PhD juries:

- S. Boyaval was in the jury for the PhD of Viljami Laurmaa ("An octree-based adaptive semi-Lagrangian free surface flow solver"), defended at EPFL in May 2016.
- S. Boyaval was in the jury for the PhD of Joubine Aghili ("Méthodes de discrétisation et de réduction de modèle pour des EDP à coefficients variables"), defended in Montpellier in December 2016.
- E. Cancès was a referee of the PhD of Maxime Morinière ("États résonants en théorie de perturbation à plusieurs corps"), defended in Grenoble in December 2016.
- V. Ehrlacher was in the jury for the PhD of Luca Nenna ("Numerical methods for Multi-Marginal Optimal Transport"), defended at Paris Dauphine in December 2016.
- V. Ehrlacher was in the jury for the PhD of Jean-Léopold Vié ("Second-order derivatives for shape optimization with a level-set method"), defended at CERMICS in December 2016.
- V. Ehrlacher was in the jury for the PhD of Faizan Nazar ("Electronic Structure of Defects in the Thomas-Fermi-von Weiszäcker Model"), defended at Warwick in December 2016.
- F. Legoll was in the jury for the PhD of Dena Kazerani ("Études mathématiques de fluides à frontières libres en dynamique incompressible"), defended at UPMC in November 2016.
- T. Lelièvre was a referee for the PhD of Tomasz Badowski on "Adaptive importance sampling via minimization of estimators of cross-entropy, mean square and inefficiency constants" defended at Freie Universität Berlin.
- T. Lelièvre was a referee for the PhD of Arthur Talpaert on "Direct Numerical Simulation of bubbles with Adaptive Mesh Refinement with distributed algorithms" defended at École Polytechnique.
- T. Lelièvre was in the jury for the PhD of Ahmed-Amine Homman ("Développement de schémas numériques d'intégration de méthodes multi-échelles"), defended at CERMICS in June 2016.
- T. Lelièvre was in the jury for the PhD of Gang Liu on "Rare event simulation by shaking transformation and Non-intrusive stratified resampling method for dynamic programming", defended at École Polytechnique in November 2016.

Project-team members have participated in the following habilitation juries:

- E. Cancès was a referee of Stéphane Redon's HdR, defended in Grenoble in May 2016.
- E. Cancès participated in the HdR jury of Nicolas Rougerie, defended in Grenoble in November 2016.
- T. Lelièvre was a referee of Pierre Etoré's HDR ("Quelques contributions à l'étude et à la simulation des diffusions asymétriques"), defended in Grenoble in December 2016.

9.3. Conference participation

Members of the project-team have delivered lectures in the following seminars, workshops and international conferences:

- S. Boyaval, Atelier Sillage et ondes de surface, Orsay, April 2016,
- S. Boyaval, Séminaire Modélisation mathématique et calcul scientifique, ENS Lyon and Institut Camille Jordan, May 2016,
- S. Boyaval, Workshop on Stochastic Partial Differential Equations, Pau, June 2016,
- S. Boyaval, Symposium on shallow-water flows at HYP 2016, Aachen, August 2016,
- S. Boyaval, Symposium on network models in PDEs at SIMAI 2016, Milan, September 2016,
- E. Cancès, seminar, Groupe de travail numérique, Laboratoire Jacques-Louis Lions, UPMC, January 2016,
- E. Cancès, workshop on computation of quantum systems in cold-matter physics and chemistry, Fields Institute, Toronto, February 2016,

- E. Cancès, Colloquium, University of Delaware, May 2016,
- E. Cancès, SIAM MS conference (invited lecture), May 2016,
- E. Cancès, KiNet workshop on mathematical and computational methods in quantum chemistry, Yale, New Haven, May 2016,
- E. Cancès, Solid Math workshop, Aalborg, May 2016,
- E. Cancès, KiNet workshop on quantum and kinetic transport, Jiatong University, Shanghai, June 2016,
- E. Cancès, workshop on coupled mathematical models for physical and nanoscale systems and their applications, Banff, Canada, August 2016,
- E. Cancès, Multiscale Modelling of Materials (MMM 2016) conference (invited lecture), Dijon, October 2016,
- L. Chamoin, Séminaire de l'équipe MISES, UPMC, Paris, April 2016,
- L. Chamoin, SIAM Conference on Uncertainty Quantification, Lausanne, April 2016,
- L. Chamoin, Workshop "New Challenges in Computational Mechanics", Cachan, May 2016,
- L. Chamoin, ECCOMAS conference, Hersonissos, Greece, June 2016,
- G. Di Gesù, IST Austria, Vienne, June 2016,
- G. Di Gesù, Eurandom YEP Workshop on Large Deviations for Interacting Particle Systems and Partial Differential Equations, Eindhoven, March 2016,
- V. Ehrlacher, Workshop on "Challenges in High-Dimensional Analysis and Computation", San Servolo, Italy, May 2016,
- V. Ehrlacher, Séminaire du CEREMADE, Université Paris-Dauphine, September 2016,
- V. Ehrlacher, Institute for Computational and Applied Mathematics seminar, University of Münster, Germany, October 2016,
- V. Ehrlacher, MATHCCES seminar, RWTH Aachen University, Germany, October 2016,
- V. Ehrlacher, EMI 2016 conference, Metz, France, October 2016,
- V. Ehrlacher, MMM 2016 conference, Dijon, France, October 2016,
- V. Ehrlacher, Workshop on "Recent developments in numerical methods for model reduction", Institut Henri Poincaré, France, November 2016,
- V. Ehrlacher, Séminaire "Problèmes spectraux en physique mathématique", December 2016,
- G. Ferré, seminar of the IPAM program 'Understanding Many Particle Systems with Machine Learning, Los Angeles, November 2016,
- G. Ferré, workshop 'Collective Variables in Quantum Mechanics', Los Angeles, November 2016,
- T. Hudson, 7th European Congress of Mathematics, Berlin, July 2016,
- T. Hudson, Applied Math Seminar, UNC Charlotte, May 2016,
- M. Josien, CANUM 2016, Obernai, France, May 2016,
- M. Josien, MMM 2016 conference, Dijon, France, October 2016,
- C. Le Bris, Edinburgh Mathematical Society Lecturer, Dundee, UK, March 18, 2016,
- C. Le Bris, Sissa workshop on Homogenization, SISSA, Trieste, June 6-10, 2016,
- C. Le Bris, PDE Seminar of the University of Chicago, February and November 2016,
- C. Le Bris, Seminar at the Applied Mathematics Department of the University of Washington at Seattle, October 2016,
- F. Legoll, EMI-PMC 2016 conference, Nashville, USA, May 2016,
- F. Legoll, ECCOMAS conference, Hersonissos, Greece, June 2016,
- F. Legoll, Weekly seminar of the LMS laboratory, Ecole Polytechnique, June 2016,

- F. Legoll, AIMS conference, Orlando, USA, July 2016,
- F. Legoll, WCCM conference, Seoul, South Korea, July 2016,
- F. Legoll, ECCM conference, Brussels, Belgium, Sept. 2016,
- F. Legoll, MMM 2016 conference, Dijon, France, October 2016,
- F. Legoll, Workshop on "Recent developments in numerical methods for model reduction", Paris, France, November 2016,
- F. Legoll, CASA weekly seminar, Eindhoven, the Netherlands, November 2016,
- F. Legoll, IFPEn weekly seminar, Paris, France, December 2016,
- T. Lelièvre, Plenary speaker at MCMSKI 2016, Lenzerheide, January 2016,
- T. Lelièvre, Séminaire Laboratoire Jacques-Louis Lions, February 2016,
- T. Lelièvre, Séminaire "Incertitudes" à EDF, March 2016,
- T. Lelièvre, Séminaire Institut de Biologie Physico-Chimique, March 2016,
- T. Lelièvre, Workshop "Particle methods for the management of risks", Paris, April 2016,
- T. Lelièvre, Séminaire équipe Inria ABS, April 2016,
- T. Lelièvre, SIAM Uncertainty Quantification, Lausanne, April 2016,
- T. Lelièvre, Workshop "Challenges in High-Dimensional Analysis and Computation", San Servolo, May 2016,
- T. Lelièvre, CANUM, May 2016,
- T. Lelièvre, Séminaire du laboratoire MICS, Centrale Supélec, June 2016,
- T. Lelièvre, Warwick Mathematics Colloquium, June 2016,
- T. Lelièvre, Workshop "Extreme events in the Earth and planetary sciences", Warwick, July 2016,
- T. Lelièvre, Faraday discussion "Reaction rate theory", Cambridge, September 2016,
- T. Lelièvre, Workshop MMM2016, October 2016,
- A. Levitt, Workshop on computation of quantum systems in cold-matter physics and chemistry, Toronto, Canada, February 2016,
- A. Levitt, Parallel Processing '16, Paris, April 2016,
- A. Levitt, Remise de prix Bull-Fourier 2015, Paris, April 2016,
- A. Levitt, GDR REST (rencontres de spectroscopie théorique) meeting, Roscoff, May 2016,
- A. Levitt, PASC 2016, Lausanne, May 2016,
- A. Levitt, 2016 ECMI Congress, Santiago de Compostela, May 2016,
- A. Levitt, Mathematical and numerical analysis of electronic structure models, Roscoff, July 2016,
- A. Levitt, Mathematics seminar, Aachen, March 2016,
- A. Levitt, Mathematical physics seminar, Texas A&M, November 2016,
- A. Levitt, Séminaire de mathématiques appliquées, Collège de France, December 2016,
- F. Madiot, CANUM 2016, Obernai, France, May 2016,
- P. Monmarché, Inria junior seminar, September 2016,
- B. Nectoux, groupe de travail "chimie quantique", Université Pierre et Marie Curie, January 2016,
- G. Stoltz, IPAM workshop "Collective Variables in Classical Mechanics", Los Angeles, October 2016,
- G. Stoltz, The 8th Multiscale Materials Modelling international conference, Dijon, November 2016,
- G. Stoltz, SIAM Conference on Mathematical Aspects of Materials Science 2016, Philadelphie, May 2016,
- P. Terrier, The 8th Multiscale Materials Modelling international conference, Dijon, November 2016.

Members of the project-team have delivered the following series of lectures:

- E. Cancès, Optical and electronic excitations in molecules and solids, 3h, Physics department, Harvard University, April 2016,
- E. Cancès, First-principle molecular simulation, 4h, French-Spanish Jacques-Louis Lions Summer School, Gijon, June 2016,
- E. Cancès, Density Functional Theory: models and numerical methods, 4h, KiNet summer school, Santa Barbara, June 2016,
- E. Cancès, Mathematical representations of quantum states, 2h, IPAM tutorial, Los Angeles, September 2016,
- E. Cancès, Mathematical techniques for quantum chemistry, 3h, Modern wavefunction methods in electronic structure theory, Gelsenkirchen, October 2016,
- V. Ehrlacher, Lecture on "Theoretical results on the Progressive Generalized Decomposition algorithm", 2h, cours GdR AMORE, IHP, December 2016,
- C. Le Bris, Series of 6 one-hour lectures on Nonperiodic multiscale problems, Winter school on Calculus of Variations in Physics and Materials Science, Würzburg, Germany, 14 -19 February 2016,
- C. Le Bris, Series of 4 one-hour lectures on Stochastic homogenization, INI Workshop on "From the Grain to the Continuum: Two Phase Dynamics of a Partially Molten, Polycrystalline Aggregate", Cambridge, UK, 14 -15 March 2016,
- T. Lelièvre, Lectures on "Numerical methods in molecular dynamics" (4h30), Winterschool Universität Basel, Engelberg, February 2016,
- T. Lelièvre, Lectures on "Model reduction techniques for stochastic dynamics" (4h), Ecole GDR EGRIN, May 2016,
- T. Lelièvre, Lectures on "Stochastic differential equations in large dimension and numerical methods" (4h), RICAM Winterschool, Linz, December 2016,
- G. Stoltz, "A mathematical introduction to steady-state nonequilibrium systems", Spring school on Molecular Dynamics, Bad Belzig, April 2016.

Members of the project-team have presented posters in the following seminars, workshops and international conferences:

- A. Levitt, COSMOS workshop, Paris, February 2016,
- J. Roussel, NESC, Sheffield, July 2016.

Members of the project-team have participated (without giving talks nor presenting posters) in the following seminars, workshops and international conferences:

- T, Hudson, SIAM Conference on Mathematical Aspects of Materials Science 2016, Philadelphie, May 2016,
- M. Josien, Winterschool on Stochastic Homogenization, Augsburg, Germany, February 2016,
- A. Levitt, IPAM semester “Understanding many-particle systems with machine learning”, Los Angeles, September and November 2016,
- P.-L. Rothé, Workshop on "Recent developments in numerical methods for model reduction", Paris, France, November 2016,
- J. Roussel, MCMSki Conference, Lenzerheide, January 2016,
- J. Roussel, COSMOS Workshop, Paris, February 2016,
- J. Roussel, CEMRACS, CIRM, Marseille, July 2016,
- J. Roussel, IPAM Tutorial, Los Angeles, September 2016,
- P. Terrier, Summer school PISACMS, Paris, September 2016.

9.4. Popularization

- É. Cancès has delivered a conference in Nancy in November 2016, in the framework of the series of lectures “Sciences et société”.
- A. Levitt participated in the “Young doctors” session of the Salon Culture & Jeux Mathématiques in May 2016.

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] T. Lelièvre, M. Rousset, G. Stoltz.*Free Energy Computations: A Mathematical Perspective*, Imperial College Press, 458 p., 2010.

Publications of the year

Articles in International Peer-Reviewed Journal

[7] H. Alrachid, T. Lelièvre, R. Talhouk.*Local and global solution for a nonlocal Fokker–Planck equation related to the adaptive biasing force process*, in "Journal of Differential Equations", 2016, vol. 260, n^o 9, p. 7032 - 7058 [*DOI :* 10.1016/j.jde.2016.01.020], https://hal.archives-ouvertes.fr/hal-01244976.

[8] F. Aviat, A. Levitt, B. Stamm, Y. Maday, P. Ren, J. W. Ponder, L. Lagardere, J.-P. Piquemal.*Truncated Conjugate Gradient (TCG): an optimal strategy for the analytical evaluation of the many-body polarization energy and forces in molecular simulations*, in "Journal of Chemical Theory and Computation", November 2016 [*DOI :* 10.1021/acs.jctc.6b00981], https://hal.archives-ouvertes.fr/hal-01395833.

[9] X. Blanc, C. Le Bris, F. Legoll.*Some variance reduction methods for numerical stochastic homogenization*, in "Philosophical Transactions of the Royal Society of London. A (1887–1895)", 2016, vol. 374, n^o 2066, 20150168, https://hal.archives-ouvertes.fr/hal-01196499.

[10] F. BOUCHUT, S. BOYAVAL.*Unified derivation of thin-layer reduced models for shallow free-surface gravity flows of viscous fluids*, in "European Journal of Mechanics - B/Fluids", January 2016, vol. 55, n^{o} 1, p. 116-131 [*DOI :* 10.1016/J.EUROMECHFLU.2015.09.003], https://hal-enpc.archives-ouvertes.fr/hal-00833468.

[11] C.-E. BRÉHIER, M. GAZEAU, L. GOUDENÈGE, T. LELIÈVRE, M. ROUSSET.*Unbiasedness of some generalized adaptive multilevel splitting algorithms*, in "The Annals of Applied Probability : an official journal of the institute of mathematical statistics", 2016, vol. 26, n^{o} 6, p. 3559 - 3601 [*DOI :* 10.1214/16-AAP1185], https://hal.archives-ouvertes.fr/hal-01142704.

[12] E. CANCÈS, D. GONTIER, G. STOLTZ.*A mathematical analysis of the GW0 method for computing electronic excited state energies of molecules*, in "Reviews in Mathematical Physics", June 2016 [*DOI :* 10.1142/S0129055X16500082], https://hal.archives-ouvertes.fr/hal-01160926.

[13] G. DI GESÙ, T. LELIÈVRE, D. LE PEUTREC, B. NECTOUX.*Jump Markov models and transition state theory: the quasi-stationary distribution approach*, in "Faraday Discussions", 2016, vol. 260, n^{o} 9, p. 7032 - 7058 [*DOI :* 10.1039/C6FD00120C], https://hal.archives-ouvertes.fr/hal-01314233.

[14] A. B. DUNCAN, T. LELIÈVRE, G. A. PAVLIOTIS.*Variance Reduction Using Nonreversible Langevin Samplers*, in "Journal of Statistical Physics", 2016, vol. 163, n^{o} 3, p. 457 - 491, 29 pages [*DOI :* 10.1007/S10955-016-1491-2], https://hal.archives-ouvertes.fr/hal-01164466.

[15] M. FATHI, G. STOLTZ.*Improving dynamical properties of metropolized discretizations of overdamped Langevin dynamics*, in "Numerische Mathematik", November 2016 [*DOI :* 10.1007/S00211-016-0849-3], https://hal.archives-ouvertes.fr/hal-01153573.

[16] G. FAURE, J. ROUSSEL, J. -B. MAILLET, G. STOLTZ.*Size consistency in Smoothed Dissipative Particle Dynamics*, in "Physical Review E ", October 2016 [*DOI :* 10.1103/PHYSREVE.94.043305], https://hal.archives-ouvertes.fr/hal-01345009.

[17] A.-A. HOMMAN, J.-B. MAILLET, J. ROUSSEL, G. STOLTZ.*New parallelizable schemes for integrating the Dissipative Particle Dynamics with Energy Conservation*, in "The Journal of Chemical Physics", January 2016 [*DOI :* 10.1063/1.4937797], https://hal.archives-ouvertes.fr/hal-01180558.

[18] J. INFANTE ACEVEDO, T. LELIÈVRE.*A non linear approximation method for solving high dimensional partial differential equations: Application in finance*, in "Mathematics and Computers in Simulation", 2016 [*DOI :* 10.1016/J.MATCOM.2016.07.013], https://hal-enpc.archives-ouvertes.fr/hal-00861892.

[19] T. JOURDAN, G. STOLTZ, F. LEGOLL, L. MONASSE.*An accurate scheme to solve cluster dynamics equations using a Fokker-Planck approach*, in "Computer Physics Communications", June 2016, vol. 207, 170 [*DOI :* 10.1016/J.CPC.2016.06.001], https://hal-cea.archives-ouvertes.fr/cea-01302986.

[20] C. LE BRIS, F. LEGOLL.*Examples of computational approaches for elliptic, possibly multiscale PDEs with random inputs*, in "Journal of Computational Physics", 2017, vol. 328, p. 455-473, https://hal.archives-ouvertes.fr/hal-01304040.

[21] C. LE BRIS, F. LEGOLL, F. MADIOT.*A numerical comparison of some Multiscale Finite Element approaches for advection-dominated problems in heterogeneous media*, in "ESAIM: Mathematical Modelling and Numerical Analysis", 2016, https://hal.archives-ouvertes.fr/hal-01235642.

[22] C. LE BRIS, F. LEGOLL, F. MADIOT.*Stabilisation de problèmes non coercifs via une méthode numérique utilisant la mesure invariante (Stabilization of non-coercive problems using the invariant measure)*, in "Comptes Rendus Mathématique", 2016, vol. 354, n^{o} 8, p. 799-803, https://hal.archives-ouvertes.fr/hal-01348854.

[23] C. LE BRIS, F. LEGOLL, W. MINVIELLE.*Special Quasirandom Structures: a selection approach for stochastic homogenization*, in "Monte Carlo Methods and Applications", 2016, vol. 22, n^{o} 1, p. 25-54, https://hal.archives-ouvertes.fr/hal-01194480.

[24] T. LELIÈVRE, G. STOLTZ.*Partial differential equations and stochastic methods in molecular dynamics*, in "Acta Numerica", May 2016, vol. 25, p. 681-880 [*DOI :* 10.1017/S0962492916000039], https://hal.archives-ouvertes.fr/hal-01334088.

[25] B. MARCHAND, L. CHAMOIN, C. REY.*Real-time updating of structural mechanics models using Kalman filtering, modified Constitutive Relation Error and ProperGeneralized Decomposition*, in "International Journal for Numerical Methods in Engineering", 2016 [*DOI :* 10.1002/NME.5197], https://hal.archives-ouvertes.fr/hal-01241738.

[26] S. REDON, G. STOLTZ, Z. TRSTANOVA.*Error Analysis of Modified Langevin Dynamics*, in "Journal of Statistical Physics", June 2016 [*DOI :* 10.1007/S10955-016-1544-6], https://hal.archives-ouvertes.fr/hal-01263700.

[27] K. SAB, F. LEGOLL, S. FOREST.*Stress gradient elasticity theory : existence and uniqueness of solution*, in "Journal of Elasticity", 2016, vol. 123, n^{o} 2, p. 179-201 [*DOI :* 10.1007/S10659-015-9554-1], https://hal.archives-ouvertes.fr/hal-01288563.

[28] I. G. TEJADA, L. BROCHARD, T. LELIÈVRE, G. STOLTZ, F. LEGOLL, E. CANCÈS.*Coupling a reactive potential with a harmonic approximation for atomistic simulations of material failure*, in "Computer Methods in Applied Mechanics and Engineering", 2016, vol. 305, p. 422 - 440 [*DOI :* 10.1016/J.CMA.2016.03.011], https://hal.archives-ouvertes.fr/hal-01196505.

[29] I. TEO, C. G. MAYNE, K. SCHULTEN, T. LELIÈVRE.*Adaptive Multilevel Splitting Method for Molecular Dynamics Calculation of Benzamidine-Trypsin Dissociation Time*, in "Journal of Chemical Theory and Computation", June 2016, vol. 12, n^{o} 6, p. 2983 - 2989 [*DOI :* 10.1021/ACS.JCTC.6B00277], https://hal.archives-ouvertes.fr/hal-01389074.

Conferences without Proceedings

[30] F. PLED, L. CHAMOIN, P.-E. ALLIER, P. LADEVÈZE.*On the control of PGD reduced-order approximations: error estimation and adaptivity*, in "European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2016)", Hersonissos, Crete Island, Greece, June 2016, https://hal.archives-ouvertes.fr/hal-01306402.

Other Publications

[31] X. ANTOINE, A. LEVITT, Q. TANG.*Efficient spectral computation of the stationary states of rotating Bose-Einstein condensates by the preconditioned nonlinear conjugate gradient method*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01393094.

[32] A. BAKHTA, V. EHRLACHER.*Cross-diffusion systems with non-zero flux and moving boundary conditions*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01397682.

[33] N. BERGLUND, G. DI GESÙ, H. WEBER.*An Eyring–Kramers law for the stochastic Allen–Cahn equation in dimension two*, April 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01304559.

[34] X. BLANC, M. JOSIEN.*From the Newton equation to the wave equation : the case of shock waves: De l'équation de Newton à l'équation des ondes dans le cas d'ondes de chocs*, May 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01314690.

[35] S. BOYAVAL.*Johnson-Segalman – Saint-Venant equations for viscoelastic shallow flows in the elastic limit*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01402628.

[36] E. CANCÈS, P. CAZEAUX, M. LUSKIN.*Generalized Kubo Formulas for the Transport Properties of Incommensurate 2D Atomic Heterostructures*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01403588.

[37] E. CANCÈS, G. DUSSON, Y. MADAY, B. STAMM, M. VOHRALÍK.*Guaranteed and robust a posteriori bounds for Laplace eigenvalues and eigenvectors: conforming approximations*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01194364.

[38] É. CANCÈS, A. LEVITT, G. PANATI, G. STOLTZ.*Robust determination of maximally-localized Wannier functions*, May 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01323700.

[39] G. DI GESÙ, M. MARIANI.*Full metastable asymptotic of the fisher information*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01400800.

[40] V. EHRLACHER, D. LOMBARDI.*A dynamical adaptive tensor method for the Vlasov-Poisson system*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01335507.

[41] G. FORT, B. JOURDAIN, T. LELIEVRE, G. STOLTZ.*Convergence and efficiency of adaptive importance sampling techniques with partial biasing*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01389996.

[42] C. L. HALL, T. HUDSON, P. V. MEURS.*Asymptotic analysis of boundary layers in a repulsive particle system*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01398449.

[43] T. HUDSON.*Upscaling a model for the thermally-driven motion of screw dislocations*, July 2016, Major revision, including overhaul of notation, significant additions to Large Deviations results in Section 6, and resolution of conjecture in original version. 45 pages, 2 figures, 1 table, https://hal.archives-ouvertes.fr/hal-01236494.

[44] C. LE BRIS, F. LEGOLL, S. LEMAIRE.*On the best constant matrix approximating an oscillatory matrix-valued coefficient in divergence-form operators*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01420187.

[45] C. LE BRIS, F. LEGOLL, F. MADIOT.*Stable approximation of the advection-diffusion equation using the invariant measure*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01367417.

[46] A. LE COËNT, L. FRIBOURG, N. MARKEY, F. DE VUYST, L. CHAMOIN.*Distributed Synthesis of State-Dependent Switching Control*, March 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01295738.

[47] D. LE PEUTREC.*On Witten Laplacians and Brascamp-Lieb's inequality on manifolds with boundary*, July 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01349786.

[48] F. LEGOLL, T. LELIEVRE, S. OLLA.*Pathwise estimates for an effective dynamics*, May 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01314221.

[49] A. LEVITT, C. ORTNER.*Convergence and Cycling in Walker-type Saddle Search Algorithms*, July 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01350514.

[50] G. STOLTZ.*Stable schemes for dissipative particle dynamics with conserved energy*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01416108.

[51] G. STOLTZ, Z. TRSTANOVA.*Stable and accurate schemes for Langevin dynamics with general kinetic energies*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01364821.

[52] P. TERRIER, M. ATHÈNES, T. JOURDAN, G. ADJANOR, G. STOLTZ.*Cluster dynamics modelling of materials: a new hybrid deterministic/stochastic coupling approach*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01378916.

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:

6. - Modeling, simulation and control
6.1. - Mathematical Modeling
6.1.2. - Stochastic Modeling (SPDE, SDE)
6.2.1. - Numerical analysis of PDE and ODE
6.2.2. - Numerical probability
6.2.3. - Probabilistic methods
6.4.2. - Stochastic control

Other Research Topics and Application Domains:

3.1. - Sustainable development
9.5.3. - Economy, Finance
9.9. - Risk management

1. Members

Research Scientists

Agnès Bialobroda Sulem [Team leader, Inria, Senior Researcher, HDR]
Aurélien Alfonsi [ENPC, Senior Researcher, HDR]
Benjamin Jourdain [ENPC, Senior Researcher, HDR]
Bernard Lapeyre [ENPC, 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]

PhD Student

Rui Chen [Université Paris-Dauphine]

Post-Doctoral Fellow

Arnaud Lionnet [Inria, until Aug 2016]

Administrative Assistant

Martine Verneuille [Inria]

Others

Céline Labart [Université de Savoie, Associate professor, until Aug 2016, Partner Researcher]
Jérôme Lelong [ENSIMAG, Associate professor, Partner Researcher]
Jean-Philippe Chancelier [ENPC, Research Scientist, Partner Researcher]
Babacar Diallo [Inria, Trainee, from Mar 2016 until Aug 2016]
Oleg Kudryavtsev [Roskov University, Russia, Invited professor, until Aug 2016]
Nicolas Le Mouel [Inria, Trainee, from Jul 2016 until Oct 2016]
Mouad Ramil [Inria, Trainee, from Mar 2016 until Aug 2016]
Antonino Zanette [University of Udine, Italy, Professor, Partner Researcher]

2. Overall Objectives

2.1. Overall Objectives

MathRisk is a joint Inria project-team with ENPC (CERMICS Laboratory) and the University Paris Est Marne-la-Vallée (UPEMLV, LAMA Laboratory), located in Paris and Marne-la-Vallée.
http://www.inria.fr/en/teams/mathrisk. Mathrisk is based on the former Mathfi project team. Mathfi was founded in 2000, and was devoted to financial mathematics. The project was focused on advanced stochastic analysis and numerical techniques motivated by the development of increasingly complex financial products. Main applications concerned evaluation and hedging of derivative products, dynamic portfolio optimization in incomplete markets, and calibration of financial models.

2.1.1. Crisis, deregulation, and impact on the research in finance

The starting point of the development of modern finance theory is traditionally associated to the publication of the famous paper of Black and Scholes in 1973 [65]. Since then, in spite of sporadic crises, generally well overcome, financial markets have grown in a exponential manner. More and more complex exotic derivative products have appeared, on equities first, then on interest rates, and more recently on credit markets. The period between the end of the eighties and the crisis of 2008 can be qualified as the "golden age of financial mathematics": finance became a quantitative industry, and financial mathematics programs flourished in top universities, involving seminal interplays between the worlds of finance and applied mathematics. During its 12 years existence, the Mathfi project team has extensively contributed to the development of modeling and computational methods for the pricing and hedging of increasingly complex financial products.

Since the crisis of 2008, there has been a critical reorientation of research priorities in quantitative finance with emphasis on risk. In 2008, the "subprime" crisis has questioned the very existence of some derivative products such as CDS (credit default swaps) or CDOs (collateralized debt obligations), which were accused to be responsible for the crisis. The nature of this crisis is profoundly different from the previous ones. It has negatively impacted the activity on the exotic products in general, - even on equity derivative markets-, and the interest in the modeling issues for these products. The perfect replication paradigm, at the origin of the success of the Black and Scholes model became unsound, in particular through the effects of the lack of liquidity. The interest of quantitative finance analysts and mathematicians shifted then to more realistic models taking into account the multidimensional feature and the incompleteness of the markets, but as such getting away from the "lost paradi(gm)" of perfect replication. These models are much more demanding numerically, and require the development of hedging risk measures, and decision procedures taking into account the illiquidity and various defaults.

Moreover, this crisis, and in particular the Lehman Brothers bankruptcy and its consequences, has underlined a systemic risk due to the strong interdependencies of financial institutions. The failure of one of them can cause a cascade of failures, thus affecting the global stability of the system. Better understanding of these interlinkage phenomena becomes crucial.

At the same time, independently from the subprime crisis, another phenomenon has appeared: deregulation in the organization of stock markets themselves. This has been encouraged by the Markets in Financial Instruments Directive (MIFID) which is effective since November, 1st 2007. This, together with the progress of the networks, and the fact that all the computers have now a high computation power, have induced arbitrage opportunities on the markets, by very short term trading, often performed by automatic trading. Using these high frequency trading possibilities, some speculating operators benefit from the large volatility of the markets. For example, the flash crash of May, 6 2010 has exhibited some perverse effects of these automatic speculating trading strategies. These phenomena are not well understood and the theme of high frequency trading needs to be explored.

To summarize, financial mathematics is facing the following new evolutions:

- the complete market modeling has become unsatisfactory to provide a realistic picture of the market and is replaced by incomplete and multidimensional models which lead to new modeling and numerical challenges.
- quantitative measures of risk coming from the markets, the hedging procedures, and the lack of liquidity are crucial for banks,
- uncontrolled systemic risks may cause planetary economic disasters, and require better understanding,
- deregulation of stock markets and its consequences lead to study high frequency trading.

The project team MathRisk is designed to address these new issues, in particular dependence modeling, systemic risk, market microstructure modeling and risk measures. The research in modeling and numerical analysis remain active in this new context, motivated by new issues.

The MathRisk project team develops the software Premia dedicated to pricing and hedging options and calibration of financial models, in collaboration with a consortium of financial institutions. https://www.rocq.inria.fr/mathfi/Premia/index.html.

The MathRisk project is part of the Université Paris-Est **“Labex” BÉZOUT**.

3. Research Program

3.1. Dependence modeling

Participants: Aurélien Alfonsi, Benjamin Jourdain, Damien Lamberton, Bernard Lapeyre.

The volatility is a key concept in modern mathematical finance, and an indicator of the market stability. Risk management and associated instruments depend strongly on the volatility, and volatility modeling has thus become a crucial issue in the finance industry. Of particular importance is the assets *dependence* modeling. The calibration of models for a single asset can now be well managed by banks but modeling of dependence is the bottleneck to efficiently aggregate such models. A typical issue is how to go from the individual evolution of each stock belonging to an index to the joint modeling of these stocks. In this perspective, we want to model stochastic volatility in a *multidimensional* framework. To handle these questions mathematically, we have to deal with stochastic differential equations that are defined on matrices in order to model either the instantaneous covariance or the instantaneous correlation between the assets. From a numerical point of view, such models are very demanding since the main indexes include generally more than thirty assets. It is therefore necessary to develop efficient numerical methods for pricing options and calibrating such models to market data. As a first application, modeling the dependence between assets allows us to better handle derivatives products on a basket. It would give also a way to price and hedge consistenly single-asset and basket products. Besides, it can be a way to capture how the market estimates the dependence between assets. This could give some insights on how the market anticipates the systemic risk.

3.2. Liquidity risk

Participants: Aurélien Alfonsi, Agnès Bialobroda Sulem, Antonino Zanette.

The financial crisis has caused an increased interest in mathematical finance studies which take into account the market incompleteness issue and the liquidity risk. Loosely speaking, liquidity risk is the risk that comes from the difficulty of selling (or buying) an asset. At the extreme, this may be the impossibility to sell an asset, which occurred for “junk assets” during the subprime crisis. Hopefully, it is in general possible to sell assets, but this may have some cost. Let us be more precise. 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. Instead, 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), this 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.

3.2.1. Long term liquidity risk.

On a long-term perspective, illiquidity can be approached via various ways: transactions costs [57], [58], [64], [71], [74], [89], [85], delay in the execution of the trading orders [90], [88], [67], trading constraints or restriction on the observation times (see e.g. [73] and references herein). As far as derivative products are concerned, one has to understand how delta-hedging strategies have to be modified. This has been considered for example by Cetin, Jarrow and Protter [87]. We plan to contribute on these various aspects of liquidity risk modeling and associated stochastic optimization problems. Let us mention here that the price impact generated by the trades of the investor is often neglected with a long-term perspective. This seems acceptable since the investor has time enough to trade slowly in order to eliminate its market impact. Instead, when the investor wants to make significant trades on a very short time horizon, it is crucial to take into account and to model how prices are modified by these trades. This question is addressed in the next paragraph on market microstructure.

3.2.2. Market microstructure.

The European directive MIFID has increased the competition between markets (NYSE-Euronext, Nasdaq, LSE and new competitors). As a consequence, the cost of posting buy or sell orders on markets has decreased, which has stimulated the growth of market makers. Market makers are posting simultaneously bid and ask orders on a same stock, and their profit comes from the bid-ask spread. Basically, their strategy is a "round-trip" (i.e. their position is unchanged between the beginning and the end of the day) that has generated a positive cash flow.

These new rules have also greatly stimulated research on market microstructure modeling. From a practitioner point of view, the main issue is to solve the so-called "optimal execution problem": given a deadline T, what is the optimal strategy to buy (or sell) a given amount of shares that achieves the minimal expected cost? For large amounts, it may be optimal to split the order into smaller ones. This is of course a crucial issue for brokers, but also market makers that are looking for the optimal round-trip.

Solving the optimal execution problem is not only an interesting mathematical challenge. It is also a mean to better understand market viability, high frequency arbitrage strategies and consequences of the competition between markets. For example when modeling the market microstructure, one would like to find conditions that allow or exclude round trips. Beyond this, even if round trips are excluded, it can happen that an optimal selling strategy is made with large intermediate buy trades, which is unlikely and may lead to market instability.

We are interested in finding synthetic market models in which we can describe and solve the optimal execution problem. A. Alfonsi and A. Schied (Mannheim University) [59] have already proposed a simple Limit Order Book model (LOB) in which an explicit solution can be found for the optimal execution problem. We are now interested in considering more sophisticated models that take into account realistic features of the market such as short memory or stochastic LOB. This is mid term objective. At a long term perspective one would like to bridge these models to the different agent behaviors, in order to understand the effect of the different quotation mechanisms (transaction costs for limit orders, tick size, etc.) on the market stability.

3.3. Contagion modeling and systemic risk

Participants: Benjamin Jourdain, Agnès Bialobroda Sulem.

After the recent financial crisis, systemic risk has emerged as one of the major research topics in mathematical finance. 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. Ideally, one would like to be able to find capital requirements (such as the one proposed by the Basel committee) that ensure that the probability of multiple defaults is below some level.

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. [60] and A. Minca [79]) or mean field interaction models (Garnier-Papanicolaou-Yang [72]). The recent approach in [60] seems very promising. It describes the financial network approach as a weighted directed graph, in which nodes represent financial institutions and edges the exposures between them. Distress propagation in a financial system may be modeled as an epidemics on this graph. In the case of incomplete information on the structure of the interbank network, cascade dynamics may be reduced to the evolution of a multi-dimensional Markov chain that corresponds to a sequential discovery of exposures and determines at any time the size of contagion. Little has been done so far on the *control* of such systems in order to reduce the systemic risk and we aim to contribute to this domain.

3.4. Stochastic analysis and numerical probability

3.4.1. Stochastic control

Participants: Vlad Bally, Jean-Philippe Chancelier, Marie-Claire Quenez, Agnès Bialobroda Sulem.

The financial crisis has caused an increased interest in mathematical finance studies which take into account the market incompleteness issue and the default risk modeling, the interplay between information and performance, the model uncertainty and the associated robustness questions, and various nonlinearities. We address these questions by further developing the theory of stochastic control in a broad sense, including stochastic optimization, nonlinear expectations, Malliavin calculus, stochastic differential games and various aspects of optimal stopping.

3.4.2. Optimal stopping

Participants: Aurélien Alfonsi, Benjamin Jourdain, Damien Lamberton, Agnès Bialobroda Sulem, Marie-Claire Quenez.

The theory of American option pricing has been an incite for a number of research articles about optimal stopping. Our recent contributions in this field concern optimal stopping in models with jumps, irregular obstacles, free boundary analysis, reflected BSDEs.

3.4.3. Simulation of stochastic differential equations

Participants: Benjamin Jourdain, Aurélien Alfonsi, Vlad Bally, Damien Lamberton, Bernard Lapeyre, Jérôme Lelong, Céline Labart.

Effective numerical methods are crucial in the pricing and hedging of derivative securities. The need for more complex models leads to stochastic differential equations which cannot be solved explicitly, and the development of discretization techniques is essential in the treatment of these models. The project MathRisk addresses fundamental mathematical questions as well as numerical issues in the following (non exhaustive) list of topics: Multidimensional stochastic differential equations, High order discretization schemes, Singular stochastic differential equations, Backward stochastic differential equations.

3.4.4. Monte-Carlo simulations

Participants: Benjamin Jourdain, Aurélien Alfonsi, Damien Lamberton, Vlad Bally, Bernard Lapeyre, Ahmed Kebaier, Céline Labart, Jérôme Lelong, Antonino Zanette.

Monte-Carlo methods is a very useful tool to evaluate prices especially for complex models or options. We carry on research on *adaptive variance reduction methods* and to use *Monte-Carlo methods for calibration* of advanced models.

This activity in the MathRisk team is strongly related to the development of the Premia software.

3.4.5. Malliavin calculus and applications in finance

Participants: Vlad Bally, Arturo Kohatsu-Higa, Agnès Bialobroda Sulem, Antonino Zanette.

The original Stochastic Calculus of Variations, now called the Malliavin calculus, was developed by Paul Malliavin in 1976 [77]. It was originally designed to study the smoothness of the densities of solutions of stochastic differential equations. One of its striking features is that it provides a probabilistic proof of the celebrated Hörmander theorem, which gives a condition for a partial differential operator to be hypoelliptic. This illustrates the power of this calculus. In the following years a lot of probabilists worked on this topic and the theory was developed further either as analysis on the Wiener space or in a white noise setting. Many applications in the field of stochastic calculus followed. Several monographs and lecture notes (for example D. Nualart [80], D. Bell [63] D. Ocone [82], B. Øksendal [91]) give expositions of the subject. See also V. Bally [61] for an introduction to Malliavin calculus.

From the beginning of the nineties, applications of the Malliavin calculus in finance have appeared : In 1991 Karatzas and Ocone showed how the Malliavin calculus, as further developed by Ocone and others, could be used in the computation of hedging portfolios in complete markets [81].

Since then, the Malliavin calculus has raised increasing interest and subsequently many other applications to finance have been found [78], such as minimal variance hedging and Monte Carlo methods for option pricing. More recently, the Malliavin calculus has also become a useful tool for studying insider trading models and some extended market models driven by Lévy processes or fractional Brownian motion.

We give below an idea why Malliavin calculus may be a useful instrument for probabilistic numerical methods.

We recall that the theory is based on an integration by parts formula of the form $E(f'(X)) = E(f(X)Q)$. Here X is a random variable which is supposed to be "smooth" in a certain sense and non-degenerated. A basic example is to take $X = \sigma\Delta$ where Δ is a standard normally distributed random variable and σ is a strictly positive number. Note that an integration by parts formula may be obtained just by using the usual integration by parts in the presence of the Gaussian density. But we may go further and take X to be an aggregate of Gaussian random variables (think for example of the Euler scheme for a diffusion process) or the limit of such simple functionals.

An important feature is that one has a relatively explicit expression for the weight Q which appears in the integration by parts formula, and this expression is given in terms of some Malliavin-derivative operators.

Let us now look at one of the main consequences of the integration by parts formula. If one considers the *Dirac* function $\delta_x(y)$, then $\delta_x(y) = H'(y-x)$ where H is the *Heaviside* function and the above integration by parts formula reads $E(\delta_x(X)) = E(H(X-x)Q)$, where $E(\delta_x(X))$ can be interpreted as the density of the random variable X. We thus obtain an integral representation of the density of the law of X. This is the starting point of the approach to the density of the law of a diffusion process: the above integral representation allows us to prove that under appropriate hypothesis the density of X is smooth and also to derive upper and lower bounds for it. Concerning simulation by Monte Carlo methods, suppose that you want to compute $E(\delta_x(y)) \sim \frac{1}{M}\sum_{i=1}^{M}\delta_x(X^i)$ where $X^1, ..., X^M$ is a sample of X. As X has a law which is absolutely continuous with respect to the Lebesgue measure, this will fail because no X^i hits exactly x. But if you are able to simulate the weight Q as well (and this is the case in many applications because of the explicit form mentioned above) then you may try to compute $E(\delta_x(X)) = E(H(X-x)Q) \sim \frac{1}{M}\sum_{i=1}^{M}E(H(X^i-x)Q^i)$. This basic remark formula leads to efficient methods to compute by a Monte Carlo method some irregular quantities as derivatives of option prices with respect to some parameters (the *Greeks*) or conditional expectations, which appear in the pricing of American options by the dynamic programming). See the papers by Fournié et al [70] and [69] and the papers by Bally et al., Benhamou, Bermin et al., Bernis et al., Cvitanic et al., Talay and Zheng and Temam in [76].

L. Caramellino, A. Zanette and V. Bally have been concerned with the computation of conditional expectations using Integration by Parts formulas and applications to the numerical computation of the price and the Greeks (sensitivities) of American or Bermudean options. The aim of this research was to extend a paper of Reigner and Lions who treated the problem in dimension one to higher dimension - which represent the real challenge in this field. Significant results have been obtained up to dimension 5 [62] and the corresponding algorithms have been implemented in the Premia software.

Moreover, there is an increasing interest in considering jump components in the financial models, especially motivated by calibration reasons. Algorithms based on the integration by parts formulas have been developed in order to compute Greeks for options with discontinuous payoff (e.g. digital options). Several papers and two theses (M. Messaoud and M. Bavouzet defended in 2006) have been published on this topic and the corresponding algorithms have been implemented in Premia. Malliavin Calculus for jump type diffusions - and more general for random variables with locally smooth law - represents a large field of research, also for applications to credit risk problems.

The Malliavin calculus is also used in models of insider trading. The "enlargement of filtration" technique plays an important role in the modeling of such problems and the Malliavin calculus can be used to obtain general results about when and how such filtration enlargement is possible. See the paper by P. Imkeller in [76]). Moreover, in the case when the additional information of the insider is generated by adding the information about the value of one extra random variable, the Malliavin calculus can be used to find explicitly the optimal portfolio of an insider for a utility optimization problem with logarithmic utility. See the paper by J.A. León, R. Navarro and D. Nualart in [76]).

A. Kohatsu Higa and A. Sulem have studied a controlled stochastic system whose state is described by a stochastic differential equation with anticipating coefficients. These SDEs can be interpreted in the sense of *forward integrals*, which are the natural generalization of the semimartingale integrals, as introduced by Russo and Valois [84]. This methodology has been applied for utility maximization with insiders.

4. Application Domains

4.1. Financial Mathematics, Insurance

The applications domains 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

- Creation of a joint seminar on Numerical probability and Mathematical Finance with the LPMA laboratory, University Paris-Diderot.
- Organization by B. Jourdain with B. Bouchard (Université Paris-Dauphine) and E. Gobet (Ecole Polytechnique) of the 2015-2016 thematic semester on Monte Carlo methods (financed by the Institute Louis Bachelier) at Institut Henri Poincaré, Paris https://www.ceremade.dauphine.fr/montecarlo/MonteCarlo.html, and the international closing conference in July 2016. https://montecarlo16.sciencesconf.org

6. New Software and Platforms

6.1. PREMIA

KEYWORDS: Financial products - Computational finance - Option pricing

6.2. Scientific Description

Premia is a software designed for option pricing, hedging and financial model calibration.
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 & 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.

6.3. Functional Description

- Participants: Mathrisk project team and contributors
- Partners: Ecole des Ponts ParisTech - Inria - Université Paris-Est - Consortium Premia
- Contact: Agnès Sulem
- URL: http://www.premia.fr
- AMS: 91B28;65Cxx;65Fxx;65Lxx;65Pxx
- License: Licence Propriétaire (genuine license for the Consortium Premia)
- OS/Middelware: Linux, Mac OS X, Windows
- APP: The development of Premia started in 1999 and 17 are released up to now and registered at the APP agency. Premia 16 has been registered on 0303/2015 under the number IDDN.FR.001.190010.013.S.C.2001.000.31000
- Programming language: C/C++
- Documentation: scientific documentation of all the algorithm implemented. PNL has a 100 pages user documentation
- Size of the software: For the Src part of Premia : 337046 lines , that is 14 Mbyte of code, and 117 Mbyte of PDF files of documentation; For PNL: 747952 lines , that is 25 MO.
- interfaces : Nsp for Windows/Linux/Mac, Excel, binding Python, and a Web interface.
- Publications: [15], [68], [75], [83], [86], [56], [66].

6.4. Content

Premia contains various numerical algorithms (Finite-differences, trees and Monte-Carlo) for pricing vanilla and exotic options on equities, interest rate, credit and energy derivatives.

- **Equity derivatives:**

  The following models are considered:

  Black-Scholes model (up to dimension 10), stochastic volatility models (Hull-White, Heston, Fouque-Papanicolaou-Sircar), models with jumps (Merton, Kou, Tempered stable processes, Variance gamma, Normal inverse Gaussian), Bates model.

  For high dimensional American options, Premia provides the most recent Monte-Carlo algorithms: Longstaff-Schwartz, Barraquand-Martineau, Tsitsklis-Van Roy, Broadie-Glassermann, quantization methods and Malliavin calculus based methods.

  Dynamic Hedging for Black-Scholes and jump models is available.

Calibration algorithms for some models with jumps, local volatility and stochastic volatility are implemented.

- **Interest rate derivatives**

  The following models are considered:

  HJM and Libor Market Models (LMM): affine models, Hull-White, CIR++, Black-Karasinsky, Squared-Gaussian, Li-Ritchken-Sankarasubramanian, Bhar-Chiarella, Jump diffusion LMM, Markov functional LMM, LMM with stochastic volatility.

  Premia provides a calibration toolbox for Libor Market model using a database of swaptions and caps implied volatilities.

- **Credit derivatives: Credit default swaps (CDS), Collateralized debt obligations (CDO)**

  Reduced form models and copula models are considered.

  Premia provides a toolbox for pricing CDOs using the most recent algorithms (Hull-White, Laurent-Gregory, El Karoui-Jiao, Yang-Zhang, Schönbucher)

- **Hybrid products**

  A PDE solver for pricing derivatives on hybrid products like options on inflation and interest or change rates is implemented.

- **Energy derivatives: swing options**

  Mean reverting and jump models are considered.

  Premia provides a toolbox for pricing swing options using finite differences, Monte-Carlo Malliavin-based approach and quantization algorithms.

6.5. PNL numerical library

To facilitate contributions, a standardized numerical library (PNL) has been developed by J. Lelong under the LGPL since 2009, which offers a wide variety of high level numerical methods for dealing with linear algebra, numerical integration, optimization, random number generators, Fourier and Laplace transforms, and much more. Everyone who wishes to contribute is encouraged to base its code on PNL and providing such a unified numerical library has considerably eased the development of new algorithms which have become over the releases more and more sophisticated.

This year, Jérome Lelong has performed the following tasks on the development of PNL:
Releases 1.7.3 and 1.7.4. of the *PNL* library (http://pnl.gforge.inria.fr/).

1. Simplify the use of PNL under Visual Studio. It can either be compiled using CMake or added as an external library to an existing project.
2. Improve the construction of large `PnlBasis` objects and make it possible to deal with non tensor functions.
3. Add complex error functions.

6.6. Consortium Premia

The software Premia is supported by a Consortium of financial institutions created in 1999. The members of the Consortium give an annual financial contribution and receive every year a new version enriched with new algorithms. They participate to the annual meeting where future new developments are discussed.

6.7. Diffusion

All releases of the software Premia (18 in 2016) are registered at the French agency APP. The most recent is provided to the Consortium with an appropriate licence. An opensource version is also available for academic purposes. The software is thus used in many universities, with in France and abroad.

6.8. Algorithms implemented in Premia in 2016

Premia 18 has been delivered to the consortium members in March 2016.

It contains the following new algorithms:

6.8.1. Insurance, Risk Management, Optimal Trade Execution

- A Forward Solution for Computing Derivatives Exposure. M. Ben Taarit, B. Lapeyre.
- Monte Carlo Calculation of Exposure Profiles and Greeks for Bermudan and Barrier Options under the Heston Hull-White Model. Q. Feng, C.W. Oosterlee.
- Dynamic optimal execution in a mixed-market-impact Hawkes price model. A. Alfonsi, P. Blanc. *Finance & Stochastics*
- A Hamilton Jacobi Bellman approach to optimal trade execution. P. Forsyth, *Applied Numerical Mathematics 61, 2011.*

6.8.2. Equity Derivatives

- Value function approximation or stopping time approximation: a comparison of two recent numerical methods for American option pricing using simulation and regression. L. Stentoft, *Journal of Computational Finance, 18(1), 2014.*
- Pricing American-Style Options by Monte Carlo Simulation: Alternatives to Ordinary Least Squares. S. Tompaidis, C. Yang, *Journal of Computational Finance, 18(1), 2014,*
- Solving Optimal Stopping Problems using Martingale Bases. J.Lelong
- The Stochastic Grid Bundling Method: Efficient Pricing of Bermudan Options and their Greeks. S. Jain, C.W. Oosterlee.
- Two-dimensional Fourier cosine series expansion method for pricing financial options. C.W.Oosterlee M.J. Ruijter, *SIAM J. Sci. Comput., 34(5), 2012.*
- Estimation of the parameters of the Wishart process. A.Alfonsi, A.Kebaier, C.Rey, *Preprint*
- The 4/2 Stochastic Volatility Model. M. Grasselli, *Preprint.*
- Ninomiya Victoir Scheme and Multi Level Scheme. A. Al Gerbi, E. Clement, B. Jourdain.
- Importance Sampling for Multilevel Monte Carlo. A.Kebaier J.Lelong
- The evaluation of barrier option prices under stochastic volatility. C.Chiarella, B.Kang, G.H.Meyer, *Computers and Mathematics with Applications 64, 2012.*
- Volatility swaps and volatility options on discretely sampled realized variance. G.Lian, C.Chiarella, P.S.Kalev, *Journal of Economic Dynamics Control 47 2014*
- Efficient variations of the Fourier transform in applications to option pricing. S. Boyarchenko and S.Levendorski, *Journal of Computational Finance, 18(2), 2014.*
- Model-free implied volatility: from surface to index. M. Fukasawa et al., *Int. J. Theor. Appl. Finan. 14, 433, 2011*
- Stratified approximations for the pricing of options on average. N.Privault J.Yu *Journal of Computational Finance.*

Moreover, J. Lelong has ensured everyday maintenance to fix various bugs, especially related to Visual C++; and has get rid of the old bunch of scripts to generate the HTML documentation by implementing the required mechanism directly in TeX. This makes the system much more robust; He has also worked on the continuous integration process with Sébastien Hinderer. Moreover, part of Premia documentation is now generated directly from the source code. The 3000 lines of undocumented C code used so far had become unmaintainable. Now, it is replaced by a much more flexible and efficient Python script.

7. New Results

7.1. Systemic risk

Participants: Agnès Bialobroda Sulem, Andreea Minca [Cornell University)], Rui Chen.

Our objective is to study the magnitude of default contagion in a large financial system, in which banks receive benefits from their connections, and to investigate how the institutions choose their connectivities by weighing the default risk and the benefits induced by connectivity. We study two versions of the model. In the first version (static) the benefits are received at the end of the contagion. In this case, each bank either receives fixed benefits per link if it survives, otherwise its payoff is zero. In the second version, which is a dynamic model, banks receive cash-flows from their connections, spread over time. Effectively, these cash flows increase the threshold of the bank over the time of contagion. We call this model contagion with intrinsic recovery features. In the first model, there is no calendar time. In the second model, the cash flows arrive at a certain rate in calendar time, while the losses come with each revealed link. We thus need to relate the intensity of revealing a link with calendar time. Both models have new features compared to past literature. The most important feature is that banks choose their connectivities optimally. The second model is dynamic and introduces growth over time. Computing the magnitude of contagion in this case is challenging, and we provide an iterative solution for this.

7.2. Backward stochastic (partial) differential equations with jumps, optimal stopping and stochastic control with nonlinear expectation

Participants: Agnès Bialobroda Sulem, Roxana Dumitrescu, Marie-Claire Quenez [(Univ Paris 7)], Bernt Øksendal, Arnaud Lionnet.

7.2.1. Nonlinear pricing in imperfect financial markets with default.

We pursue the development of the theory of stochastic control and optimal stopping with nonlinear expectation induced by a nonlinear BSDE with (default) jump, and the application to nonlinear pricing in financial markets with default. To that purpose we have studied nonlinear BSDE with default and proved several properties for these equations. We have also addressed the case with ambiguity on the model, in particular ambiguity on the default probability. In this context, we study robust superhedging strategies for the seller of a game optimal stopping problem by proving some duality results, and characterize the robust seller's price of a game option as the value function of a *mixed generalized* Dynkin game.

7.2.2. Stochastic control of mean-field SPDEs with jumps

We study stochastic maximum principles, both necessary and sufficient, for SPDE with jumps with a general mean-field operator.

7.2.3. Numerical methods for Forward-Backward SDEs

The majority of the results on the numerical methods for FBSDEs relies on the global Lipschitz assumption, which is not satisfied for a number of important cases such as the Fisher-KPP or the FitzHugh-Nagumo equations. In a previous work, A. Lionnet with Gonzalo Dos Reis and Lukasz Szpruch showed that for BSDEs with monotone drivers having polynomial growth in the primary variable y, only the (sufficiently) implicit schemes converge. But these require an additional computational effort compared to explicit schemes. They have thus developed a general framework that allows the analysis, in a systematic fashion, of the integrability properties, convergence and qualitative properties (e.g. comparison theorem) for whole families of modified explicit schemes. These modified schemes are characterized by the replacement of the driver by a driver that depends on the time-grid, and converge to the original driver as the size of the time-steps goes to 0. The framework yields the convergence of some modified explicit scheme with the same rate as implicit schemes and with the computational cost of the standard explicit scheme [55].

7.3. Optimal transport

Participants: Aurélien Alfonsi, Benjamin Jourdain.

With J. Corbetta (postdoc financed by the chair financial risks), A. Alfonsi and B. Jourdain are interested in the time derivative of the Wasserstein distance between the marginals of two Markov processes. 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 proved that the evolution of the Wasserstein distance is actually given by this candidate. In dimension one, they show that this remains true for Piecewise Deterministic Markov Processes [45].

7.4. Option Pricing and Calibration

7.4.1. Calibration of regime-switching local volatility models

Participant: Benjamin Jourdain.

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 (2011), 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 where 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 prove existence to the associated Fokker-Planck equation. Thanks to Figalli (2008), they then deduce existence of a new class of fake Brownian motions. They extend these results to the special case of the LSV model called Regime Switching Local Volatility, where the stochastic volatility factor is a jump process taking finitely many values and with jump intensities depending on the spot level. Under the same condition on the range of its square, they prove existence to the associated Fokker-Planck PDE. They finally deduce existence of the calibrated model by extending the results in Figalli (2008).

7.4.2. American options

Participant: Damien Lamberton.

With Mihail Zervos, D. Lamberton has worked on American options involving the maximum of the underlying asset. With Giulia Terenzi, he has been working on American options in Heston's model. They obtained results about existence and uniqueness for the associated variational inequality, in suitable weighted Sobolev spaces (see Feehan and co-authors for recent results on elliptic problems).

7.5. Dependence modeling

7.5.1. Estimation of the parameters of a Wishart process

Participants: Aurélien Alfonsi, Ahmed Kebaier, Clément Rey.

We have studied the Maximum Likelihood Estimator for the Wishart processes and in particular its convergence in the ergodic and in some non ergodic cases. In the non ergodic cases, our analysis rely on refined results on the Laplace transform for Wishart processes. Our work also extends the recent paper by Ben Alaya and Kebaier on the maximum likelihood estimation for the CIR process.

7.6. Numerical Probability

7.6.1. Parametrix method for reflected SDEs

With A. Kohatsu-Higa and M. Hayashi, Aurelien Alfonsi is investigating how to apply the parametrix method recently proposed by V. Bally and A. Kohatsu-Higa for reflected SDEs. This method allows them to obtain an unbiased estimator for expectations of general functions of the process.

7.6.2. Regularity of probability laws using an interpolation method

Participants: Vlad Bally, Lucia Caramellino.

This work was motivated by previous papers of Nicolas Fournier, J. Printemps, E. Clément, A. Debussche and V. Bally on the regularity of the law of the solutions of some equations with coefficients with little regularity - for example diffusion processes with Hölder coefficients (but also many other examples including jump type equations, Bolzmann equation or Stochastic PDE's). Since we do not have sufficient regularity the usual approach by Malliavin calculus fails in this framework. Then one may use an alternative idea which roughly speaking is the following: We approximate the law of the random variable X (the solution of the equation at hand) by a sequence $X(n)$ of random variables which are smooth and consequently we are able to establish integration by parts formulas for $X(n)$ and we are able to obtain the absolutely continuity of the law of $X(n)$ and to establish estimates for the density of the law of $X(n)$ and for its derivatives. Notice that the derivatives of the densities of $X(n)$ generally blow up - so we can not derive directly results concerning the density of the law of X. But, if the speed of convergence of $X(n)$ to X is stronger than the blow up, then we may obtain results concerning the density of the law of X. It turns out that this approach fits in the framework of interpolation spaces and that the criterion of regularity for the law of X amounts to the characterization of an interpolation space between a space of distributions and a space of smooth functions. Although the theory of interpolation spaces is very well developed and one already know to characterize the interpolation spaces for Sobolev spaces of positive and negative indices, we have not found in the (huge) literature a result which covers the problem we are concerned with. So, although our result may be viewed as an interpolation result, it is a new one. As an application we discussed the regularity of the law of a Wiener functional under a Hörmander type non degeneracy condition. These papers will appear in Annals of Probability.

7.6.3. Regularity of the solution of jump type equations

Continuing the above work we study, in collabration with Lucia Caramellino, the regularity of the solution of jump type equations. This subject has been extensively treated in the literarure using different hypothesis and different variants of Malliavin calculus adapted to equations with jumps. The case of Poisson Point measures with absolutely continuous intensity measure is already well understood with the paper of Bichteler, Garereux and Jacod in the 80's. But the case of discrete intensity measures is more subtle. In this case J. Picard has succeded to obtain regularity results using a variant of Malliavin Calculus based on finite differences. We work also in this framework but we do not use directly some variant of Malliavin calculus but we use an interpolation argument . These are still working papers.

7.6.4. An invariance principle for stochastic series (U- Statistics)

In collaboration with L. Caramellino we work on invariance principles for stochastic series of polynomial type. In the case of polynomials of degree one we must have the classical Central Limit Theorem (for random variables which are not identically distributed). For polynomials of higher order we are in the framework of the so called U statistics which have been introduced by Hoffdings in the years 1948 and which play an important role in modern statistics. Our contribution in this topic concerns convergence in total variation distance for this type of objects. We use abstract Malliavin calculus and more generally, the methods mentioned in the above paragraph.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Consortium PREMIA, Natixis - Inria
- Consortium PREMIA, Crédit Agricole CIB - Inria

8.2. Bilateral Grants with Industry

Chair Ecole Polytechnique-ENPC-UPMC-Société Générale "Financial Risks" of the Risk fondation : A. Alfonsi, B. Jourdain, B. Lapeyre

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

- ANR Stab 2013-2016, Participant : B. Jourdain, Partners : Lyon 1, Paris-Dauphine
- ANR Cosmos 2015-2018, Participant: B. Jourdain ; Partners : Ecole des Ponts, Telecom, INIRIA Rennes and IBPC

9.1.2. 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).
- Department of Mathematics, University of Manchester (Tusheng Zhang, currently in charge of an EU-ITN program on BSDEs and Applications).
- Kensas University (Yaozhong Hu)
- Mannheim University (Alexander Schied, Chair of Mathematics in Business and Economics, Department of Mathematics)
- 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), 2 months

9.3.1.1. Internships

- Babacar Diallo [Inria, Trainee, from Mar 2016 until Aug 2016]
- Nicolas Le Mouel [Inria, Trainee, from Jul 2016 until Oct 2016]
- Mouad Ramil [Inria, Trainee, from Mar 2016 until Aug 2016]

9.3.2. Visits to International Teams

9.3.2.1. Research Stays Abroad

- Vlad Bally visited Tor Vergata University, Roma. (Collboration with Lucia Caramellino)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

- B. Jourdain (with B. Bouchard and E. Gobet): organization of the 2015-2016 thematic semester on Monte Carlo methods financed by the Institute Louis Bachelier, and the closing conference.
- A. Alfonsi: Co-organizer of the working group seminar of MathRisk “Méthodes stochastiques et finance”. http://cermics.enpc.fr/~alfonsi/GTMSF.html
- A.Sulem : Co-organiser of the joint working group seminar MathRisk/LPMA, University Paris-Diderot :"Mathematical finance and numerical probability".

10.1.1.1. Member of the Organizing Committees

J. Lelong:

- journées MAS 2016, Grenoble.
- CEMRACS 2017
- Les Journées de Probabilités 2017
- session organizer at CANUM 2016
- Session organizer at “The International Conference on Monte Carlo techniques”, 2016, Paris

10.1.2. Scientific Events Selection

10.1.2.1. Reviewer

A. Sulem is Reviewer for *Mathematical Reviews*

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- R. Elie

  Associate editor of *SIAM Journal on Financial Mathematics (SIFIN)* (since November 2014)
- D. Lamberton

  Associate editor of
  - Mathematical Finance,
  - Associate editor of ESAIM Probability & Statistics
- A. Sulem

  Associate editor of
  - 2011- Present: *Journal of Mathematical Analysis and Applications (JMAA)*
  - 2009- Present: *International Journal of Stochastic Analysis (IJSA)*
  - 2008- Present: *SIAM Journal on Financial Mathematics (SIFIN)*

10.1.3.2. Reviewer - Reviewing Activities

The members of the team reviewed numerous papers for numerous journals .

10.1.4. Invited Talks

- A. Alfonsi
  - January 15th 2016: "Wishart processes: MLE estimation and interest rate modelling", North British Probability seminar, Edinburgh.
  - February 5th 2016: "Dynamic optimal execution in a mixed-market-impact Hawkes price model", Frontiers in Stochastic Modeling for Finance, Padua, Italy
  - October 7th 2016: "Maximum Likelihood Estimation for Wishart processes" WORKSHOP USPC-NUS Models and numerical methods for financial risk management, Paris Diderot.
  - November 17th 2016: "Extension and calibration of a Hawkes-based optimal execution model", SIAM Conference on Financial Mathematics & Engineering, Austin.
  - December 9th 2016: "Optimal Execution in a Hawkes Price Model and Calibration", Market Microstructure Confronting many viewpoints #4, Paris.
- B. Jourdain
  - Seminar of the chair“financial risks”, June 3rd 2016 : Strong convergence properties of the Ninomiya-Victoir scheme and applications to multilevel Monte Carlo
  - Seminar Mathrisk P7, September 22th 2016 : Existence for a calibrated regime-switching local volatility model
- C. Labart
  - Frontiers in stochastic modelling for finance, Padua and Venice, Italy, February 2016.
  - Closing International Conference of Thematic cycle on Monte-Carlo techniques, Paris, July 2016.
- J. Lelong
  - CANUM 2016
  - Seminar on Insurance Mathematics and Stochastic Finance at ETH Zurich, May 2016.
  - Journées MAS 2016.
  - Closing International Conference of Thematic cycle on Monte-Carlo techniques, Paris, July 2016
  - Frontiers in stochastic modelling for finance, Padua and Venice, Italy, February 2016
- A. Sulem
  - Stochastic analysis, control and games with applications to financial economics, University of Leeds, November 2016.
  - National University of Singapore/ Université Paris-Diderot workshop on quantitative finance, October 2016, Paris.
  - Abel Symposium 2016 "Computation and Combinatorics in Dynamics, Stochastics and Control", August 2016, Barony Rosendal, Norway. http://hans.munthe-kaas.no/AbelSymp2016/
  - Simulation of Stochastic graphs and applications, closing conference on “ Monte-Carlo techniques”, Paris, July 2016
  - Conference "Frontiers in Stochastic Modelling for Finance", Padua and Venice, Italy, February, 2016. https://www.maths.univ-evry.fr/conferences/padova2016/index_ws.htm
  - "Actuarial and Financial Mathematics Conference" (Plenary talk) , February 2016, Brussels, Belgium.
    http://www.afmathconf.ugent.be
- A. Zanette

  "Hybrid tree-finite difference methods for the Heston, Bates and Heston Hull-White models". SIMAI Politecnico di Milano 2016.

10.1.5. Research Administration

- A. Sulem : Member of the Committee for technology development, Inria Paris
- B. Jourdain : Head of the doctoral school MSTIC, university Paris-Est

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Undergraduate programs

A. Alfonsi: ‘Probabilités”, first year course at the Ecole des Ponts.

B. Jourdain :

- course "Mathematical finance", 2nd year ENPC

- course "Introduction to probability theory", 1st year, Ecole Polytechnique

B. Jourdain, B. Lapeyre course "Monte-Carlo methods", 3rd year ENPC and Master Recherche Mathématiques et Application, Université Paris-Est Marne-la-Vallée

Graduate programs

A. Alfonsi:

- “Traitement des données de marché : aspects statistiques et calibration”, lecture for the Master at UPEMLV.

- “Mesures de risque”, Master course of UPEMLV and Paris VI.

- Professeur chargé de cours at Ecole Polytechnique

J.-F. Delmas, B.Jourdain course "Jump processes with applications to energy markets", 3rd year ENPC and Master Recherche Mathématiques et Application, Université Paris-Est Marne-la-Vallée

B.Jourdain

- course "Stochastic numerical methods", 3rd year, Ecole Polytechnique

- projects in finance and numerical methods, 3rd year, Ecole Polytechnique

A. Sulem

- "Finite difference for PDEs in Finance", Master 2 MASEF, Université Paris IX-Dauphine, Département Mathématiques et Informatique de la Décision et des Organisations (MIDO), 18h.

- Master of Mathematics, University of Luxembourg, 22 h lectures and responsible of the module "Numerical Methods in Finance".

Doctoral programs

A. Sulem: International summer school in mathematical finance, University of Alberta in Edmonton, Canada "Informational and Imperfect Financial Markets", https://www.pims.math.ca/scientific-event/160625-pssmf (5 lectures)

10.2.2. Supervision

- PhD :

  Anis Al Gerbi : "Ninomiya-Victoir scheme: strong convergence, asymptotics for the normalized error and multilevel Monte Carlo methods", Université Paris-Est supervised by B. Jourdain and E. Clément, defended on October 10 2016
- PhD in progress :

  Rui Chen (Fondation Sciences Mathématiques de Paris grant), "Stochastic Control of mean field systems and applications to systemic risk, from September 2014, Université Paris-Dauphine, Superviser: A. Sulem

Marouen Iben Taarit , “ On CVA and XVA computations ”, CIFRE Natixis/ENPC, Adviser: Bernard Lapeyre

Giulia Terenzi , "American options in complex financial models", Université Paris-Est Marne-la-Vallée, Supervisors: Damien Lamberton and Lucia Caramellino, from University Tor Vergata, Rome

Alexandre Zhou (started November 2015) "Analysis of stochastic particle methods applied to finance", supervised by B.Jourdain

10.2.3. Juries

- B. Jourdain

  - PhD of Khaled Salhi, defended on December 5, University of Lorraine

  - Reviewer for the PhD of Anthony Le Cavil, defended on December 9, University Paris-Saclay

- A. Sulem

  PhD Richàrd Fischer, *Modélisation de la dépendance pour des statistiques d'ordre et estimation non-paramétrique*, (Modelling the dependence of order statistics and nonparametric estimation), (Jury chair), defended on September 30 2016, Ecole des Ponts.

11. Bibliography

Major publications by the team in recent years

[1] L. ABBAS-TURKI, B. LAPEYRE.*American options by Malliavin calculus and nonparametric variance and bias reduction methods*, in "SIAM J. Financ. Math.", 2012, vol. 3, n^{o} 1, p. 479-510.

[2] A. AHDIDA, A. ALFONSI.*Exact and high order discretization schemes for Wishart processes and their affine extensions*, in "Annals of Applied Probability", 2013, vol. 23, n^{o} 3, p. 1025-1073 [*DOI :* 10.1214/12-AAP863], http://hal.inria.fr/hal-00491371.

[3] A. ALFONSI.*High order discretization schemes for the CIR process: Application to affine term structure and Heston models*, in "Stochastic Processes and their Applications", 2010, vol. 79, p. 209-237, http://www.ams.org/journals/mcom/2010-79-269/S0025-5718-09-02252-2/home.html.

[4] A. ALFONSI.*Affine diffusions and related processes: simulation, theory and applications*, Bocconi and Springer Series, Mathematics statistics, finance and economics, Springer, 2015.

[5] A. ALFONSI, B. JOURDAIN, A. KOHATSU-HIGA.*Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme*, in "Annals of Applied Probability", 2014, https://hal-enpc.archives-ouvertes.fr/hal-00727430.

[6] A. ALFONSI, A. SCHIED.*Optimal Trade Execution and Absence of Price Manipulations in Limit Order Book Models*, in "SIAM J. Finan. Math.", 2010, vol. 1, n^{o} 1, p. 490-522, http://epubs.siam.org/doi/abs/10.1137/090762786.

[7] 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.

[8] V. BALLY, N. FOURNIER.*Regularization properties od the 2D homogenuos Bolzmann equation without cutoff*, in "PTRF", 2011, n^{o} 151, p. 659-670.

[9] M. JEUNESSE, B. JOURDAIN.*Regularity of the American put option in the Black-Scholes model with general discrete dividends*, in "Stochastic Processes and their Applications", 2012, vol. 112, p. 3101-3125, DOI:10.1016/j.spa.2012.05.009, http://hal.archives-ouvertes.fr/hal-00633199.

[10] B. JOURDAIN.*Probabilités et statistique*, Ellipses, 2009.

[11] D. LAMBERTON, M. MIKOU.*Exercise boundary of the American put near maturity in an exponential Lévy model*, in "Finance and Stochastics", 2013, vol. 17, n^o 2, p. 355-394.

[12] D. LAMBERTON, M. ZERVOS.*On the optimal stopping of a one-dimensional diffusion*, in "Electronic Journal of Probability", 2013, vol. 18, n^o 34, p. 1-49.

[13] M.-C. QUENEZ, A. SULEM.*BSDEs with jumps, optimization and applications to dynamic risk measures*, in "Stochastic Processes and their Applications", March 2013, vol. 123, n^o 8, p. 3328-3357 [*DOI :* 10.1016/J.SPA.2013.02.016], https://hal.inria.fr/hal-00709632.

[14] 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^o 9, 23, https://hal.inria.fr/hal-00773708.

[15] A. SULEM.*Numerical Methods implemented in the Premia Software*, March-April 2009, vol. 99, Special issue of the Journal “Bankers, Markets, Investors”, Introduction by Agnès Sulem (Ed) and A. Zanette.

[16] B. ØKSENDAL, A. SULEM.*Applied Stochastic Control of Jump Diffusions*, Universitext, Second Edition, Springer, Berlin, Heidelberg, New York, 257 pages 2007.

[17] B. ØKSENDAL, A. SULEM.*Singular stochastic Control and Optimal stopping with partial information of Itô-Lévy processes*, in "SIAM J. Control & Optim.", 2012, vol. 50, n^o 4, p. 2254–2287, http://epubs.siam.org/doi/abs/10.1137/100793931.

[18] B. ØKSENDAL, A. SULEM, T. ZHANG.*Singular Control and Optimal Stopping of SPDEs, and Backward SPDEs with Reflection*, in "Mathematics of Operations Research", June 2013, https://hal.inria.fr/hal-00919136.

Publications of the year

Articles in International Peer-Reviewed Journal

[19] 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://hal-enpc.archives-ouvertes.fr/hal-01188675.

[20] A. ALFONSI, P. BLANC.*Dynamic optimal execution in a mixed-market-impact Hawkes price model*, in "Finance and Stochastics", January 2016 [*DOI :* 10.1007/S00780-015-0282-Y], https://hal-enpc.archives-ouvertes.fr/hal-00971369.

[21] A. ALFONSI, P. BLANC.*Extension and calibration of a Hawkes-based optimal execution model*, in "Market microstructure and liquidity", August 2016 [*DOI :* 10.1142/S2382626616500052], https://hal-enpc.archives-ouvertes.fr/hal-01169686.

[22] A. Alfonsi, A. Kebaier, C. Rey.*Maximum Likelihood Estimation for Wishart processes*, in "Stochastic Processes and their Applications", November 2016 [*DOI :* 10.1016/j.spa.2016.04.026], https://hal-enpc.archives-ouvertes.fr/hal-01184310.

[23] A. Alfonsi, C. Labart, J. Lelong.*Stochastic Local Intensity Loss Models with Interacting Particle Systems*, in "Mathematical Finance", 2016, vol. 26, n^o 2, p. 366–394 [*DOI :* 10.1111/mafi.12059], https://hal.archives-ouvertes.fr/hal-00786239.

[24] A. Alfonsi, A. Schied, F. Klöck.*Multivariate transient price impact and matrix-valued positive definite functions*, in "Mathematics of Operations Research", March 2016 [*DOI :* 10.1287/moor.2015.0761], https://hal-enpc.archives-ouvertes.fr/hal-00919895.

[25] L. Badouraly Kassim, J. Lelong, I. Loumrhari.*Importance sampling for jump processes and applications to finance*, in "Journal of Computational Finance", December 2016, vol. 19, n^o 2, p. 109-139, https://hal.archives-ouvertes.fr/hal-00842362.

[26] 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 [*DOI :* 10.1137/140965910], https://hal-upec-upem.archives-ouvertes.fr/hal-00979936.

[27] M. Clausel, J.-F. Coeurjolly, J. Lelong.*Stein estimation of the intensity of a spatial homogeneous Poisson point process*, in "Annals of Applied Probability", June 2016, vol. 26, n^o 3, p. 1495-1534 [*DOI :* 10.1214/15-AAP1124], https://hal.archives-ouvertes.fr/hal-01024648.

[28] C. De Luigi, J. Lelong, S. Maire.*Adaptive numerical integration and control variates for pricing Basket Options*, in "Applied Numerical Mathematics", 2016, vol. 100, 17, https://hal.archives-ouvertes.fr/hal-00746872.

[29] 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://hal.archives-ouvertes.fr/hal-01006131.

[30] 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://hal.archives-ouvertes.fr/hal-01114996.

[31] R. Dumitrescu, M.-C. Quenez, A. Sulem.*A Weak Dynamic Programming Principle for Combined Optimal Stopping/Stochastic Control with Ef-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.

[32] 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.

[33] 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.

[34] G. E. Espinosa, C. Hillairet, B. Jourdain, M. Pontier.*Reducing the debt : is it optimal to outsource an investment?*, in "Mathematics and Financial Economics", March 2016, vol. 10, n$^{\rm o}$ 4, p. 457-493, https://hal-enpc.archives-ouvertes.fr/hal-00824390.

[35] 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$^{\rm o}$ 1, p. 131-170, https://hal.archives-ouvertes.fr/hal-00608977.

[36] L. Goudenège, M. Andrea, A. Zanette.*Pricing and hedging GLWB in the Heston and in the Black–Scholes with stochastic interest rate models* , in "Insurances : Mathematics and Economics", September 2016 [*DOI :* 10.1016/j.insmatheco.2016.05.018], https://hal.archives-ouvertes.fr/hal-01390968.

[37] B. Jourdain, J. Reygner.*A multitype sticky particle construction of Wasserstein stable semigroups solving one-dimensional diagonal hyperbolic systems with large monotonic data*, in "Journal of Hyperbolic Differential Equations", September 2016, vol. 13, n$^{\rm o}$ 3, p. 441-602 [*DOI :* 10.1142/S0219891616500144], https://hal-enpc.archives-ouvertes.fr/hal-01100604.

[38] 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$^{\rm o}$ 9, p. 4963-4996 [*DOI :* 10.3934/dcds.2016015], https://hal-enpc.archives-ouvertes.fr/hal-01171261.

[39] J. Mint Moustapha, B. Jourdain, D. Daucher.*A probabilistic particle approximation of the "Paveri-Fontana" kinetic model of traffic flow*, in "SMAI Journal of Computational Mathematics", November 2016, vol. 2, p. 229-253 [*DOI :* 10.5802/smai-jcm.15], https://hal.archives-ouvertes.fr/hal-01415636.

[40] 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.

Scientific Books (or Scientific Book chapters)

[41] 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, 319 [*DOI :* 10.1007/978-3-319-23425-0], https://hal.inria.fr/hal-01406649.

[42] B. Øksendal, A. Sulem, T. Zhang.*A stochastic HJB equation for optimal control of forward-backward SDEs*, in "The Fascination of Probability, Statistics and their Applications", Springer Verlag, 2016, 11, https://hal.inria.fr/hal-01406655.

Other Publications

[43] A. Al Gerbi, B. Jourdain, E. Clément.*Asymptotic error distribution for the Ninomiya-Victoir scheme in the commutative case*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01390897.

[44] A. Al Gerbi, B. Jourdain, E. Clément.*Asymptotics for the normalized error of the Ninomiya-Victoir scheme*, January 2016, working paper or preprint, https://hal-enpc.archives-ouvertes.fr/hal-01259915.

[45] A. ALFONSI, J. CORBETTA, B. JOURDAIN.*Evolution of the Wasserstein distance between the marginals of two Markov processes*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01390887.

[46] V. BALLY, L. CARAMELLINO.*An Invariance Principle for Stochastic Series II. Non Gaussian Limits*, December 2016, working paper or preprint, https://hal-upec-upem.archives-ouvertes.fr/hal-01413533.

[47] V. BALLY, L. CARAMELLINO.*Regularity of Wiener functionals under a Hörmander type condition of order one*, December 2016, working paper or preprint, https://hal-upec-upem.archives-ouvertes.fr/hal-01413556.

[48] V. BALLY, L. CARAMELLINO, P. PIGATO.*Diffusions under a local strong Hörmander condition. Part I: density estimates*, December 2016, working paper or preprint, https://hal-upec-upem.archives-ouvertes.fr/hal-01413546.

[49] V. BALLY, L. CARAMELLINO, P. PIGATO.*Diffusions under a local strong Hörmander condition. Part II: tube estimates*, December 2016, working paper or preprint, https://hal-upec-upem.archives-ouvertes.fr/hal-01413545.

[50] G. FORT, B. JOURDAIN, T. LELIEVRE, G. STOLTZ.*Convergence and efficiency of adaptive importance sampling techniques with partial biasing*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01389996.

[51] B. JOURDAIN, A. ZHOU.*Fake Brownian motion and calibration of a Regime Switching Local Volatility model*, July 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01341212.

[52] A. KEBAIER, J. LELONG.*Coupling Importance Sampling and Multilevel Monte Carlo using Sample Average Approximation*, May 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01214840.

[53] S. LABBÉ, J. LELONG.*Stochastic modelling of thermal effects on a ferromagnetic nano particle*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01337197.

[54] J. LELONG.*Pricing American options using martingale bases*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01299819.

[55] A. LIONNET, G. DOS REIS, L. SZPRUCH.*Convergence and qualitative properties of modified explicit schemes for BSDEs with polynomial growth*, July 2016, working paper or preprint, https://hal.inria.fr/hal-01404132.

References in notes

[56] *PREMIA: un outil d'évaluation pour les options*, NextOption, 2006.

[57] M. AKIAN, J. MENALDI, A. SULEM.*On an Investment-Consumption model with transaction costs*, in "SIAM J. Control and Optim.", 1996, vol. 34, p. 329-364.

[58] M. AKIAN, A. SULEM, M. TAKSAR.*Dynamic optimisation of long term growth rate for a portfolio with transaction costs - The logarithmic utility case*, in "Mathematical Finance", 2001, vol. 11, p. 153-188.

[59] A. ALFONSI, A. SCHIED.*Optimal Trade Execution and Absence of Price Manipulations in Limit Order Book Models*, in "SIAM J. Finan. Math.", 2010, vol. 1, p. 490-522.

[60] H. AMINI, R. CONT, A. MINCA.*Resilience to Contagion in Financial Networks*, in "Mathematical Finance", 2013.

[61] V. BALLY.*An elementary introduction to Malliavin calculus*, Inria, Rocquencourt, February 2003, n° 4718, http://hal.inria.fr/inria-00071868.

[62] V. BALLY, L. CARAMELLINO, A. ZANETTE.*Pricing American options by a Monte Carlo method using a Malliavin calculus approach*, in "Monte Carlo methods and applications", 2005, vol. 11, n° 2, p. 97–133.

[63] D. BELL.*The Malliavin Calculus*, Pitman Monographs and Surveys in Pure and Applied Math., Longman and Wiley, 1987, n° 34.

[64] T. BIELECKI, J.-P. CHANCELIER, S. PLISKA, A. SULEM.*Risk sensitive portfolio optimization with transaction costs*, in "Journal of Computational Finance", 2004, vol. 8, p. 39-63.

[65] F. BLACK, M. SCHOLES.*The pricing of Options and Corporate Liabibilites*, in "Journal of Political Economy", 1973, vol. 81, p. 637-654.

[66] J.-P. CHANCELIER, B. LAPEYRE, J. LELONG.*Using Premia and Nsp for Constructing a Risk Management Benchmark for Testing Parallel Architecture*, in "Concurrency and Computation: Practice and Experience", June 2014, vol. 26, n° 9, p. 1654-1665 [*DOI :* 10.1002/CPE.2893], https://hal.archives-ouvertes.fr/hal-00447845.

[67] I. ELSANOSI, B. ØKSENDAL, A. SULEM.*Some Solvable Stochastic control Problems with Delay*, in "Stochastics and Stochastics Reports", 2000.

[68] J. D. FONSECA, M. MESSAOUD.*Libor Market Model in Premia: Bermudan pricer, Stochastic Volatility and Malliavin calculus*, in "Bankers, Markets, Investors", March-April 2009, vol. Special report: Numerical Methods implemented in the Premia Software, n° 99, p. 44–57.

[69] E. FOURNIÉ, J.-M. LASRY, J. LEBUCHOUX, P.-L. LIONS.*Applications of Malliavin calculus to Monte Carlo methods in Finance, II*, in "Finance & Stochastics", 2001, vol. 2, n° 5, p. 201-236.

[70] E. FOURNIÉ, J.-M. LASRY, J. LEBUCHOUX, P.-L. LIONS, N. TOUZI.*An application of Malliavin calculus to Monte Carlo methods in Finance*, in "Finance & Stochastics", 1999, vol. 4, n° 3, p. 391-412.

[71] N. C. FRAMSTAD, B. ØKSENDAL, A. SULEM.*Optimal Consumption and Portfolio in a Jump Diffusion Market with Proportional Transaction Costs*, in "Journal of Mathematical Economics", 2001, vol. 35, p. 233-257.

[72] J. GARNIER, G. PANANICOLAOU, T.-W. YANG.*Large deviations for a mean field model of systemic risk*, 2012, Manuscript, arXiv:1204.3536.

[73] P. GASSIAT, H. PHAM, M. SIRBU.*Optimal investment on finite horizon with random discrete order flow in illiquid markets*, in "International Journal of Theoretical and Applied Finance", 2010, vol. 14, p. 17-40.

[74] Y. KABANOV, M. SAFARIAN.*Markets with Transaction Costs: Mathematical Theory*, Springer Verlag, 2009.

[75] C. LABART, J. LELONG.*Pricing Parisian Options using Laplace transforms*, in "Bankers, Markets, Investors", March-April 2009, vol. Special report: Numerical Methods implemented in the Premia Software, n^{o} 99, p. 29–43.

[76] D. LAMBERTON, B. LAPEYRE, A. SULEM.*Application of Malliavin Calculus to Finance*, in "special issue of the journal Mathematical Finance", January 2003.

[77] P. MALLIAVIN.*Stochastic calculus of variations and hypoelliptic operators*, in "Proc. Inter. Symp. on Stoch. Diff. Equations", Kyoto, Wiley 1978, 1976, p. 195-263.

[78] P. MALLIAVIN, A. THALMAIER.*Stochastic Calculus of variations in Mathematical Finance*, Springer Finance, Springer Verlag, 2006.

[79] A. MINCA.*Modélisation mathématique de la contagion de défaut; Mathematical modeling of financial contagion*, Université Pierre et Marie Curie, Paris 6, September 5 2011.

[80] D. NUALART.*The Malliavin Calculus and Related Topics*, Springer–Verlag, 1995.

[81] D. OCONE, I. KARATZAS.*A generalized representation formula with application to optimal portfolios*, in "Stochastics and Stochastic Reports", 1991, vol. 34, p. 187-220.

[82] D. OCONE.*A guide to the stochastic calculus of variations*, in "Stochastic Analysis and Related Topics", H. KOERZLIOGLU, S. ÜSTÜNEL (editors), Lecture Notes in Math.1316, 1987, p. 1-79.

[83] N. PRIVAULT, X. WEI.*Calibration of the LIBOR market model - implementation in Premia*, in "Bankers, Markets, Investors", March-April 2009, vol. Special report: Numerical Methods implemented in the Premia Software, n^{o} 99, p. 20–29.

[84] F. RUSSO, P. VALLOIS.*Stochastic calculus with respect to continuous finite quadratic variation processes*, in "Stochastics and Stochastics Reports", 2000, vol. 70, p. 1–40.

[85] A. SULEM.*Dynamic Optimisation for a mixed Portfolio with transaction costs*, in "Numerical methods in Finance", 1997, p. 165-180, edited by L.C.G. Rogers and D.Talay, Cambridge University Press, Publications of the Newton Institute.

[86] A. SULEM, A. ZANETTE.*Premia: A Numerical Platform for Pricing Financial Derivatives*, in "Ercim News", July 2009, vol. 78.

[87] U. ÇETIN, R. JARROW, P. PROTTER.*Liquidity risk and arbitrage pricing theory*, in "Finance and Stochastics", 2004, vol. 8.

[88] B. ØKSENDAL, A. SULEM, T. ZHANG.*Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations*, in "Advances in Applied Probability", 2011, vol. 43, p. 572-596.

[89] B. ØKSENDAL, A. SULEM.*Optimal Consumption and Portfolio with both fixed and proportional transaction costs: A Combined Stochastic Control and Impulse Control Model*, in "SIAM J. Control and Optim.", 2002, vol. 40, p. 1765-1790.

[90] B. ØKSENDAL, A. SULEM.*Optimal stochastic impulse control with delayed reaction*, in "Applied Mathematics and Optimization", 2008, vol. 58, p. 243-255.

[91] B. ØKSENDAL.*An Introduction to Malliavin Calculus with Applications to Economics*, in "Lecture Notes from a course given 1996 at the Norwegian School of Economics and Business Administration (NHH)", September 1996, NHH Preprint Series.

Team MIMOVE

Middleware on the Move

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
Distributed Systems and middleware

Table of contents

Team MIMOVE

Creation of the Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

1.2.1. - Dynamic reconfiguration
1.2.4. - QoS, performance evaluation
1.2.5. - Internet of things
1.2.6. - Sensor networks
1.2.7. - Cyber-physical systems
1.3. - Distributed Systems
1.4. - Ubiquitous Systems
1.5. - Complex systems
1.5.1. - Systems of systems
1.5.2. - Communicating systems
2.5. - Software engineering
2.6.2. - Middleware

Other Research Topics and Application Domains:

6.4. - Internet of things
8.2. - Connected city
8.5.1. - Participative democracy

1. Members

Research Scientists

Nikolaos Georgantas [Team leader, Inria, Researcher]
Valérie Issarny [Inria, Inria@SiliconValley, Senior Researcher, HDR]

Faculty Member

Françoise Sailhan [CNAM, Associate Professor, *on sabbatical leave (CRCT)*]

Technical Staff

Cong Kinh Nguyen [Inria, EIT Digital Env&You project]
Pierre-Guillaume Raverdy [Inria, SED, from Jun 2016, *part-time*]
Fadwa Rebhi [Inria, ADT CityLab Platform until Oct 2016; then, EIT Digital Env&You project]

PhD Students

Amel Belaggoun [CEA, *co-advised PhD*, *part-time*]
Georgios Bouloukakis [Inria, UPMC]
Radha Pallavali [Inria, UPMC, from Nov 2016]

Post-Doctoral Fellows

Rachit Agarwal [Inria, H2020 FIESTA-IoT project]
Rafael Angarita Arocha [Inria, EIT Digital CivicBudget project, from May 2016]
Eya Ben Ahmed [Inria]
Cristhian Parra Trepowski [Inria, EIT Digital CivicBudget project]

Visiting Scientist

Garvita Bajaj [Indraprastha Institute of Information Technology, New Delhi, *PhD intern*, from May 2016 until Sept 2016]

Administrative Assistant

Maryse Desnous [Inria]

Others

Siddhartha Dutta [Intern, from May 2016 until July 2016]

Christelle Rohaut [Intern, from Mar 2016 until Jun 2016]

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. In addition, mobile distributed systems are most often characterized by the absence of any centralized control. This results in peer interaction between system entities, ad hoc or opportunistic relations between them, and relations reflecting the social behavior of the systems' users. The above features span the application, middleware and higher network layers of such systems in a cross-layer fashion.

This challenging environment is characterized by high complexity raising key research questions:

- How to deal with the extreme uncertainty, when developing and running mobile distributed systems, resulting from the openness and constant evolution of their execution environment?
- How to manage the ultra large scale and dynamicity resulting from millions or even billions of mobile devices that interact with the physical environment through sensing and actuation?
- How to leverage the social aspects arising out of billions of users carrying personal devices in order to enable powerful, critical-mass social sensing and actuation?

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:

- **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 [23], [29]. 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 [28].

- **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 [68], 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 [65]. MiMove's research investigates techniques for efficient coordination of future mobile sensing and actuation systems with a special focus on their dependability.
- **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 [82], [46], [79], 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 [32], [59], [60].

Outcomes of the three identified research topics are implemented as middleware-level functionalities giving rise to software architectures for mobile distributed systems and enabling practical application and assessment of our research. Furthermore, although our research results can be exploited in numerous application domains, we focus in particular on the domain of smart cities, which is an area of rapidly growing social, economic and technological interest.

3. Research Program

3.1. Introduction

MiMove targets research enabling next-generation mobile distributed systems, from their conception and design to their runtime support. These systems are challenged by their own success and consequent massive growth, as well as by the present and future, fast evolving, global networking and computing environment. This context is well-captured by the Future Internet vision, whose mobile constituents are becoming the norm rather than the exception. MiMove's research topics relate to a number of scientific domains with intensive ongoing research, such as ubiquitous computing, self-adaptive systems, wireless sensor networks, participatory sensing and social networks. In the following, we discuss related state-of-the-art research – in particular work focusing on middleware for mobile systems – and we identify the open research challenges that drive our work.

3.2. Emergent mobile distributed systems

Emergent mobile distributed systems promise to provide solutions to the complexity of the current and future computing and networking environments as well as to the ever higher demand for ubiquitous mobile applications, in particular being a response to the volatile and evolving nature of both the former and the latter. Hence, such systems have gained growing interest in the research literature. Notably, research communities have been formed around *self-adaptive systems* and *autonomic systems*, for which various overlapping definitions exist [72]. Self-adaptive systems are systems that are able to adapt themselves to uncertain execution environments, while autonomic systems have been defined as having one or more characteristics known as *self-** properties, including self-configuring, self-healing, self-optimizing and self-protecting [54]. Self-adaptive or autonomic systems typically include an adaptation loop comprising *modeling*, *monitoring*, *analyzing*, *deciding* and *enactment* processes. The adaptation loop provides feedback about changes in the system and its environment to the system itself, which adjusts itself in response. Current research on emergent distributed systems, including mobile ones, addresses all the dimensions of the adaptation loop [31], [25], [61], [83].

In our previous work, we introduced the paradigm of *emergent middleware*, which enables networked systems with heterogeneous behaviors to coordinate through adequate interaction protocols that emerge in an automated way [50], [28], [26]. A key point of that work is the combined study of the application- and middleware-layer behaviors, while current efforts in the literature tend to look only at one layer, either the application [48] or the middleware [19], [49], and take the other for granted (i.e., homogeneous, allowing direct coordination). Furthermore, the uncertainty of the computing and networking environments that is intrinsic to emergent mobile distributed systems [41] calls for taking into account also the underlying network and computational resources in a cross-layer fashion. In another line of work, we studied cross-integration of heterogeneous interaction paradigms at the middleware layer (message passing versus event-based and data sharing), where we investigate functional and QoS semantics of paradigms across their interconnections [43], [53]. Our focus there is to grasp the relation between individual and end-to-end semantics when bridging heterogeneous interaction protocols. In contrast, existing research efforts typically focus on emergent or evolving properties in homogeneous settings [42]. Last but not least, integrating heterogeneous mobile distributed systems into emergent compositions raises the question of dependability. More specifically, the overall correctness of the composition with respect to the individual requirements of the constituent systems can be particularly hard to ensure due to their heterogeneity. Again, current approaches typically deal with homogeneous constraints for dependability [39], [85], [40] with few exceptions [38].

As evident from the above, there is considerable interest and intensive research on emergent mobile distributed systems, while at the same time there are key research questions that remain open despite initial relevant work, including ours, which are summarized in the following:

- How to effectively deal with the combined impact on emergent properties of the different functional layers of mobile distributed systems (e.g., [50], [28], [26], [69])?
- How to perceive and model emergent properties in space and in time across volatile compositions of heterogeneous mobile distributed systems (e.g., [43], [53])?
- How to produce dependable emergent mobile distributed systems, i.e., systems that correctly meet their requirements, despite uncertainty in their emergence and execution exacerbated by heterogeneity (e.g., [38])?

3.3. Large-scale mobile sensing and actuation

In the past decade, the increasingly low cost of MEMS [0] devices and low-power microprocessors has led to a significant amount of research into mobile sensing and actuation. The results of this are now reaching the general public, going beyond the largely static use of sensors in scenarios such as agriculture and waste-water management, into increasingly *mobile* systems. These include sensor-equipped smartphones and personal wearable devices focused on the idea of a “quantified self", gathering data about a user's daily habits in order to enable them to improve their well-being. However, in spite of significant advances, the key challenges of these systems arise from largely the same attributes as those of early envisioned mobile systems, introduced in [76] and re-iterated in [75]: relative resource-poverty in terms of computation and communication, variable

[0] Micro-Electro-Mechanical Systems.

and unreliable connectivity, and limitations imposed by a finite energy source. These remain true even though modern mobile devices are significantly more powerful compared to their ancestors; the work we expect them to do has increased, and the computation and storage abilities available through fixed infrastructure such as the cloud are larger by order of magnitudes than any single mobile device. The design of algorithms and protocols to efficiently coordinate the sensing, processing, and actuation capabilities of the large number of mobile devices in future systems is a core area of MiMove's research.

Precisely, the focus of MiMove's research interests lies mostly in the systems resulting from the increased popularity of sensor-equipped smart devices that are carried by people, which has led to the promising field of *mobile phone sensing* or *mobile crowd-sensing* [58], [55]. The paradigm is powerful, as it allows overcoming the inherent limitation of traditional sensing techniques that require the deployment of dedicated fixed sensors (e.g., see work on noise mapping using the microphones in users' telephones [70]). Specifically, we are interested in the challenges below, noting that initial work to address them already exists, including that by team members:

- How to efficiently manage the large scale that will come to the fore when millions, even billions of devices will need to be managed and queried simultaneously (e.g., [81], [45])?
- How to efficiently coordinate the available devices, including resource-poor mobile devices and the more-capable cloud infrastructure (e.g., [68], [36], [74], [64])?
- How to guarantee dependability in a mobile computing environment (e.g., [34], [80], [30])?
- How to ensure that the overhead of sensing does not lead to a degraded performance for the user (e.g., [56], [36])?

3.4. Mobile social crowd-sensing

Mobile crowd-sensing as introduced in Section 3.3 is further undergoing a transformation due to the widespread adoption of social networking. The resulting mobile *social* crowd-sensing may be qualified as "*people-centric sensing*" and roughly subdivides into two categories [57]: i) *participatory sensing*, and ii) *opportunistic sensing*. Participatory sensing entails direct involvement of humans controlling the mobile devices, while opportunistic sensing requires the mobile device itself to determine whether or not to perform the sensing task. Orthogonally to the above categorization, mobile sensing can be [55]: i) *personal sensing*, mostly to monitor a person's context and well-being; ii) *social sensing*, where updates are about the social and emotional statuses of individuals; or iii) *urban (public) sensing*, where public data is generated by the public and for the public to exploit. Personal sensing is aimed towards personal monitoring and involves one or just a few devices in direct relationship with their custodian. For instance, SoundSense [62] is a system that enables each person's mobile device to learn the types of sounds the owner encounters through unsupervised learning. Another application example relates to the sensing-based detection of the users' transportation mode by using their smartphones [47]. In social sensing, the mobile device or its owner decides what social information to share about the owner or the owner's environment, with an individual or group of friends [55], [37], [52], [21], [66]. Social sensing is mostly participatory. Therefore, it is the custodian of the device who determines when and where data should be generated. Social participatory sensing is closely related to social networking [63]. On the other hand, within opportunistic social sensing, the underlying system is in charge of acquiring needed data through relevant probes, as opposed to having the end-user providing them explicitly [24], [51], [22]. In urban sensing, also known as public sensing, data can be generated by everyone (or their devices) and exploited by everyone for public knowledge, including environment monitoring, or traffic updates [55]. In participatory urban sensing, users participate in providing information about the environment by exploiting the sensors/actuators embedded in their devices (which can be smartphones, vehicles, tablets, etc.) [55]. However data is only generated according to the owner's willingness to participate. Participatory urban sensing is especially characterized by scale issues at the data level, where data is generated by numerous individuals and should be processed and aggregated for knowledge to be inferred, involving adequate data scaling approaches [44]. Ikarus [84] is an example of participatory sensing, where data is collected by a large number of paragliders throughout their flights. The focus is on aggregating the data and rendering the results on a thermal map.

As outlined above, mobile social crowd-sensing has been a very active field of research for the last few years with various applications being targeted. However, effectively enabling mobile social crowd-sensing still raises a number of challenges, for which some early work may be identified:

- How to ensure that the system delivers the right quality of service, e.g., in terms of user-perceived delay, in spite of the resource constraints of mobile systems (e.g., [71])?
- How to guarantee the right level of privacy (e.g., [33], [73])?
- How to ensure the right level of participation from end-users so that mobile sensing indeed becomes a relevant source of accurate knowledge, which relates to eliciting adequate incentive mechanisms [86], in particular based on the understanding of mobile application usage [78], [77]?
- How to enrich sensor-generated content that is quantitative with user-generated one, thereby raising the issue of leveraging highly unstructured data while benefiting from a rich source of knowledge (e.g., sensing the crowdedness of a place combined with the feeling of people about the crowdedness, which may hint on the place's popularity as much as on discomfort)?

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. However, given the importance of conducting empirical studies to assess and nurture our research, we focus on one application area that 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 is 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 intend to focus on, further leveraging our research on emergent mobile distributed systems, large-scale mobile sensing & actuation, and mobile social crowd-sensing.

In a first step, we concentrate on the following specialized applications, which we investigate in close collaboration with other researchers, in particular as part of the dedicated Inria Project Lab *CityLab@Inria*:

- **Democratization of urban data for healthy cities.** The objective here is to integrate the various urban data sources, especially by way of crowd-Xing, to better understand city nuisances 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).

- **Socially-aware urban mobility.** Mobility within mega-cities is known as one of the major challenges to face urgently due to the fact that today's mobility patterns do not scale and to the negative effect on the environment and health. It is our belief that mobile social and physical sensing may significantly help in promoting the use of public transport, which we have started to investigate through empirical study based on the development and release of dedicated apps.
- **Social applications.** Mobile applications are being considered by sociologists as a major vehicle to actively involve citizens and thereby prompt them to become activists. This is especially studied with the Social Apps Lab at UC Berkeley. Our objective is to study such a vehicle from the ICT perspective and in particular elicit relevant middleware solutions to ease the development and development of such "*civic apps*".

Acknowledging the need for collaborative research in the application domain of smart cities, MiMove is heavily involved and actually leading CityLab@Inria [0]. The Inria Project Lab CityLab is focused on the study of ICT solutions promoting social sustainability in smart cities, and involves the following Inria project-teams in addition to MiMove: CLIME, DICE, FUN, MYRIADS, SMIS, URBANET and WILLOW. CityLab further involves strong collaboration with California universities affiliated with CITRIS (Center for Information Technology Research in the Interest of Society) and especially UC Berkeley, in relation with the *Inria@SiliconValley* program. We note that Valérie Issarny acts as scientific manager of Inria@SiliconValley and is currently visiting scholar at CITRIS at UC Berkeley. In this context, MiMove researchers are working closely with colleagues of UC Berkeley, including researchers from various disciplines interested in smart cities (most notably sociologists).

5. Highlights of the Year

5.1. Highlights of the Year

Members of MiMove are co-founders of the Ambiciti start-up (http://ambiciti.io) together with the Inria team CLIME, and the NUMTECH and the Civic Engine SMEs. Ambiciti's technology is a single platform delivering real-time data on street-by-street exposure and risks on multiple environmental pollutants. The platform's technology leverages open data along with cloud, IoT, mobile and data analytics technologies. Ambiciti collects real-time, street-by-street pollution data and provides urban citizens with a means to personalize their decisions with regard to environmental hazards. The aim is to enable citizens to make more informed choices about their activities, personal behavior and location, and to protect their own health. Ambiciti also supplies businesses with crucial data that allows to better inform consumers and to increase the valuation of services (e.g., real estate). Eventually, Ambiciti supports governments in protecting citizens' health and in growing cities more sustainably in providing the necessary urban pollution data. Key elements of the Ambiciti platform include the Ambiciti mobile app that leverages mobile phone sensing middleware solutions to monitor the individual and collective exposure of citizens to environmental pollutions in a resource-efficient way (more at https://www.inria.fr/en/centre/paris/news/ambiciti-an-application-a-start-up). The first version of the Ambiciti App (successor of SoundCity) deals with noise and air pollution. In particular, Inria and the Paris city council were awarded a *Décibel d'Argent* prize for the App (more at https://www.inria.fr/en/centre/paris/news/2016-decibel-d-or-golden-decibel-competition-ambiciti-receives-the-decibel-d-argent-silver-decibel-prize-in-the-research-category).

6. New Software and Platforms

6.1. Introduction

In order to validate our research results and, in certain cases, make them available to specific communities or to the public, our research activities encompass the development of related software as surveyed below.

[0] http://citylab.inria.fr

6.2. VSB: eVolution Service Bus for the Future Internet

Participants: Georgios Bouloukakis, Nikolaos Georgantas [contact], Siddhartha Dutta.

URL: https://tuleap.ow2.org/plugins/git/chorevolution/evolution-service-bus

The *eVolution Service Bus (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 [0]. 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. In this paradigm, a common intermediate bus protocol is used to facilitate interconnection between multiple heterogeneous middleware protocols. Conversion of each protocol to the common bus protocol is done by a component associated to the service/Thing in question and its middleware, called a Binding Component (BC), as it binds the service/Thing to the service bus. We introduce a generic architecture for VSB, which relies on the notion of *Generic Middleware (GM)* connector. GM abstracts interactions among peer components that employ the same middleware protocol through generic *post* and *get* operations, in a unifying fashion for any middleware protocol. We propose an API (application programming interface) for GM and a related generic interface description, which we call *GIDL*, for application components that (abstractly) employ GM. Concrete middleware protocols and related interface descriptions of application components that employ these middleware protocols can be mapped to GM API and GIDL, respectively. Based on these abstractions, we elaborate a generic architecture for BCs, as well as a related method for BC synthesis and refinement for a concrete choreography that includes services/Things with heterogeneous middleware protocols.

The eVolution Service Bus (VSB) presents a significant rethinking of the architecture and the implementation of a service bus destined to serve dynamic choreographies of services but also Things as first-class entities. More specifically, VSB presents the following advancements:

- VSB is a unified interoperability solution for both services and Things participating in choreographies;
- VSB is flexible and lightweight: it is a completely decentralized network of BCs that are deployed as necessary; hence, no BC is needed when a service/Thing employs the same middleware protocol as the one used as common bus protocol;
- VSB provides support for the client-server, publish/subscribe, tuple space and data streaming interaction paradigms;
- Different protocols can be introduced as VSB's common bus protocol with the same easiness as for integrating support for a new middleware protocol of a service/Thing; additionally, there is no need for relying on and/or providing a full-fledged ESB platform;
- While very modular, VSB's architecture includes only few levels of indirection in the processing of primitives when converting between protocols; this makes it simple, lightweight and fast;
- In VSB, mapping between a concrete middleware protocol and the GM paradigm can be performed in different ways, thus enabling to cover all possible interaction cases; there is no unique, fixed mapping limiting the applicability of the solution;
- BC synthesis follows a systematic method allowing for its automation: we have developed related tools that support the automated generation of a BC for a service/Thing from its GIDL interface description.

[0] https://tools.ietf.org/html/rfc7252

VSB is being developed within the H2020 CHOReVOLUTION project (see § 8.2.1.1). It is also based on previous development carried out in the FP7 CHOReOS project [0]. VSB is available for download under open source license.

6.3. Ambiciti App & Platform: Monitoring the Exposure to Environmental Pollution

Participants: Valerie Issarny [contact], Cong Kinh Nguyen, Pierre-Guillaume Raverdy, Fadwa Rebhi.

URL: http://ambiciti.io

Is your exposure to noise too high on certain days? How is air pollution in your street? Will air quality improve in the next hours? Do you want to measure the noise pollution on the way between your home and your office? What pollution levels are considered harmful for your health? Ambiciti (previously SoundCity) provides answers to these questions and many others through dedicated Apps and Platforms that leverage Inria research results in the area of mobile distributed systems (from MiMove team) and data assimilation (from Inria CLIME team). The Ambiciti app is available for download on both the App and the Play stores. Starting December 2016, the Ambiciti software solutions are licensed to the Ambiciti start-up.

Monitoring exposure to noise pollution: Noise pollution is a major environmental health problem, with an estimated number of 10,000 premature deaths each year in Europe. The 2002 European environmental noise directive defines a common approach intended to avoid, prevent or reduce the harmful effects of noise. It requires the determination of exposure to environmental noise in major cities, through noise mapping. Until recently, this has been done solely through numerical simulation. Daytime, evening and nighttime averages are generally produced, without distinction between the different days of the year. Also, it is difficult to fill the gap between a noise map and the actual exposure of people where they live and stay. This motivates to monitor noise pollution where and when people are exposed. One promising direction is to make use of the noise sensors that people carry most of the time, i.e., the microphones embedded in their mobile phones.

Ambiciti (previously called SoundCity as it was initially focused on the monitoring of noise pollution) measures the actual noise levels to which individuals are exposed using such an approach, while taking into account the relatively low quality of the collected measurements. Ambiciti can then monitor noise levels throughout the day and inform users about their instantaneous, hourly and daily exposures.

In addition to the monitoring of the individual exposure to the noise pollution using mobile phones, the collective exposure may be derived from crowd-sensing. The adoption rate of mobile phones makes it possible to collect a huge amount of observational data about the noise pollution at the city scale. Recent studies have indeed highlighted the emergence of new environmental monitoring schemes leveraging the combination of mobile phones-embedded sensors and citizen participation. Ambiciti then leverages a mobile phone sensing middleware for collecting noise measurements at the urban scale, which are then assimilated toward the production of real-time pollution maps.

Monitoring exposure to air pollution: The Ambiciti app delivers information about the exposure to air pollution, providing hourly air quality maps, which are computed using numerical simulation. Depending on the user's location, the user may have access to hourly air quality maps, at street resolution, in real time and for the next two days. Currently, only Paris (France) and the Bay area enjoy such high resolution maps, but other cities are on the way to be included.

[0] http://www.choreos.eu

Mobile app features: Ambiciti is easy to use, while featuring various functions to meet different levels of user engagement (from passive monitoring to citizen scientists):

- Measuring noise level, anytime on demand or automatically during the day,
- Providing air quality indexes (according to the EU definition), in the past, present and future hours or days, together with pollution levels for nitrogen dioxide, fine particulate matter and ozone,
- Displaying statistics on exposure to pollutions, hourly, daily, during daytime and nighttime, for both noise and air.
- Displaying maps with own's noise measurements and/or hourly NO2 levels (including street-level resolution for Paris and the San Francisco Bay area)
- Promoting citizen science where communities of users may engage into the intensive measurement of noise in order to analyze a given journey or to map neighborhoods.
- Providing pollution-aware routing.
- Calibration of the smartphone for noise monitoring, automatically with Ambiciti database or manually with a sound level meter,

Privacy: The Ambiciti app has been designed with privacy in mind, which especially holds in the case of noise pollution monitoring. It is in particular important to stress that actual sound samples are never stored on the phone or uploaded to the Ambiciti servers. Only the amplitude of the sound in dB(A) is calculated, and uploaded provided the user's permission. An anonymous identifier is further created for each device for distinguishing between the data sent by different users, while no identifying information is collected. Further detail may be found from the App information page.

The Ambiciti platform is developed in collaboration with the Inria CLIME team together with The Civic Engine and the NUMTECH SMEs in the context of CityLab@Inria and Inria@SiliconValley, and with the support of the EIT Digital Env&You activity.

6.4. AppCivist-PB: A Platform for Democratic Assembly Customized for Participatory Budgeting

Participants: Valérie Issarny [contact], Cristhian Parra Trepowski, Rafael Angarita.

Participatory budgeting processes are among the most illustrative, real-life experiences of participatory democracy. Participatory Budgeting (PB) has its beginnings in the late 1980s, when some Brazilian cities started to experiment with processes of citizen participation in decisions about how to better allocate part of the city's budget. Although PB takes different forms, they can all be considered as refining the following base process: residents of a city propose spending ideas, volunteers or delegates develop those ideas into proposals, residents then vote on the proposals, and the government finally implements the winning projects. Since the 1980s, PB processes have spread around the world as a set of administrative reforms and, more recently, as a "best practice" in mainstream international development.

With AppCivist-PB, we want to enable city governments to configure the software assemblies that best match the requirements of the kind of PB campaign they want to support, while leveraging existing software services and components. However, from the overall perspective of participatory democracy, our goal is primarily to facilitate the elaboration of proposals by citizen assemblies that form according to the citizen interests. In other words, we want to support a process that emphasizes collaborative contribution making at all stages of the elaboration of proposals by diverse citizen assemblies, which are primarily created by and for citizens. The collaborative process must in particular facilitate the assembly of groups (or sub-assemblies) on the basis of commonalities among the proposals, which is essential if one wants to sustain city-scale participation and be inclusive of citizen contributions.

AppCivist-PB helps users assemble proposal making and selection workflows, using service-oriented architecture (SOA) principles. The composition principles of SOA allow for various implementations and instances of these workflows, including the possibility of integrating and linking different workflows for the same PB campaign. For example, a city might create and manage its own workflow to receive proposals and facilitate deliberation and voting by registered residents; at the same time, citizen groups (typically activists) can create their own, independent, workflows to co-create, develop, and promote proposals for the city, following their own collaboration practices. Compared to traditional SOA, AppCivist-PB distinguishes itself by enabling the assembly of software services dedicated to the support of online-facilitated participatory democracy by and for relevant citizen assemblies.

The AppCivist-PB platform is developed in collaboration with the Social Apps Labs at CITRIS at University of California Berkeley (USA) in the context of CityLab@Inria and Inria@SiliconValley, together with the support of the EIT Digital CivicBudget activity.

7. New Results

7.1. Introduction

MiMove's research activities in 2016 have focused on a set of areas directly related to the team's research topics. Hence, we have worked on QoS for Emergent Mobile Systems (§ 7.2) in relation to our research topic regarding Emergent Mobile Distributed Systems (§ 3.2). Furthermore, our effort on Ambiciti (§ 7.3) is linked to our research on Mobile Social Crowd-sensing (§ 3.4). Still in the context of Mobile Social Crowd-sensing (§ 3.4), we have developed AppCivist-PB (§ 7.4) related to our interest in social applications aiming to actively involve citizens (see § 4.1); this is further linked to our research on composition of Emergent Mobile Distributed Systems (§ 3.2). Finally, we have worked on the Fiesta-IoT ontology (§ 7.5) and on the Sarathi platform (§ 7.6), related to our research on both Large-scale Mobile Sensing & Actuation (§ 3.3) and Mobile Social Crowd-sensing (§ 3.4).

7.2. QoS for Emergent Mobile Systems

Participants: Georgios Bouloukakis, Nikolaos Georgantas, Siddhartha Dutta, Valérie Issarny.

With the emergence of Future Internet applications that connect web services, sensor-actuator networks and service feeds into open, dynamic, mobile choreographies, heterogeneity support of interaction paradigms is of critical importance. Heterogeneous interactions can be abstractly represented by client-server, publish/subscribe, tuple space and data streaming middleware connectors that are interconnected via bridging mechanisms providing interoperability among the choreography peers. We make use of the *eVolution Service Bus (VSB)* (see § 6.2) as the connector enabling interoperability among heterogeneous choreography participants [15]. VSB models interactions among peers through generic *post* and *get* operations that represent peer behavior with varying time/space coupling.

Within this context, we study end-to-end Quality of Service (QoS) properties of choreographies, where in particular we focus on the effect of middleware interactions on QoS. We consider both homogeneous and heterogeneous (via VSB) interactions. We report in the following our results in two complementary directions:

- Choreography peers deployed in mobile environments are typically characterized by intermittent connectivity and asynchronous sending/reception of data. In such environments, it is essential to guarantee acceptable levels of timeliness between sending and receiving mobile users. In order to provide QoS guarantees in different application scenarios and contexts, it is necessary to model the system performance by incorporating the intermittent connectivity. Queueing Network Models (QNMs) offer a simple modeling environment, which can be used to represent various application scenarios, and provide accurate analytical solutions for performance metrics, such as system response time. We provide an analytical solution regarding the end-to-end response time between users sending and receiving data by modeling the intermittent connectivity of mobile users with QNMs.

We utilize the publish/subscribe middleware as the underlying communication infrastructure for the mobile users. To represent the user's connections/disconnections, we model and solve analytically an ON/OFF queueing system by applying a mean value approach. Finally, we validate our model using simulations with real-world workload traces. The deviations between the performance results foreseen by the analytical model and the ones provided by the simulator are shown to be less than 5% for a variety of scenarios [16].

– Based on the QoS models and analyses outlined in the previous paragraph, we go one step further towards realistic QoS modeling and analysis of choreographies integrating heterogeneous interaction paradigms. We introduce QoS modeling patterns that correspond to each one of the interaction paradigms – client-server, publish/subscribe, tuple space and data streaming – and for different interaction styles – one way, two way synchronous, two way asynchronous. Our patterns rely on Queueing Network Models (QNMs) and represent the following characteristics of choreography peers and their middleware protocols: (i) reliable or unreliable interactions supported by the middleware and underlying transport layers; (ii) application-level (user) and middleware-level disconnections; (iii) application-level and middleware-level buffering of messages with finite capacity; (iv) limited lifetime of messages; and (v) timing of synchronous interactions. These QoS patterns enable the analysis and evaluation of the performance and success rates characterizing the modeled interactions. By combining several QoS patterns, we can further evaluate the end-to-end QoS of choreography interactions among heterogeneous peers. Based on our QoS models, we statistically analyze through simulations the effects on QoS when varying the parameters found in (i) to (v). We can also in this way evaluate the interconnection effectiveness, i.e., the degree of mapping of QoS semantics and expectations, when interconnecting heterogeneous choreography peers.

7.3. Mobile Phone Sensing Middleware for Urban Pollution Monitoring

Participants: Valerie Issarny, Cong Kinh Nguyen, Pierre-Guillaume Raverdy, Fadwa Rebhi.

Mobile Phone Sensing (MPS) is a powerful solution for massive-scale sensing at low cost. The ubiquity of phones together with the rich set of sensors that they increasingly embed make mobile phones the devices of choice to sense our environment. Further, thanks to the – even sometimes unconscious – participation of people, MPS allows for leveraging both quantitative and qualitative sensing. And, still thanks to the participation of people who are moving across space, mobile phones may conveniently act as opportunistic proxies for the sensors in their communication range, which includes the fast developing wearables.

However, despite the numerous research work since the end 2000s, MPS keeps raising key challenges among which: How to make MPS resource-efficient? How to mitigate mobile sensing heterogeneities? How to involve and leverage the crowd? How to leverage prior experiences?

Addressing the above MPS challenges primarily lies in taming the high heterogeneity not only of the computing system but also the crowd. The latter introduces a new dimension compared to traditional middleware research that has been concentrating on overcoming the heterogeneities of the computing infrastructure. In order to tackle these two dimensions together, we have been conducting a large scale empirical study in co-operation with the city of Paris (see http://tinyurl.com/soundcity-paris). Our experiment revolves around the public release of a MPS app for noise pollution monitoring that is built upon our dedicated mobile crowd-sensing middleware. Building on the Paris experiment, we systematically studied the influence of resource-efficiency and sensing accuracy on the effectiveness of the crowd participation [18]. In a complementary way, we analyzed user participation across time, so as to derive participation patterns that MPS middleware and application design may leverage.

Key take-away for MPS middleware and application design following our analysis includes:

- While contributors exhibit high heterogeneity regarding the accuracy of their sensors, they overall exhibit similar patterns. Location accuracy leads to discard about 60% of the observations and most observations are in the $[20 - 50]$ meters accuracy range. Noise sensing accuracy varies but calibration may be achieved per model rather than per device; calibration may then combine a number of techniques from comparison using a high-quality reference sensor to automated techniques leveraging assimilation and machine learning. Although our experiment is focused on noise sensing, we may expect similar results for other physical sensors. Overall, MPS allows collecting and assimilating relevant observations/measures. Still, the number of contributed measures by the MPS system needs to be high enough to overcome the low accuracy of the phone sensors.

- Although not specifically related to heterogeneity, energy efficiency is critical for the adoption of MPS. Our study confirms that energy-delay tradeoffs is a valuable approach; hence, the middleware must enable the buffering of the observations while the frequency of the transfers must be tuned by the application. Still, we notice that 30% of the observations reach the server after 2 hours even when observations are not buffered and are sent every 5mns, which indicates long periods of disconnection. Hence, if the timeliness of the observation is critical, then participatory sensing is most likely the approach to follow to ensure that the user is conscious about the sensing and activates appropriate network connection.
- The heterogeneity of the contributing crowd is obvious. However, it turns out to be an asset rather than a shortcoming of MPS. Indeed, the crowd overall exhibits similar contribution patterns across time. However, in the detail, each individual has different contribution patterns. This allows for the collection of complementary contributions over the whole day.
- The users appear to be still most of the time, while the user's activity cannot be qualified for 20% of the observations. This should be accounted for in the design of mobility-dependent MPS.
- One design issue that arises for MPS is whether to promote participatory or opportunistic sensing. It is our belief that a system (and thus supporting app) must support both. This enables to collect as many observations as possible from a large diversity of people, while participatory sensing guarantees contributions of higher quality.

7.4. Computer-mediated Social Communication Interoperability

Participants: Rafael Angarita, Nikolaos Georgantas, Valerie Issarny, Cristhian Parra Trepowski, Christelle Rohaut.

People increasingly rely on computer-mediated communication for their social interactions. This is a direct consequence of the global reach of the Internet combined with the massive adoption of social media and mobile technologies that make it easy for people to view, create and share information within their communities almost anywhere, anytime. The success of social media has further led – and is still leading – to the introduction of a large diversity of social communication services (e.g., Skype, Facebook, Google Plus, Telegram, Instagram, WhatsApp, Twitter, Slack, ...). These services differ according to the types of communities and interactions they primarily aim at supporting. However, existing services are not orthogonal and users ultimately adopt one service rather than another based on their personal experience. As a result, users who share similar interests from a social perspective may not be able to interact in a computer-mediated social sphere because they adopt different technologies. This is particularly exacerbated by the fact that the latest social media are proprietary services that offer an increasingly rich set of functionalities, and the function of one service does not easily translate -both socially and technically- into the function of another. As an illustration, compare the early and primitive social media that is the Email with the richer social network technology. Protocols associated with the former are rather simple and email communication between any two individuals is now trivial, independent of the mail servers used at both ends. On the other hand, protocols associated with today's social networks involve complex interaction processes, which prevent communication across social networks.

The above issue is no different than the long-standing issue of interoperability in distributed computing systems, which require to mediate (or translate) the protocols run by the interacting parties for them to be able to exchange meaningful messages and coordinate. And, while interoperability in the early days of distributed systems was essentially relying on the definition of standards, the increasing complexity and diversity of networked systems has led to the introduction of various interoperability solutions, among which the (Enterprise) Service Bus paradigm.

In the above context, we have specifically introduced the "*social communication bus*" paradigm so as to allow interoperability across computer-mediated social communication protocols. Our work is motivated by our research effort within the AppCivist project. AppCivist provides a software platform for participatory democracy that leverages the reach of the Internet and the powers of computation to enhance the experience and efficacy of civic participation. Its first instance, AppCivist-PB, targets participatory budgeting, an exemplary process of participatory democracy that let citizens prepare and select projects to be implemented with public funds by their cities [17]. For city-wide engagement, AppCivist-PB must enable citizens to participate with the Internet-based communication services they are the most confortable with. The need for interoperability in this context is indeed paramount since the idea is to include people in the participatory processes without leaving anyone behind. This has led us to revisit the service bus paradigm for the sake of social communication across communities, so as to gather together the many communities of our cities.

Our contributions span:

- *Social communication paradigm:* Based on the survey of the various forms of computer-mediated social communication supported by today's software services and tools, we have derived how the approaches to middleware interoperability may apply to social communication interoperability.
- *Social Communication Bus architecture:* We leverage the VSB bus (see § 6.2) that supports interoperability across interaction paradigms as opposed to interoperability across heterogeneous middleware protocols implementing the same paradigm. The proposed bus architecture features the traditional concepts of bus protocols and binding components, but those are customized for the sake of social interaction whose coupling differs along the social and presence dimensions.
- *Social Communication Bus instance for participatory democracy:* We have refined our bus architecture, introducing the Social-MQ implementation that leverages the RabbitMQ message broker. The resulting implementation has been integrated within the AppCivist-PB platform for evaluation.

In order to inform the further study of the "Social Communication Bus" paradigm, we have analyzed existing practices and supporting technologies promoting citizen collaboration. In relation with our work on the AppCivist-PB platform, our study has concentrated on Participatory Budgeting (PB) campaigns, with a special focus on US-related initiatives, as a mean to understand the current and future design space of ICT for participatory democracy. We then derived new design opportunities for ICT to facilitate citizen collaboration in the PB process, and by extension, to reflect on how these technologies could better foster deliberative decision-making at a scale that is both small and large.

This research is carried out in collaboration with the Social Apps Lab at CITRIS at UC Berkeley in the context of CityLab@Inria and Inria@SiliconValley.

7.5. FIESTA-IoT Ontology: Semantic Model for Federation & Interoperability among Platforms

Participants: Rachit Agarwal, Valérie Issarny, Nikolaos Georgantas.

Plethora of heterogeneous data is being generated and made available by diverse platforms. Such platforms can be those that are formed by the use of mobile application that act as interface between sensing devices and storage or between users and storage. The diversity and openness in the data generated isolate platforms and lead to interoperability issues between platforms, where much work has to be done in order to ensure compatibility. One has to understand the other's format, parse different data formats, and create the mapping between different data formats. One method to accomplish this interoperability is by attaching semantics to this data. Semantics provides meaning to the data and helps in (a) achieving common understanding and (b) performing analysis and reasoning. Many IoT-related semantic models [0] propose interoperability but have many issues like: observation graph is missing, are highly domain specific, and do not follow best practices. In order to address the above, we focused our research on: the identification of a unified semantic model that addresses the above, creation of a prototype application, and identification of guidelines for storing semantic data [13]. We report our following key results:

[0] http://sensormeasurement.appspot.com/?p=ontologies

- *State of art survey of semantic models that are available in literature in the domain of the Internet of Things*: This survey gave us required knowledge needed for the semantic model from which concepts can be reused to create a unified ontology. This helps the semantic community by not overloading the domain with concepts similar to already existing concepts, and allows us to reuse concepts as much as possible. We identified that recent trends show more and more use of the SSN [35] and oneM2M [67] ontologies. However, these models are currently far from being able to address observation-related issues and lack domain taxonomy.
- *Unified semantic model for enabling interoperability and federation of testbeds*: Based on the analysis of the concepts from various onotologies identified, we unify specific concepts from these identified ontologies into one ontology. These ontologies being: SSN, oneM2M, IoT-lite [27], WGS84 [0], DUL [0], TIME [0] and M3-lite taxonomy (created as a part of this research). Such unification gives our ontology the power to define meta data about the sensor that is producing the observation and the observation itself. The federation is achieved by the use of the taxonomy that each platform should follow.
- *Best practices to publish data based on the unified model*: In order to enable full interoperability, federation and usage of data, it is essential that best practices are followed while storing the data based on the unified model. We identify various best practices which form our recommendations to the platform owners towards annotating the data with respect to the ontology. This is supported by a reference annotator that also acts as a guide for developers to publish data.

These above-mentioned results are currently applied in the frame of the EU funded H2020 FIESTA-IoT project (see § 8.2.1.2).

7.6. Sarathi: A Platform for Personalized Mobility Service for Urban Travellers

Participants: Rachit Agarwal, Garvita Bajaj, Georgios Bouloukakis, Valérie Issarny, Nikolaos Georgantas.

Thanks to the increased abundance of mobile phones, the recent field of mobile participatory sensing could be leveraged towards providing a more fine-grained and up-to-date view of a city's transportation system. Thus, in order to address problems like dynamicity (unexpected faults, stoppages, etc.) and unexpected load (number of people using the transportation), etc., in different societal contexts of France and India, we aimed to produce a middleware platform called "*Sarathi*" that is enriched with personalized mobility services for urban travelers and is evaluated via real-life demonstrators. Towards this, the key results include:

- *Identification of System Architecture* [14]: We first identify requirements for our system that would satisfy the objectives. The identified requirements are then mapped to specific components that would carry out specific tasks. A client-server system architecture is then created by connecting the identified components. Some components that we identified are: UI component that would run at the client side, recommendation system and knowledgebase component that would run at the server, and a communication component that would ensure communication of the client with the server. To realise these components, we also identify tools and techniques that would ensure best runtime performance.
- *Modeling Passenger convenience in Metro transit* [20]: This effort builds upon existing research in the area, studied during our joint survey of related work, and applies the work to the context of the Paris and New Delhi metro system. This work captures 'personalized' experience of passengers during a multi-leg journey and models the convenience for commuters. A leg in a journey is defined as a segment of a journey traveled on a metro line. The work proposes a mathematical model for commuter convenience and validates it using data collected from metro commuters. The convenience model uses 3 convenience measures namely *seat availability*, *wait time* and *comfort*. The work

[0] https://www.w3.org/2003/01/geo/
[0] http://www.ontologydesignpatterns.org/ont/dul/DUL.owl
[0] https://www.w3.org/TR/owl-time/

also aims to identify the best mobile interaction paradigm for enabling timely data collection and dissemination and outlines a middleware architecture to achieve this (aiming at acceptable response times for mobile apps).

- *Mobile Application*: An Android application called *MetroCognition* for gathering commuters convenience rating during their metro transit based on the three above described measures has been developed, deployed and made available on Google Play Store [0] for beta testing.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. Inria Support

8.1.1.1. Inria IPL CityLab@Inria

Participants: Valérie Issarny [correspondent], Fadwa Rebhi.

- **Name:** CityLab@Inria – *Overcoming the Smart City Challenge – Toward Environmental and Social Sustainability*
- **Related activities:** § 7.3 and § 6.3
- **Period:** [January 2014 – December 2018]
- **Inria teams:** CLIME, DICE, FUN, MIMOVE, MYRIADS, SMIS, URBANET, WILLOW
- **URL:** http://citylab.inria.fr

The Inria Project Lab (IPL) CityLab@Inria studies ICT solutions toward smart cities that promote both social and environmental sustainability. A strong emphasis of the Lab is on the undertaking of a multi-disciplinary research program through the integration of relevant scientific and technology studies, from sensing up to analytics and advanced applications, so as to actually enact the foreseen smart city Systems of Systems. Obviously, running experiments is a central concern of the Lab, so that we are able to confront proposed approaches to actual settings.

8.1.1.2. Inria ADT CityLab Platform

Participants: Valérie Issarny [correspondent], Fadwa Rebhi.

- **Name:** CityLab Platform – *A Platform for Smarter Cities Promoting Social and Environmental Sustainability*
- **Related activities:** § 7.3 and § 6.3
- **Period:** [November 2014 – October 2016]
- **Partners:** Inria MiMove, Inria CLIME.

The CityLab Platform ADT is part of the CityLab Inria Project Lab. The ADT is more specifically focused on the development of a middleware platform supporting mobile crowd-Xing for environmental pollution monitoring through user-led observations.

[0] https://play.google.com/apps/testing/edu.sarathi.metroCognition

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. H2020 ICT CHOReVOLUTION

Participants: Nikolaos Georgantas [correspondent], Valérie Issarny [correspondent].

Name: CHOReVOLUTION – *Automated Synthesis of Dynamic and Secured Choreographies for the Future Internet*

URL: http://www.chorevolution.eu

Type: Research & Innovation Action (ICT)

Topic: Tools and Methods for Software Development

Related activities: § 7.2 and § 6.2

Period: [January 2015 - December 2017]

Partners: CEFRIEL (Italy), Inria MiMove, OW2 Consortium (France), Thales Communications S.A. (France) [**coordinator**], Università degli Studi dell'Aquila (Italy) [**scientific leader**], Softeco Sismat SrL (Italy), Tirasa (Italy), Viktoria Swedish ICT (Sweden).

The Future Internet (FI) represents an age of unprecedented opportunities for social, economic, and business growth thanks to the global scale connectivity of the virtual as well as of the physical world. This indeed opens up a wealth of innovative and revolutionary real-life scenarios, as for instance illustrated by the smarter cities perspectives where envisioned scenarios significantly ease daily human activities and give support for the growth of new markets and employment opportunities. However, leveraging the FI for the development of innovative software applications remain a challenging task even though major enablers are readily available by ways of service-oriented and cloud computing. It is in particular our vision that enabling the choreography of FI services shall play a significant role in the provisioning of innovative applications. However, existing choreography-based service composition approaches are rather static and are poorly suited to the need of the FI that is a highly dynamic networking environment, further bringing together highly heterogeneous services ranging from Thing- to Business-based services that span different security domains. As a result, the technology is not mature enough for market take-up. CHOReVOLUTION elevates the Readiness Level of existing choreography technologies in order to drop the dynamism and cross-organization security barriers via the automated synthesis of dynamic and secured choreographies in the FI. To meet its objectives, CHOReVOLUTION undertakes both research and innovation tasks. The former concentrates on choreography modelling, synthesis, adaptation, service bus, security, and cloud; the latter focus on industrial validation, development support and integration platform, and the establishment of a CHOReVOLUTION community and market take- up. Last but not least CHOReVOLUTION outcomes are assessed by experimenting with new applications in the field of Intelligent Transportation Systems.

8.2.1.2. H2020 ICT FIESTA-IoT

Participants: Valérie Issarny [correspondent], Nikolaos Georgantas [correspondent].

Name: FIESTA-IoT – *Federated Interoperable Semantic IoT/cloud Testbeds and Applications*

URL: http://fiesta-iot.eu

Type: Research & Innovation Action (ICT)

Topic: FIRE+ (Future Internet Research & Experimentation)

Related activities: § 7.5

Period: [February 2015 - January 2018]

Partners: Fraunhofer FOKUS (Germany) [**coordinator**], INSIGHT @ National University of Galway (Ireland) [**co-coordinator**], University of Southampton IT Innovation Centre (UK), Inria MiMove, University of Surrey (UK), Unparallel Innovation Lda (Portugal), Easy Global Market (France), NEC Europe Ltd (UK), University of Cantabria (Spain), Com4innov (France), Athens Information Technology (Greece), SOCIEDAD PARA EL DESARROLLO REGIONAL DE CANTABRIA (Spain), Ayuntamiento de Santander (Spain), Korea Electronics Technology Institute (Korea).

Despite the proliferation of IoT and smart cities testbeds, there is still no easy way to conduct large scale experiments that leverage data and resources from multiple geographically and administratively distributed IoT platforms. Recent advances in IoT semantic interoperability provide a sound basis for implementing novel cloud-based infrastructures that could allow testbed-agnostic access to IoT data and resources. FIESTA will open new horizons in IoT experimentation at a global scale, based on the interconnection and interoperability of diverse IoT testbeds. FIESTA will produce a first-of-a-kind blueprint experimental infrastructure (tools, techniques and best practices) enabling testbed operators to interconnect their facilities in an interoperable way, while at the same time facilitating researchers in deploying integrated experiments, which seamlessly transcend the boundaries of multiple IoT platforms. FIESTA will be validated and evaluated based on the interconnection of four testbeds (in Spain, UK, France and Korea), as well as based on the execution of novel experiments in the areas of mobile crowd-sensing, IoT applications portability, and dynamic intelligent discovery of IoT resources. In order to achieve global outreach and maximum impact, FIESTA will integrate an additional testbed and experiments from Korea, while it will also collaborate with IoT experts from USA. The participation of a Korean partner (based its own funding) will maximize FIESTA's value for EC money. Moreover, the project will take advantage of open calls processes towards attracting third-parties that will engage in the integration of their platforms within FIESTA or in the conduction of added-value experiments. As part of its sustainability strategy, FIESTA will establish a global market confidence programme for IoT interoperability, which will enable innovative platform providers and solution integrators to ensure/certify the openness and interoperability of their developments.

8.2.2. Collaborations in European Programs, Except FP7 & H2020

8.2.2.1. EIT Digital Env&You

Participant: Valérie Issarny [correspondent].

Name: Env&You – *Personalizing environmental science for your home, your neighborhood and your life*

URL: http://ambiciti.io

Related activities: § 7.3 and § 6.3

Period: [January 2016 - December 2016]

Partners: Ambientic (F), CapDigital (F), Forum Virium Helsinki (FI), Inria CLIME, Inria MIMOVE [**coordinator**], NumTech (F), TheCivicEngine (USA).

There is a clear, and probably increasing, desire from the citizens to better know their individual exposure to pollution. Partial solutions exist to the exposure data problem but each focuses on one or another domain of information – crowdsourcing exposure, translating government open data to usable consumer information, harnessing social media information, harnessing biometrics – what is unique about Env&You is that we assimilate a multi-dimensional picture of exposure and provide the integrated information to citizen, government, and business use (spanning: B2G, B2B and B2C business cases).

8.2.2.2. EIT Digital CivicBudget

Participants: Valérie Issarny [correspondent], Nikolaos Georgantas [correspondent].

Name: CivicBudget – *Software platform supporting Internet-based participatory budgeting campaigns*

Related activities: § 7.4 and § 6.4

Period: [January 2016 - December 2016]

Partners: CITRIS (USA), Inria MIMOVE, MissionsPubliques (F) [**coordinator**], Nexus (DE), and TU Berlin (DE).

Many cities in Europe and the U.S.A, and around the world, commit a percentage of their annual budget (often 5%) to implement citizen-proposed projects through a process called Participatory Budgeting (PB). However, supporting urban-scale participatory budgeting campaigns is greatly challenged as it still principally relies on physical meetings. CivicBudget addresses this challenge by leveraging latest ICT so as to promote urban-scale inclusion. CivicBudget fosters a new and inclusive urban public sphere of citizenship. It is especially designed for community groups and activists who want to participate in the PB process. City governments will also be able to promote its use. CivicBudget will facilitate the mobilization of residents both to promote their proposals and to monitor their progress through the PB process to implementation.

8.3. International Initiatives

8.3.1. Inria International Labs

Valérie Issarny acts as scientific manager of the Inria@Silicon Valley program (https://project.inria.fr/inria-siliconvalley/) since summer 2013; she is visiting scholar at the EECS Department of University of California, Berkeley, and hosted by CITRIS.

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

8.3.2.1. Inria DRI/DST-CEFIPRA Associate Team: SARATHI

Participant: Rachit Agarwal [correspondent].

Name: SARATHI – *Personalized Mobility Services for Urban Travelers*

Instrument: Inria DRI/DST-CEFIPRA Associate Team

Related activities: § 7.6

Period: [January 2014 - December 2016]

Partners: Indraprastha Institute of Information Technology (IIIT) Delhi (India), Inria MiMove.

Website: http://sarathi.gitlab.io/web/

The focus of the *Sarathi* project is on creating a personalized mobility service platform for urban travelers. The proposed work would require work on large scale mobile participatory sensing, urban transportation, location-aware services, machine learning, and software engineering. The individual strength of MiMove and IIIT provide complementary technical benefits for the project. MiMove leverages its work on large scale mobile participatory sensing (so far focused on EU-based transit contexts) addressing challenges brought to the fore by dynamic large scale systems in India; IIIT will build up on their previous work on mobile based system to provide route information and work on learning and mining techniques for inferring events of interest in transport systems.

Besides the complementary technical benefits, the collaboration will also help the project in evaluating the proposed solution in context of both developing and developed countries with different societal structure and preferences. Since personalized services are an integral part of the solution, the variety in social structures of India and France will help in developing solutions that are valid across continents. A deployment of the proposed solution in India will also test scalability and robustness of the solution in resource-constrained environments (e.g. intermittent network connectivity, low bandwidth) and will help in developing solutions that can be deployed in different working environments. Similarly, France (with already an advanced transit system) offers opportunities in verifying the requirements of a successful sustainable transport system.

8.3.2.2. Inria/Brazil Associate Team: ACHOR

Participant: Nikolaos Georgantas [correspondent].

Name: ACHOR – *Adaptive enactment of service choreographies*

Instrument: Inria/Brazil Associate Team

Related activities: § 7.2 and § 6.2

Period: [January 2016 - December 2018]

Partners: Universidade Federal de Goiás (UFG), Brazil, Inria MiMove.

Website: http://www.inf.ufg.br/projects/achor

Service choreographies are distributed compositions of services (e.g., Web services) that coordinate their execution and interactions without centralized control. Due to this decentralized coordination and the ability to compose third-party services, choreographies have shown great potential as an approach to automate the construction of large-scale, on-demand, distributed applications. Technologies to enable this approach are reaching maturity level, such as modeling languages for choreography specification and engines that operate the deployment of services and enactment of choreographies at Future Internet scales. Nevertheless, a number of problems remain open on the way to fully realize the approach, among them: (i) Deployment of multiple choreographies on top of a collection of shared services (considering service sharing as an effective way to increase the utilization of resources); (ii) Dynamic adaptation of functional and non-functional properties due to runtime changes in the environment and user requirements (adapting the set of services and/or the resources used to run the services in order to add/remove/change functions and maintain QoS properties, respectively); and (iii) Seamless and dynamic integration of mobile services (e.g., smartphone apps, sensors and actuators on handhelds and wearables) and cloud- based services (including the need to consider: mobility of both devices and services, resource constraints of mobile devices, temporary disconnection, interoperability between different interaction paradigms (message-passing, event-based, data-sharing) at the middleware layer, and effect of these paradigms on end-to-end QoS).

The overall goal of the project is to design an architecture for adaptive middleware to support service choreographies in large-scale scenarios that involve dynamicity and diversity in terms of application requirements, service interaction protocols, and the use of shared local, mobile and cloud resources.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

8.4.1.1. Internships

Garvita Bajaj (from May 2016 until Sept 2016)

PhD internship funded by Associate Team Sarathi and H2020 FIESTA-IoT project.

Subject: *Extending current FIESTA-IoT Ontology*

Institution: Indraprastha Institute of Information Technology (IIIT) Delhi (India)

8.4.2. Visits to International Teams

8.4.2.1. Research Stays Abroad

Valérie Issarny is visiting scholar at the EECS Department at UC Berkeley; she is hosted by CITRIS in the context of which she carries out collaborative research in the area of smart cities and acts as scientific coordinator of the Inria@SiliconValley program.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Valérie Issarny is co-chair of the BIS'16 workshop, which is the yearly workshop organized by Inria@SiliconValley to present the state of scientific collaborations and to work on proposals for future ambitious joint projects.

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

- Valérie Issarny is PC co-chair of ICSE-SEIS'2018 - The Software Engineering in Society Track of the ICSE'18 conference.

9.1.2.2. Member of the Conference Program Committees

- Valérie Issarny is PC member of the following International Conferences: COORDINATION'2016, FASE'2016 & 2017 & 2018, ICDCS'2016, ICSE-SEIS'2016 & 2017, ICSE'2018, InterIoT'2016, ISEC'2017, Middleware'2016, MOBILESoft'2017, SEAMS'17, EUROSYS'2017.
- Nikolaos Georgantas is PC member of the following international conferences: SOSE'16&'17, WETICE'16&'17, SAC'16 &'17, AmI'17, ICSE'17 Poster Track, ANT'16.
- Nikolaos Georgantas is PC member of the following international workshops: MRT'16, SERENE'16, ARM'16, IoT-ASAP'17.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Nikolaos Georgantas is associate editor of the International Journal of Ambient Computing and Intelligence (IJACI);
- Valérie Issarny is associate editor of the Springer JISA Journal of Internet Services and Applications.

9.1.4. Invited Talks

- Valérie Issarny, *Monitoring Urban Pollution using Mobile Crowd Sensing: The SoundCity Use Case*, Keynote at - UIC - The 13th International Conference on Ubiquitous Intelligence and Computing, July 19, 2016; Toulouse, France.
- Valérie Issarny, *Revisiting Service-oriented Architecture for the IoT: A Middleware Perspective*, Keynote at - ICSOC - The 14th International Conference on Service-Oriented Computing, October 12, 2016; Banff, Alberta, Canada.

9.1.5. Scientific Expertise

- Valérie Issarny is elected member of the *Commission d'Evaluation Inria*;
- Nikolaos Georgantas is member of the Inria PhD scholarship, Inria postdoc scholarship and Inria professor leave *(Délégation)* scholarship selection committees at Inria Paris.

9.1.6. Research Administration

- Valérie Issarny is scientific coordinator of Inria@Silicon Valley and CityLab@Inria;
- Nikolaos Georgantas is member of the PhD monitoring committee at Inria Paris.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

E-learning

Valérie Issarny, MOOC *Villes intelligentes : défis technologiques et sociétaux*
https://www.fun-mooc.fr/courses/inria/41009/session01/about

Class teaching

Master: Rachit Agarwal, “Urban data democratization and its application to access urban concepts" as part of “Gestion de données ambiantes et Internet des objets", 9 hours (équivalent TD), niveau M2, University of Versailles Saint-Quentin en Yvelines, France.

9.2.2. Supervision

In 2016, the following PhD thesis was successfully defended:

Emil Andriescu, *Dynamic Data Adaptation for the Synthesis and Deployment of Protocol Mediators*, defended in February 2016, UPMC-EDITE, CIFRE Inria-Ambientic, advised by Valérie Issarny and Roberto Speicys-Cardoso (Ambientic).

Additionally, the following PhD theses are currently in progress at the MiMove team:

Georgios Bouloukakis, *Enabling Emergent Mobile Systems in the IoT: from Middleware-layer Communication Interoperability to Associated QoS Analysis*, started October 2013, UPMC-EDITE, advised by Nikolaos Georgantas and Valérie Issarny.

Radha Pallavali, *Dynamic adaptation of middleware-layer protocols for emergent mobile systems*, started November 2016, UPMC-EDITE, advised by Nikolaos Georgantas and Valérie Issarny.

Also, Valérie Issarny is co-advising with Ansgar Radermacher from CEA-LISE, the PhD thesis of Amel Belaggoun on *Adaptabilité et reconfiguration des systèmes temps-réel embarquées*; this is a PhD from UPMC-EDITE with the research being undertaken at CEA-LISE.

9.2.3. Juries

Nikolaos Georgantas was examiner for the PhD thesis of Fatiha Houacine, *Architecture Orientée Service pour le Mobile Cloud Computing*, defended in November 2016, Conservatoire National des Arts et Métiers, France, advised by Ass. Prof. Samia Bouzefrane.

10. Bibliography

Major publications by the team in recent years

[1] S. Ben Mokhtar, D. Preuveneers, N. Georgantas, V. Issarny, Y. Berbers.*EASY: Efficient SemAntic Service DiscoverY in Pervasive Computing Environments with QoS and Context Support*, in "Journal of Systems and Software, Special Issue on Web Services Modelling and Testing", 2008, vol. 81, n^{o} 5, p. 785-808.

[2] 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.

[3] B. Billet, V. Issarny.*Dioptase: a distributed data streaming middleware for the future web of things*, in "Journal of Internet Services and Applications", 2014, vol. 5, n^{o} 1, 28 [*DOI :* 10.1186/s13174-014-0013-1], https://hal.inria.fr/hal-01081738.

[4] 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.

[5] 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.

[6] 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.

[7] V. Issarny, M. Caporuscio, N. Georgantas.*A Perspective on the Future of Middleware-based Software Engineering*, in "FOSE '07: 2007 Future of Software Engineering", Washington, DC, USA, IEEE Computer Society, 2007, p. 244–258, http://dx.doi.org/10.1109/FOSE.2007.2.

[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.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[9] E.-M. Andriescu.*Dynamic Data Adaptation for the Synthesis and Deployment of Protocol Mediators*, Université Pierre et Marie Curie - Paris VI, February 2016, https://tel.archives-ouvertes.fr/tel-01402011.

Articles in International Peer-Reviewed Journal

[10] N. Ben Mabrouk, N. Georgantas, V. Issarny.*Multi-Objective Service Composition in Ubiquitous Environments with Service Dependencies*, in "International Journal of Services Computing (IJSC)", April 2016, vol. 4, https://hal.inria.fr/hal-01403750.

Invited Conferences

[11] A. BELAGGOUN, V. Issarny.*Towards adaptive autosar : a system-level approach*, in "FISITA 2016 World Automotive Congress", Busan, South Korea, September 2016, https://hal.inria.fr/hal-01416879.

[12] V. Issarny, G. Bouloukakis, N. Georgantas, B. Billet.*Revisiting Service-oriented Architecture for the IoT: A Middleware Perspective*, in "14th International Conference on Service Oriented Computing (ICSOC)", Banff, Alberta, Canada, October 2016, https://hal.inria.fr/hal-01358399.

International Conferences with Proceedings

[13] R. AGARWAL, D. Gomez Fernandez, T. Elsaleh, A. Gyrard, J. Lanza, L. Sanchez, N. Georgantas, V. Issarny.*Unified IoT Ontology to Enable Interoperability and Federation of Testbeds*, in "3rd IEEE World Forum on Internet of Things", Reston, United States, December 2016, https://hal.inria.fr/hal-01386917.

[14] G. Bajaj, R. AGARWAL, G. Bouloukakis, P. Singh, N. Georgantas, V. Issarny.*Towards Building Real-Time, Convenient Route Recommendation System for Public Transit*, in "IEEE International Smart Cities Conference", Trento, Italy, September 2016, https://hal.inria.fr/hal-01351068.

[15] G. Bouloukakis, N. Georgantas, S. Dutta, V. Issarny.*Integration of Heterogeneous Services and Things into Choreographies*, in "14th International Conference on Service Oriented Computing (ICSOC)", Banff, Alberta, Canada, October 2016, https://hal.inria.fr/hal-01358043.

[16] G. Bouloukakis, N. Georgantas, A. Kattepur, V. Issarny.*Timeliness Evaluation of Intermittent Mobile Connectivity over Pub/Sub Systems*, in "8th ACM/SPEC International Conference on Performance Engineering (ICPE)", L'Aquila, Italy, April 2017, https://hal.inria.fr/hal-01415893.

[17] J. Holston, V. Issarny, C. Parra.*Engineering Software Assemblies for Participatory Democracy: The Participatory Budgeting Use Case*, in "Software Engineering in Society at ICSE", Austin, TX, United States, May 2016, https://hal.inria.fr/hal-01261012.

[18] V. ISSARNY, V. MALLET, K. NGUYEN, P.-G. RAVERDY, F. REBHI, R. VENTURA.*Dos and Don'ts in Mobile Phone Sensing Middleware: Learning from a Large-Scale Experiment*, in "ACM/IFIP/USENIX Middleware 2016", Trento, Italy, Proceedings of the 2016 International Middleware Conference, December 2016 [*DOI :* 10.1145/2988336.2988353], https://hal.inria.fr/hal-01366610.

References in notes

[19] L. ALDRED, W. M. P. VAN DER AALST, M. DUMAS, A. H. M. TER HOFSTEDE.*Dimensions of Coupling in Middleware*, in "Concurrency and Computation: Practice and Experience", 2009, vol. 21, n^o 18, p. 2233–2269, http://eprints.qut.edu.au/40797/.

[20] G. BAJAJ, G. BOULOUKAKIS, A. PATHAK, P. SINGH, N. GEORGANTAS, V. ISSARNY.*Toward Enabling Convenient Urban Transit through Mobile Crowdsensing*, in "18th IEEE International Conference on Intelligent Transportation Systems", Gran Canaria, Spain, September 2015, https://hal.inria.fr/hal-01204827.

[21] B. BALAJI, J. XU, A. NWOKAFOR, R. GUPTA, Y. AGARWAL.*Sentinel: Occupancy Based HVAC Actuation Using Existing WiFi Infrastructure Within Commercial Buildings*, in "Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems", New York, NY, USA, SenSys '13, ACM, 2013, p. 17:1–17:14, http://doi.acm.org/10.1145/2517351.2517370.

[22] M. V. BARBERA, A. EPASTO, A. MEI, V. C. PERTA, J. STEFA.*Signals from the Crowd: Uncovering Social Relationships Through Smartphone Probes*, in "Proceedings of the 2013 Conference on Internet Measurement Conference", New York, NY, USA, IMC '13, ACM, 2013, p. 265–276, http://doi.acm.org/10.1145/2504730.2504742.

[23] L. BARESI, C. GHEZZI.*The Disappearing Boundary Between Development-time and Run-time*, in "Proceedings of the FSE/SDP Workshop on Future of Software Engineering Research", New York, NY, USA, FoSER '10, ACM, 2010, p. 17–22, http://doi.acm.org/10.1145/1882362.1882367.

[24] A. BEACH, M. GARTRELL, S. AKKALA, J. ELSTON, J. KELLEY, K. NISHIMOTO, B. RAY, S. RAZGULIN, K. SUNDARESAN, B. SURENDAR, M. TERADA, R. HAN.*Whozthat? Evolving an ecosystem for context-aware mobile social networks*, in "IEEE Xplore", 2008, vol. 22, n^o 4, p. 50–55.

[25] N. BENCOMO, A. BELAGGOUN, V. ISSARNY.*Dynamic Decision Networks for Decision-making in Self-adaptive Systems: A Case Study*, in "Proceedings of the 8th International Symposium on Software Engineering for Adaptive and Self-Managing Systems", Piscataway, NJ, USA, SEAMS '13, IEEE Press, 2013, p. 113–122, http://dl.acm.org/citation.cfm?id=2487336.2487355.

[26] A. BENNACEUR.*Synthèse dynamique de médiateurs dans les environnements ubiquitaires*, Université Pierre et Marie Curie - Paris VI, July 2013, http://hal.inria.fr/tel-00849402.

[27] M. BERMUDEZ-EDO, T. ELSALEH, P. BARNAGHI, K. TAYLOR.*IoT-Lite: A Lightweight Semantic Model for the Internet of Things*, in "13th IEEE International Conference on Ubiquitous Intelligence and Computing (UIC)", Toulouse, July 2016.

[28] 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.

[29] B. BOEHM.*A View of 20th and 21st Century Software Engineering*, in "Proceedings of the 28th International Conference on Software Engineering", New York, NY, USA, ICSE '06, ACM, 2006, p. 12–29, http://doi.acm.org/10.1145/1134285.1134288.

[30] Z. BOUZID, M. G. POTOP-BUTUCARU, S. TIXEUIL.*Optimal Byzantine-resilient Convergence in Unidimensional Robot Networks*, in "Theor. Comput. Sci.", July 2010, vol. 411, n^o 34-36, p. 3154–3168, http://dx.doi.org/10.1016/j.tcs.2010.05.006.

[31] R. CALINESCU, C. GHEZZI, M. KWIATKOWSKA, R. MIRANDOLA.*Self-adaptive Software Needs Quantitative Verification at Runtime*, in "Commun. ACM", September 2012, vol. 55, n^o 9, p. 69–77, http://doi.acm.org/10.1145/2330667.2330686.

[32] L. CAPRA, P. CHÂTEL, A. PATHAK, R. SPEICYS CARDOSO.*TravelDashboard – a Framework for the Delivery of Personalized Mobility Services to Urban Travellers*, in "ERCIM News", April 2013, vol. 2013, n^o 93, http://hal.inria.fr/hal-00939031.

[33] D. CHRISTIN, C. ROSSKOPF, M. HOLLICK.*uSafe: A privacy-aware and participative mobile application for citizen safety in urban environments*, in "Pervasive and Mobile Computing", 2013, vol. 9, n^o 5, p. 695-707, http://dblp.uni-trier.de/db/journals/percom/percom9.html#ChristinRH13.

[34] V. CICIRELLO, M. PEYSAKHOV, G. ANDERSON, G. NAIK, K. TSANG, W. REGLI, M. KAM.*Designing dependable agent systems for mobile wireless networks*, in "Intelligent Systems, IEEE", 2004, vol. 19, n^o 5, p. 39–45.

[35] M. COMPTON, P. BARNAGHI, L. BERMUDEZ, R. GARCÍA-CASTRO, O. CORCHO, S. COX, J. GRAYBEAL, M. HAUSWIRTH, C. HENSON, A. HERZOG, V. HUANG, K. JANOWICZ, W. D. KELSEY, D. LE PHUOC, L. LEFORT, M. LEGGIERI, H. NEUHAUS, A. NIKOLOV, K. PAGE, A. PASSANT, A. SHETH, K. TAYLOR.*The SSN ontology of the W3C semantic sensor network incubator group*, in "Web Semantics: Science, Services and Agents on the World Wide Web", Desmber 2012, vol. 17, p. 25–32 [*DOI :* 10.1016/J.WEBSEM.2012.05.003], http://linkinghub.elsevier.com/retrieve/pii/S1570826812000571.

[36] E. CUERVO, A. BALASUBRAMANIAN, D.-K. CHO, A. WOLMAN, S. SAROIU, R. CHANDRA, P. BAHL.*MAUI: Making Smartphones Last Longer with Code Offload*, in "Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services", New York, NY, USA, MobiSys '10, ACM, 2010, p. 49–62, http://doi.acm.org/10.1145/1814433.1814441.

[37] M. DEMIRBAS, M. BAYIR, C. AKCORA, Y. YILMAZ, H. FERHATOSMANOGLU.*Crowd-sourced sensing and collaboration using Twitter*, in "World of Wireless Mobile and Multimedia Networks (WoWMoM), 2010 IEEE International Symposium on a", June 2010, http://dx.doi.org/10.1109/WOWMOM.2010.5534910.

[38] F. DI GIANDOMENICO, A. BERTOLINO, A. CALABRÒ, N. NOSTRO.*An approach to adaptive dependability assessment in dynamic and evolving connected systems*, in "International Journal of Adaptive, Resilient and Autonomic Systems (IJARAS)", March 2013, vol. Volume 4, n^o 1, p. 1-25 [*DOI :* 10.4018/JARAS.2013010101], http://www.igi-global.com/article/approach-adaptive-dependability-assessment-dynamic/75547.

[39] M. DIXIT, A. CASIMIRO, P. LOLLINI, A. BONDAVALLI, P. VERISSIMO.*Adaptare: Supporting Automatic and Dependable Adaptation in Dynamic Environments*, in "ACM Trans. Auton. Adapt. Syst.", July 2012, vol. 7, n^o 2, p. 18:1–18:25, http://doi.acm.org/10.1145/2240166.2240168.

[40] I. EPIFANI, C. GHEZZI, G. TAMBURRELLI.*Change-point Detection for Black-box Services*, in "Proceedings of the Eighteenth ACM SIGSOFT International Symposium on Foundations of Software Engineering", New York, NY, USA, FSE '10, ACM, 2010, p. 227–236, http://doi.acm.org/10.1145/1882291.1882326.

[41] N. ESFAHANI, S. MALEK.*Uncertainty in Self-Adaptive Software Systems*, in "Software Engineering for Self-Adaptive Systems II", R. LEMOS, H. GIESE, H. A. MULLER, M. SHAW (editors), Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2013, vol. 7475, p. 214-238, http://dx.doi.org/10.1007/978-3-642-35813-5_9.

[42] A. FILIERI, C. GHEZZI, G. TAMBURRELLI.*A Formal Approach to Adaptive Software: Continuous Assurance of Non-functional Requirements*, in "Form. Asp. Comput.", March 2012, vol. 24, n^{o} 2, p. 163–186, http://dx.doi.org/10.1007/s00165-011-0207-2.

[43] N. GEORGANTAS, G. BOULOUKAKIS, S. BEAUCHE, V. ISSARNY.*Service-oriented Distributed Applications in the Future Internet: The Case for Interaction Paradigm Interoperability*, in "ESOCC 2013 - European Conference on Service-Oriented and Cloud Computing", Malaga, Spain, K.-K. LAU, W. LAMERSDORF, E. PIMENTEL (editors), Lecture Notes in Computer Science, Springer, July 2013, vol. 8135, p. 134-148 [*DOI :* 10.1007/978-3-642-40651-5_11], http://hal.inria.fr/hal-00841332.

[44] S. HACHEM.*Service-oriented middleware for the large-scale mobile Internet of Things*, Université de Versailles Saint Quentin en Yvelines, February 2014.

[45] S. HACHEM, A. PATHAK, V. ISSARNY.*Probabilistic Registration for Large-Scale Mobile Participatory Sensing*, in "Proceedings of the 11th IEEE International Conference on Pervasive Computing and Communications, (PerCom)", Mar. 2013.

[46] S. HACHEM, A. TONINELLI, A. PATHAK, V. ISSARNY.*Policy-based Access Control in Mobile Social Ecosystems*, in "IEEE International Symposium on Policies for Distributed Systems and Networks", Pisa, Italie, IEEE computer society, 2011, http://hal.inria.fr/inria-00608201.

[47] S. HEMMINKI, P. NURMI, S. TARKOMA.*Accelerometer-based Transportation Mode Detection on Smartphones*, in "Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems", New York, NY, USA, SenSys '13, ACM, 2013, p. 13:1–13:14, http://doi.acm.org/10.1145/2517351.2517367.

[48] P. INVERARDI, M. TIVOLI.*Automatic Synthesis of Modular Connectors via Composition of Protocol Mediation Patterns*, in "Proceedings of the 2013 International Conference on Software Engineering", Piscataway, NJ, USA, ICSE '13, IEEE Press, 2013, p. 3–12, http://dl.acm.org/citation.cfm?id=2486788.2486790.

[49] V. ISSARNY, A. BENNACEUR, Y.-D. BROMBERG.*Middleware-layer Connector Synthesis: Beyond State of the Art in Middleware Interoperability*, in "11th International School on Formal Methods for the Design of Computer, Communication and Software Systems: Connectors for Eternal Networked Software Systems", M. BERNARDO, V. ISSARNY (editors), Lecture notes in computer science, Springer, 2011, vol. 6659, p. 217-255 [*DOI :* 10.1007/978-3-642-21455-4], http://hal.inria.fr/inria-00586630.

[50] V. ISSARNY, A. BENNACEUR.*Composing Distributed Systems: Overcoming the Interoperability Challenge*, in "FMCO 2012", F. DE BOER, M. BONSANGUE, E. GIACHINO, R. HAHNLE (editors), Lecture Notes in Computer Science, Springer, 2013, p. 168-196 [*DOI :* 10.1007/978-3-642-40615-7_6], http://hal.inria.fr/hal-00828801.

[51] J. JUN, Y. GU, L. CHENG, B. LU, J. SUN, T. ZHU, J. NIU.*Social-Loc: Improving Indoor Localization with Social Sensing*, in "Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems", New York, NY, USA, SenSys '13, ACM, 2013, p. 14:1–14:14, http://doi.acm.org/10.1145/2517351.2517352.

[52] C. KAISER, A. POZDNOUKHOV.*Enabling real-time city sensing with kernel stream oracles and MapReduce*, in "Pervasive and Mobile Computing", 2013, vol. 9, n° 5, p. 708-721, http://dblp.uni-trier.de/db/journals/percom/percom9.html#KaiserP13.

[53] A. KATTEPUR, N. GEORGANTAS, V. ISSARNY.*QoS Analysis in Heterogeneous Choreography Interactions*, in "11th International Conference on Service Oriented Computing (ICSOC)", Berlin, Germany, December 2013, http://hal.inria.fr/hal-00866190.

[54] J. O. KEPHART, D. M. CHESS.*The Vision of Autonomic Computing*, in "Computer", January 2003, vol. 36, n° 1, p. 41–50, http://dx.doi.org/10.1109/MC.2003.1160055.

[55] W. Z. KHAN, Y. XIANG, M. Y. AALSALEM, Q. ARSHAD.*Mobile Phone Sensing Systems: A Survey*, in "IEEE Communications Surveys Tutorials", 2013, vol. 15, n° 1, p. 402-427.

[56] N. D. LANE, Y. CHON, L. ZHOU, Y. ZHANG, F. LI, D. KIM, G. DING, F. ZHAO, H. CHA.*Piggyback CrowdSensing (PCS): Energy Efficient Crowdsourcing of Mobile Sensor Data by Exploiting Smartphone App Opportunities*, in "Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems", New York, NY, USA, SenSys '13, ACM, 2013, p. 7:1–7:14, http://doi.acm.org/10.1145/2517351.2517372.

[57] N. D. LANE, S. B. EISENMAN, M. MUSOLESI, E. MILUZZO, A. T. CAMPBELL.*Urban sensing systems: opportunistic or participatory?*, in "Proceedings of the 9th workshop on Mobile computing systems and applications", ACM, 2008, p. 11–16.

[58] N. D. LANE, E. MILUZZO, H. LU, D. PEEBLES, T. CHOUDHURY, A. T. CAMPBELL.*A survey of mobile phone sensing*, in "Communications Magazine, IEEE", 2010, vol. 48, n° 9, p. 140–150.

[59] LE PARISIEN.*Comment ne plus jamais être en retard*, Oct 2013, Article in "Le parisien : Comment ne plus jamais être en retard", Public article on the neverBLate app developed by team members, http://www.leparisien.fr/espace-premium/seine-et-marne-77/comment-ne-plus-jamais-etre-en-retard-23-10-2013-3250205.php.

[60] LE PARISIEN.*Fuyez les rames bondées*, Oct 2013, Article in "Le Parisien : fuyez les rames bondées, Public article on the Boîte à Sardines app developed by team members, http://www.leparisien.fr/espace-premium/paris-75/fuyez-les-rames-bondees-23-10-2013-3249981.php.

[61] X. LIU, A. BOUGUETTAYA, X. WU, L. ZHOU.*Ev-LCS: A System for the Evolution of Long-Term Composed Services*, in "IEEE Trans. Serv. Comput.", January 2013, vol. 6, n° 1, p. 102–115, http://dx.doi.org/10.1109/TSC.2011.40.

[62] H. LU, W. PAN, N. D. LANE, T. CHOUDHURY, A. T. CAMPBELL.*SoundSense: Scalable sound sensing for people-centric applications on mobile phones*, in "Proceedings of the 7th International conference on Mobile systems, applications, and services", New York, NY, USA, 2009, p. 165–178.

[63] E. MILUZZO, N. D. LANE, S. B. EISENMAN, A. T. CAMPBELL.*CenceMe – injecting sensing presence into social networking applications*, in "Smart Sensing and Context", Springer, 2007, p. 1–28.

[64] N. MITTON, S. PAPAVASSILIOU, A. PULIAFITO, K. S. TRIVEDI.*Combining Cloud and sensors in a smart city environment*, in "EURASIP Journal on Wireless Communications and Networking", 2012, vol. 2012, n^{o} 1, p. 1–10.

[65] L. MOTTOLA, G. P. PICCO.*Programming wireless sensor networks: Fundamental concepts and state of the art*, in "ACM Computing Surveys (CSUR)", 2011, vol. 43, n^{o} 3, 19.

[66] A. J. OLINER, A. P. IYER, I. STOICA, E. LAGERSPETZ, S. TARKOMA.*Carat: Collaborative Energy Diagnosis for Mobile Devices*, in "Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems", New York, NY, USA, SenSys '13, ACM, 2013, p. 10:1–10:14, http://doi.acm.org/10.1145/2517351.2517354.

[67] ONEM2M.*TS-0012 oneM2M Base Ontology*, 2016.

[68] A. PATHAK, V. K. PRASANNA.*Energy-efficient task mapping for data-driven sensor network macroprogramming*, in "Computers, IEEE Transactions on", 2010, vol. 59, n^{o} 7, p. 955–968.

[69] R. POPESCU, A. STAIKOPOULOS, A. BROGI, P. LIU, S. CLARKE.*A Formalized, Taxonomy-driven Approach to Cross-layer Application Adaptation*, in "ACM Trans. Auton. Adapt. Syst.", May 2012, vol. 7, n^{o} 1, p. 7:1–7:30, http://doi.acm.org/10.1145/2168260.2168267.

[70] R. K. RANA, C. T. CHOU, S. S. KANHERE, N. BULUSU, W. HU.*Ear-phone: An End-to-end Participatory Urban Noise Mapping System*, in "Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks", New York, NY, USA, IPSN '10, ACM, 2010, p. 105–116, http://doi.acm.org/10.1145/1791212.1791226.

[71] L. RAVINDRANATH, J. PADHYE, R. MAHAJAN, H. BALAKRISHNAN.*Timecard: Controlling User-perceived Delays in Server-based Mobile Applications*, in "Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles", New York, NY, USA, SOSP '13, ACM, 2013, p. 85–100, http://doi.acm.org/10.1145/2517349.2522717.

[72] M. SALEHIE, L. TAHVILDARI.*Self-adaptive Software: Landscape and Research Challenges*, in "ACM Trans. Auton. Adapt. Syst.", May 2009, vol. 4, n^{o} 2, p. 14:1–14:42, http://doi.acm.org/10.1145/1516533.1516538.

[73] G. SANTUCCI.*Privacy in the Digital Economy: Requiem or Renaissance? An Essay on the Future of Privacy*, 2013, Privacy Surgeon, http://www.privacysurgeon.org/blog/wp-content/uploads/2013/09/Privacy-in-the-Digital-Economy-final.pdf.

[74] M. SATYANARAYANAN, P. BAHL, R. CACERES, N. DAVIES.*The case for VM-based cloudlets in mobile computing*, in "Pervasive Computing, IEEE", 2009, vol. 8, n^{o} 4, p. 14–23.

[75] M. SATYANARAYANAN.*Mobile computing: The next decade*, in "ACM SIGMOBILE Mobile Computing and Communications Review", 2011, vol. 15, n^{o} 2, p. 2–10.

[76] M. SATYANARAYANAN.*Fundamental challenges in mobile computing*, in "Proceedings of the fifteenth annual ACM symposium on Principles of distributed computing", ACM, 1996, p. 1–7.

[77] I. SCHWEIZER, C. MEURISCH, J. GEDEON, R. BÄRTL, M. MÜHLHÄUSER.*NoiseMap: Multi-tier Incentive Mechanisms for Participative Urban Sensing*, in "Proceedings of the Third International Workshop on Sensing Applications on Mobile Phones", New York, NY, USA, PhoneSense'12, ACM, 2012, p. 9:1–9:5, http://doi.acm.org/10.1145/2389148.2389157.

[78] C. SHIN, J.-H. HONG, A. K. DEY.*Understanding and Prediction of Mobile Application Usage for Smart Phones*, in "Proceedings of the 2012 ACM Conference on Ubiquitous Computing", New York, NY, USA, UbiComp '12, ACM, 2012, p. 173–182, http://doi.acm.org/10.1145/2370216.2370243.

[79] R. SPEICYS CARDOSO, V. ISSARNY.*Architecting Pervasive Computing Systems for Privacy: A Survey*, in "Sixth Working IEEE/IFIP Conference on Software Architecture : WICSA 2007", Mumbai, Maharashtra, Inde, 2007, 26, http://hal.inria.fr/inria-00415925.

[80] M.-O. STEHR, C. TALCOTT, J. RUSHBY, P. LINCOLN, M. KIM, S. CHEUNG, A. POGGIO.*Fractionated software for networked cyber-physical systems: Research directions and long-term vision*, in "Formal Modeling: Actors, Open Systems, Biological Systems", Springer, 2011, p. 110–143.

[81] T. TEIXEIRA, S. HACHEM, V. ISSARNY, N. GEORGANTAS.*Service Oriented Middleware for the Internet of Things: A Perspective*, in "ServiceWave", Poznan, Pologne, Springer-Verlag, 2011, p. 220-229, http://hal.inria.fr/inria-00632794.

[82] A. TONINELLI, A. PATHAK, V. ISSARNY.*Yarta: A Middleware for Managing Mobile Social Ecosystems*, in "GPC 2011 : International Conference on Grid and Pervasive Computing", Oulu, Finlande, J. RIEKKI, M. YLIANTTILA, M. GUO (editors), Lecture notes in computer science, Springer, May 2011, vol. 6646, p. 209-220 [*DOI :* 10.1007/978-3-642-20754-9_22], http://hal.inria.fr/hal-00723794.

[83] S. VANSYCKEL, D. SCHAFER, V. MAJUNTKE, C. KRUPITZER, G. SCHIELE, C. BECKER.*COMITY: A framework for adaptation coordination in multi-platform pervasive systems*, in "Pervasive and Mobile Computing", 2014, vol. 10, Part A, p. 51 - 65, Selected Papers from the Eleventh Annual {IEEE} International Conference on Pervasive Computing and Communications (PerCom 2013) [*DOI :* 10.1016/J.PMCJ.2013.10.006], http://www.sciencedirect.com/science/article/pii/S1574119213001302.

[84] M. VON KAENEL, P. SOMMER, R. WATTENHOFER.*Ikarus: Large-scale Participatory Sensing at High Altitudes*, in "Proceedings of the 12th Workshop on Mobile Computing Systems and Applications", ACM, 2011, p. 63–68.

[85] C. XU, W. YANG, X. MA, C. CAO, J. LU.*Environment Rematching: Toward Dependability Improvement for Self-Adaptive Applications*, in "Automated Software Engineering (ASE), 2013 IEEE/ACM 28th International Conference on", Nov 2013, p. 592-597, http://dx.doi.org/10.1109/ASE.2013.6693118.

[86] J. ZIMMERMAN, A. TOMASIC, C. GARROD, D. YOO, C. HIRUNCHAROENVATE, R. AZIZ, N. R. THIRUVENGADAM, Y. HUANG, A. STEINFELD.*Field Trial of Tiramisu: Crowd-sourcing Bus Arrival Times to Spur Co-design*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", New York, NY, USA, CHI '11, ACM, 2011, p. 1677–1686, http://doi.acm.org/10.1145/1978942.1979187.

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:

5.3. - Image processing and analysis
5.9. - Signal processing
6.1.1. - Continuous Modeling (PDE, ODE)
6.2.1. - Numerical analysis of PDE and ODE
6.2.6. - Optimization

Other Research Topics and Application Domains:

1.3. - Neuroscience and cognitive science
9.4.2. - Mathematics
9.4.3. - Physics
9.4.4. - Chemistry
9.5.3. - Economy, Finance

1. Members

Research Scientists

Jean-David Benamou [Team leader, Inria, Senior Researcher, HDR]
Vincent Duval [Inria, Researcher]
Quentin Merigot [CNRS, Researcher]
Gabriel Peyré [CNRS, Senior Researcher, HDR]

Faculty Members

Guillaume Carlier [Univ. Paris IX, Professor, HDR]
François-Xavier Vialard [Univ. Paris IX, Associate Professor, from Feb 2016]

Technical Staff

Simon Legrand [Inria]

PhD Students

Lenaic Chizat [Univ. Paris IX, from Nov 2016]
Quentin Denoyelle [Univ. Paris IX]
Aude Genevay [Univ. Paris IX]
Maxime Laborde [Univ. Paris IX]
Marco Masoero [Univ. Paris IX]
Miao Yu [Univ. Paris VII, from Sep 2016]

Post-Doctoral Fellows

Roman Andreev [Univ. Paris VI]
Clarice Poon [Univ. Paris IX, until Aug 2016]
Bernhard Schmitzer [CNRS, from Feb 2016]

Visiting Scientist

Jonathan Vacher [Univ. Paris IX, from Apr 2016 until Sep 2016]

Administrative Assistant

Martine Verneuille [Inria]

Others

Gwendoline de Bie [Inria, Intern, from Jun 2016 until Sep 2016]
Christine Durok [Inria, Intern, from May 2016 until Jul 2016]
Jean Louet [Univ. Paris XI, ATER, from Sep 2015]
Luca Nenna [Univ. Paris XI, ATER from Sep 2016]

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 [131]). This engineering problem consists in minimizing the transport cost between two given mass densities. In the 40's, Kantorovich [139] 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 [82], 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 [183], [182], 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 [149] 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 [83] 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 [91] 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 [119] [147] [67] [68] and the references therein. Geometric approaches based on Laguerre diagrams and discrete data [156] 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 cost has been established by McCann and Gangbo [130]. Optimal Transportation techniques have been applied for example to a Coulomb ground cost in Quantum chemistry in relation with Density Functional theory [114]. 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 [160] 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 [51]. 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 [159].

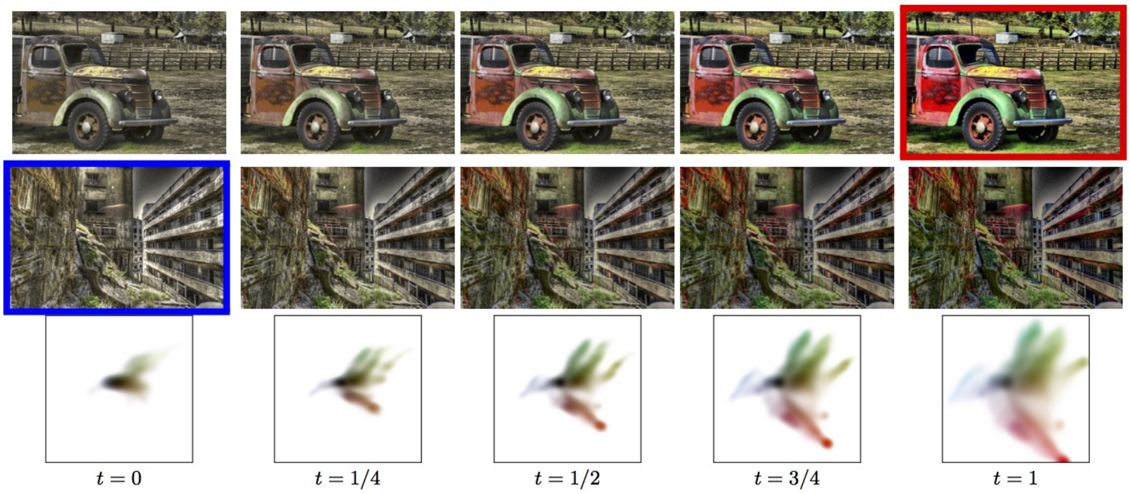

Figure 1. Example of color transfer between two images, computed using the method developed in [64], see also [176]. The image framed in red and blue are the input images. Top and middle row: adjusted image where the color of the transported histogram has been imposed. Bottom row: geodesic (displacement) 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 [136], [116], [115], [61], [146], mesh adaptation [145], the reconstruction of the early mass distribution of the Universe [127], [84] in Astrophysics, and the numerical optimisation of reflectors following the Optimal Transportation interpretation of Oliker [92] and Wang [184]. Extensions of OT such as multi-marginal transport has potential applications in Density Functional Theory , Generalized solution of Euler equations [81] (DFT) and in statistics and finance [58], [129] ...Recently, there has been a spread of interest in applications of OT methods in imaging sciences [76], statistics [72] and machine learning [118]. This is largely due to the emergence of fast numerical schemes to approximate the transportation distance and its generalizations, see for instance [64]. 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 [60]. 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 [62]. 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 [141] 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 [78].

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 [137]. 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 [138], [109] and a model of congested crowd motion proposed by Maury, Santambrogio and Roudneff-Chupin [152]. 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 in involves taking the derivative in the Wasserstein sense of the relevant energy, which in turns 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 to define the gradient flow. This gradient flow can either be understood as defining a geodesic on the (infinite dimensional) group of diffeomorphisms [57], or on a (infinite dimensional) space of curves or surfaces [185]. 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 [57], [135]. 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 problem through geodesic

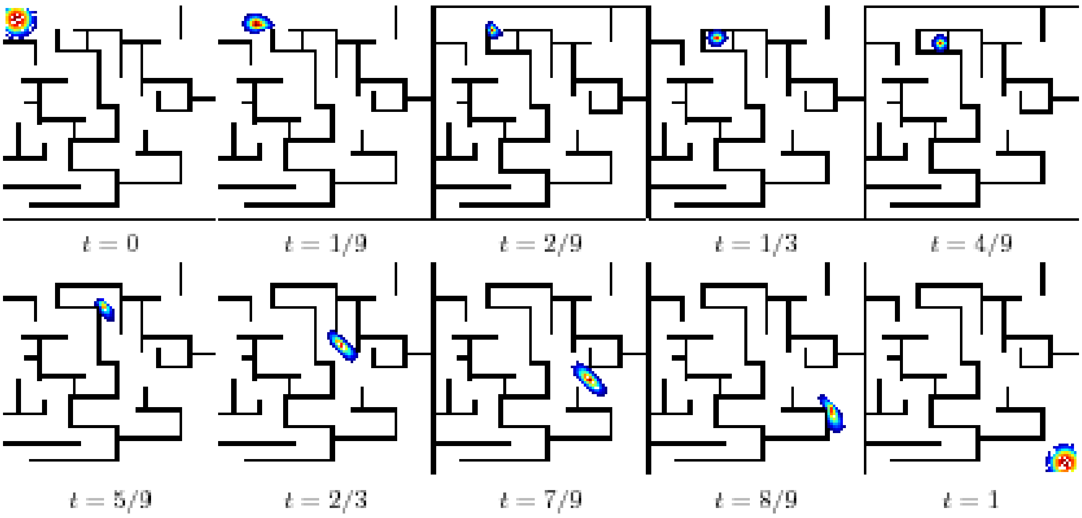

Figure 2. Examples of displacement interpolation (geodesic for optimal transport) according to a non-Euclidean Riemannian metric (the mass is constrained to move inside a maze) between to input Gaussian distributions. Note that the maze is dynamic: its topology change over time, the mass being "trapped" at time $t = 1/3$.

shooting method [153], 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 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) [93].

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 [151] and learned representation [186]).

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" [112] and in statistics by Tibshirani [177] 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 [94], 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 [96], 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 [95], [123], [120] 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 method is the total variation regularization [171], [88]. It favors low-complexity images that are piecewise constant, see Figure 3 for some example to solve some image processing problems. Beside applications in image processing, sparsity-related ideas also had a deep impact in statistics [177] and machine learning [53]. 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) [54]. 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 [144], [179] 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's 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 [60], [168]). 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 [64]) and machine learning (see [118]). 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 [169] reformulated the problem as a variational problem subject to a convexity constraint. For more general models, and using ideas from Optimal Transportation, Carlier [98] considered the more general c-convexity constraint and proved a general existence result. Using the formulation of [98] McCann, Figalli and Kim [124] gave conditions under which the principal agent problem can be written as an infinite dimensional convex variational problem. The important results of [124] 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 [104], [157], [154].

Goals: So far, the mathematical PA model can be numerically solved for simple utility functions. A Bregman approach inspired by [64] is currently being developed [101] 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* [58], [129] 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 [64], [69], [62] 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 [99] 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 usual 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) [137] 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 [137]), the porous medium equations, the granular media equation, just to give a few examples. It also finds application in image processing [87]. 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 [107])
- 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 [74]
- 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.

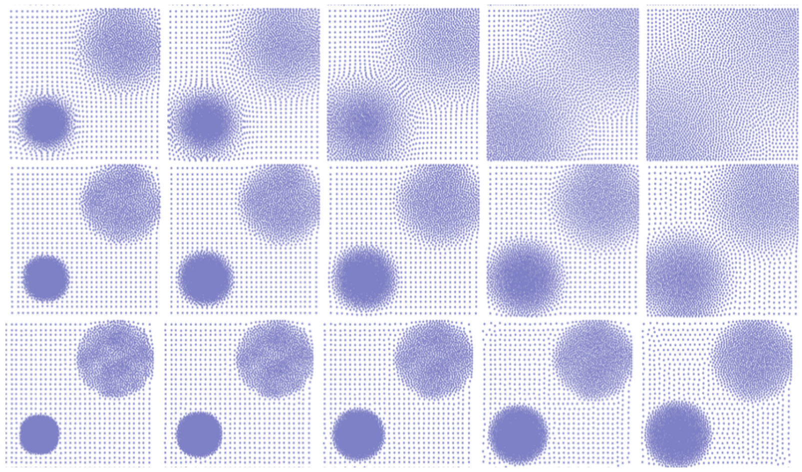

Figure 4. Example of non-linear diffusion equations solved with a JKO flow [65]. The horizontal axis shows the time evolution minimizing the functional $\int \frac{\rho^\alpha}{\alpha-1}$ on the density ρ (discretized here using point clouds, i.e. sum of Diracs' with equal mass). Each row shows a different 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 [103] which leads to very degenerate PDEs [79].

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 [69].

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 in Section 4.2.

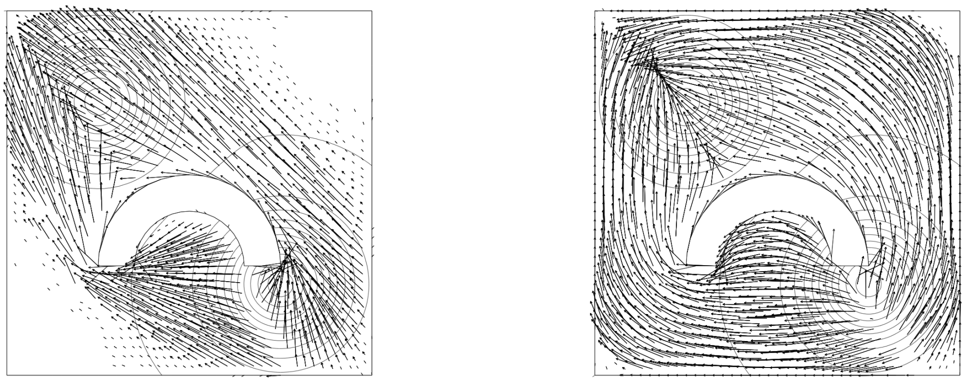

Figure 5. Monge and Wardrop flows of mass around an obstacle [62]. the source/target mass is represented by the level curves. Left : no congestion, Right : congestion.

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 [90] or tracking elastic organ movements [174]. Previous attempts [148], [165] 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 [60] and to define geodesic metrics on the space of shapes and diffeomorphisms [111].

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 [141] with an extremely wide range of potential applications [132]. 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 [63]. 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 [62] and on the analytical side [97] 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 [50].

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 [152] 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 [50], [86].
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 [62], 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) [180], [85], [178]. Our team has already studied flows and geodesics over non-Riemannian shape spaces, which allows for piecewise smooth deformations [111].

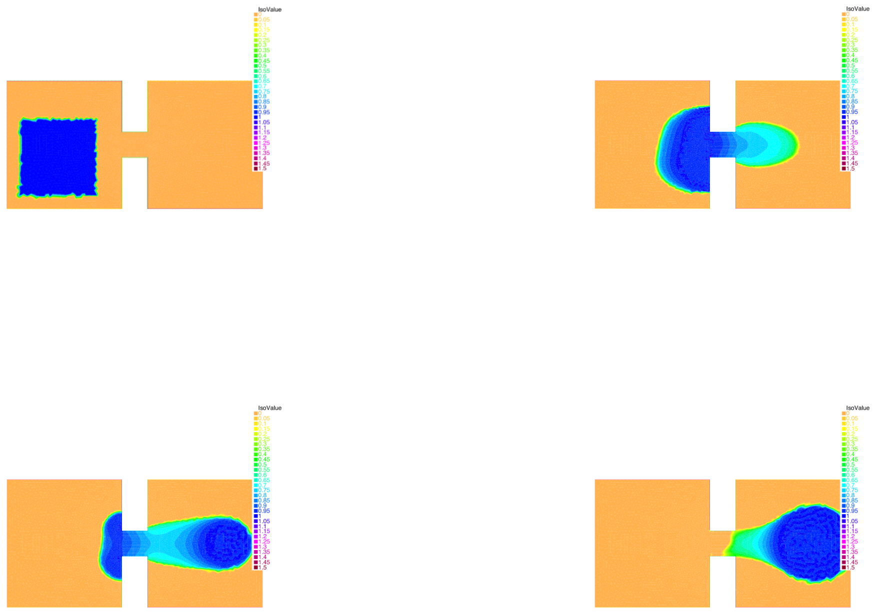

Figure 6. Example of crowd congestion with density dependent speed. The macroscopic density, at 4 different times, of people forced to exit from one room towards a meeting point in a second room.

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 [173], [174] 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 [143], 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 [181] 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 [125]. 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 [125], [126].
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 [179] 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 [75]) for discrete (finite dimensional) problems. We extended this result in a simple infinite dimensional setting, namely sparse regularization of Radon measures for deconvolution [120]. A deep understanding of those continuous inverse problems is crucial to analyze the behavior of their discrete counterparts, and in [121] 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

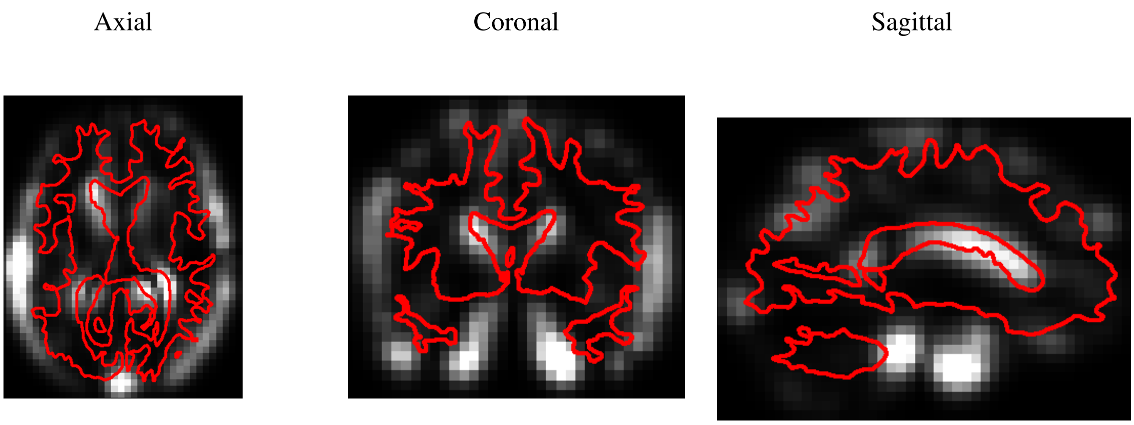

Figure 7. Learning Riemannian metrics in diffeomorphic image matching to capture the brain variability: a diagonal operator that encodes the Riemannian metric is learnt on a template brain out of a collection of brain images. The values of the diagonal operator are shown in greyscale. The red curves represent the boundary between white and grey matter. For more details, we refer the reader to [181], which was a first step towards designing effective and robust metric learning algorithms.

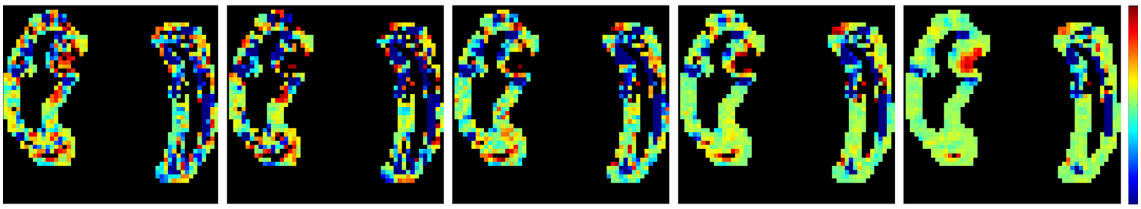

Figure 8. Statistics on initial momenta: In [125], we compared several intersubject transport methodologies to perform statistics on longitudinal evolutions. These longitudinal evolutions are represented by an initial velocity field on the shapes boundaries and these velocity fields are then compared 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 [100], [75] 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. he relevant previous work in this direction are the fundamental results of Chambolle, Caselles and co-workers [59], [52], [108]. 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 [120] 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 [164], [122]. This requirement is also at the heart of recently proposed models of cortical processing [163]. 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 [155], [122]. G. Peyré has done several contributions on the definition of geometric wavelets transform and directional texture models, see for instance [164]. Dario Prandi has recently applied methods from sub-Riemannian geometry to image restoration [77].
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 [70]) 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 [134]. 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 [172]. Typical approaches [140] try to recovered 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 [120].
Goals: We aim at extending the theoretical performance analysis obtained for sparse measures [120] to the set of piecewise regular 2-D and 3-D functions. Some interesting previous work of C. Poon et al [166] (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 [49].

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 [90], [162] 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 [150], which are connected to so-called metamorphoses models in the LDDMM framework [71].
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 [90], [162] 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 [68] for the Monge-Ampère equation and [154] for general non-linear variational problems. Both offer convergence guarantees and are amenable to fast numerical resolution techniques such as Newton solvers. Since [68] 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 [128], a new different and more accurate approach has been proposed by Mirebeau, Benamou and Collino [66]. As shown in [113], 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 [104], Quentin Mérigot [157] and J-M. Mirebeau [154]. 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 [137]) 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 [65] 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 [120], [121] 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 [105].
Our expertise: We have provided the first results on the discrete-to-continous convergence in both sparse regularization variational problems [120], [121] and the static formulation of OT and Wasserstein barycenters [105]
Goals: In a collaboration with Jérôme Malick (Inria Grenoble), our first goal is to generalized the result of [120] to generic partly-smooth convex regularizers routinely used in imaging science and machine learning, a prototypal example being the nuclear norm (see [179] for a review of this class of functionals). Our second goal is to extend the results of [105] to the novel class of entropic discretization schemes we have proposed [64], 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, C. Poon) 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 [96], [120]. 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 [120] 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 [96]). A possible, but very costly, methodology is to ressort to Lasserre’s SDP representation hierarchy [142]. 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 [171]), 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 [60], [158], congested optimal transport [89] and to sparse regularizations (see for instance [168] 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 [60], see also [158]. We have also recently developed new proximal schemes that can cope with non-smooth composite objectives functions [168].
Goals: We aim at extending these solvers to a wider class of variational problems, most notably optimization under divergence constraints [62]. 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.

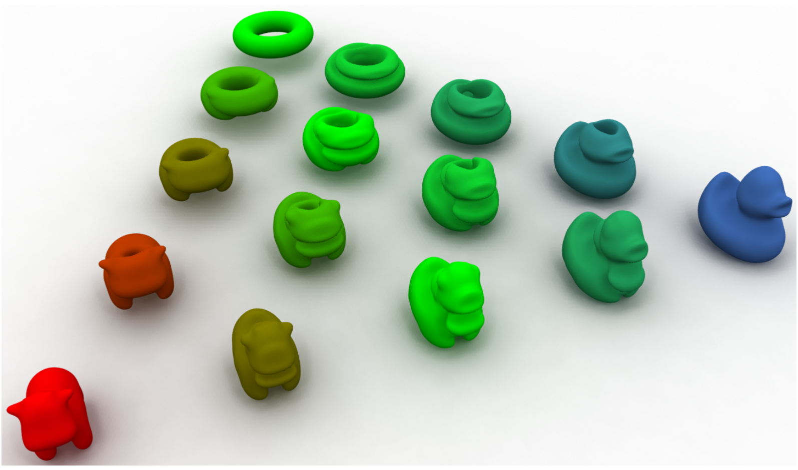

Figure 9. Example of barycenter between shapes computed using optimal transport barycenters of the uniform densities inside the 3 extremal shapes, computed as detailed in [176]. Note that the barycenters are not in general uniform distributions, and we display them as the surface defined by a suitable 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 [118]. As shown in [64], 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 [64] shown how Bregman projections [80] and Dykstra algorithm [56] offer a generic optimization framework to solve a variety of generalized OT problems. Carlier and Dupuis [101] 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 [176]. 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 [110]) 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 [114]) and fluid dynamics (Brenier's weak solutions of the incompressible Euler equation, see [81]). 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. Application Domains

4.1. Freeform Optics

Following the pioneering work of Caffarelli and Oliker [92], Wang [184] has shown that the inverse problem of freeforming a *convex* reflector which sends a prescribed source to a target intensity is a particular instance of Optimal Transportation. This is a promising approach to automatize the industrial design of optimised energy efficient reflectors (car/public lights for instance). We show in figure 10 the experiment setting and one of the first numerical simulations produced by the ADT Mokabajour.

The method developed in [68] has been used by researchers of TU Eindhoven in collaboration with Philips Lightning Labs to compute reflectors [167] in a simplified setting (directional light source). Another approach, based on a geometric discretization of Optimal Transportation has been developed in [8], and is able to handle more realistic conditions (punctual light source).

Solving the exact Optimal Transportation model for the Reflector inverse problem involves a generalized Monge-Ampère problem and is linked to the open problem of c-convexity compatible discretization we plan to work on. The corresponding software development is the topic of the ADT Mokabajour.

4.1.1. Software and industrial output.

See section 4.3 below for softwares. These methods will clearly become mainstream in reflector design but also in lense design [170]. The industrial problems are mainly on efficiency (light pollution) and security (car head lights) based on free tailoring of the illumination. The figure below is an extreme test case where we exactly reproduce an image. They may represent one of the first incursion on PDE discretisation based methods into the field of non-imaging optics.

4.2. Metric learning for natural language processing

The analysis of large scale datasets to perform un-supervised (clustering) and supervised (classification, regression) learning requires the design of advanced models to capture the geometry of the input data. We believe that optimal transport is a key tool to address this problem because (i) many of these datasets are composed of histograms (social network activity, image signatures, etc.) (ii) optimal transport makes use of a ground metric that enhances the performances of classical learning algorithms, as illustrated for instance in [118].

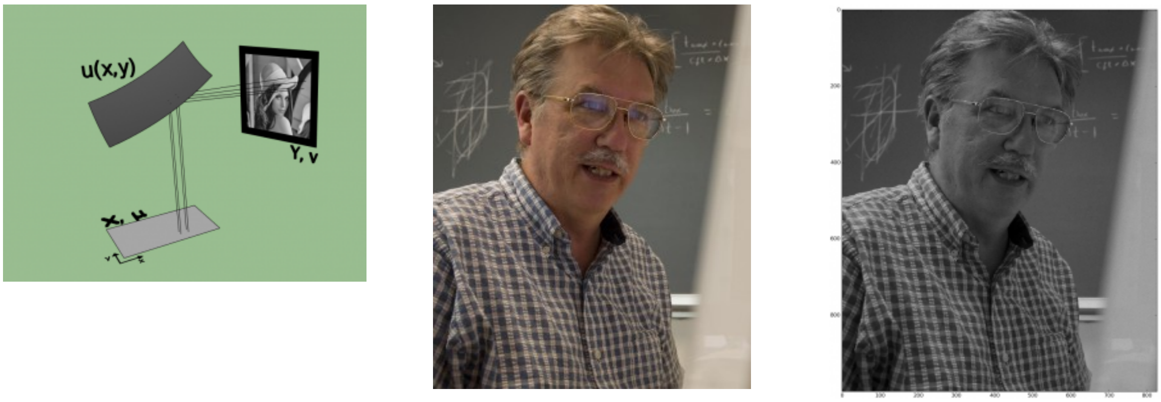

Figure 10. A constant source to a prescribed image (center). The reflector is computed (but not shown) and a resimulation using ray tracing shows the image reflected by the computed reflector.

Some of the theoretical and numerical tools developed by our team, most notably Wasserstein barycenters [51], [76], are now becoming mainstream in machine learning [72], [118]. In its simplest (convex) form where one seeks to only maximize pairwise wasserstein distances, metric learning corresponds to the congestion problem studied by G. Carlier and collaborators [106], [79], and we will elaborate on this connection to perform both theoretical analysis and develop numerical schemes (see for instance our previous work [69]).

We aim at developing novel variational estimators extending classification regression energies (SVM, logistic regression [133]) and kernel methods (see [175]). One of the key bottleneck is to design numerical schemes to learn an optimal metric for these purpose, extending the method of Marco Cuturi [117] to large scale and more general estimators. Our main targeted applications is natural language processing. The analysis and processing of large corpus of texts is becoming a key problems at the interface between linguistic and machine learning [55]. Extending classical machine learning methods to this field requires to design suitable metrics over both words and bag-of-words (i.e. histograms). Optimal transport is thus a natural candidate to bring innovative solutions to these problems. In a collaboration with Marco Cuturi (Kyoto University), we aim at unleashing the power of transportation distances by performing ground distance learning on large database of text. This requires to lift previous works on distance on words (see in particular [161]) to distances on bags-of-words using transport and metric learning.

4.3. Physics

The Brenier interpretation of the generalized solutions of Euler equations in the sense of Arnold is an instance of multi-marginal optimal transportation, a recent and expanding research field which also appears in DFT (see chemistry below). Recent numerical developments in OT provide new means of exploring these class of solutions.

In the years 2000 and after the pioneering works of Otto, the theory of *many-particle systems* has become "geometrized" thanks to the observed intimate relation between the geometric theory of geodesic convexity in the Wasserstein distance and the proof of entropy dissipation inequalities that determine the trend to equilibrium. The OT approach to the study of equilibration is still an extremely active field, in particular the various recently established connections to sharp functional inequalities and isoperimetric problems.

A third specific topic is the use of optimal transport models in *non-imaging optics*. Light intensity here plays the role of the source/target prescribed mass and the transport map defines the physical shape of specular reflector or refracting lense achieving such a transformation. This models have been around since the works of

Figure 11. Examples of two histogram (bag-of-words) extracted from the congress speech of US president. In this application, the goal is to infer a meaningful metric on the words of the english language and lift this metric to histogram using OT technics.

Oliker and Wang in the 90's. Recent numerical progresses indicate that OT may have an important industrial impact in the design of optical elements and calls for further modelisation and analysis.

4.4. Chemistry

The treatment of *chemical reactions* in the framework of OT is a rather recent development. The classical theory must be extended to deal with the transfer of mass between different particle species by means of chemical reactions.

A promising and significant recent advance is the introduction and analysis of a novel metric that combines the pure transport elements of the Wasserstein distance with the annihilation and creation of mass, which is a first approximation of chemical reactions. The logical next challenge is the extension of OT concepts to vectorial quantities, which allows to rewrite cross-diffusion systems for the concentration of several chemical species as gradient flows in the associated metric. An example of application is the modeling of a *chemical vapor deposition process*, used for the manufacturing of thin-film solar cells for instance. This leads to a degenerate cross-diffusion equations, whose analysis — without the use of OT theory — is delicate. Finding an appropriate OT framework to give the formal gradient flow structure a rigorous meaning would be a significant advance for the applicability of the theory, also in other contexts, like for biological multi-species diffusion.

A very different application of OT in chemistry is a novel approach to the understanding of *density functional theory* (DFT) by using optimal transport with "Coulomb costs", which is highly non convex and singular. Albeit this theory shares some properties with the usual optimal transportation problems, it does not induce a metric between probability measures. It also uses the multi-marginal extension of OT, which is an active field on its own right.

4.5. Biology

OT methods have been introduced in biology via gradient flows in the Wasserstein metric. Writing certain *chemotaxis* systems in variational form allowed to prove sharp estimates on the long time asymptotics of the bacterial aggregation. This application had a surprising payback on the theory: it lead to a better understanding and novel proofs of important functional inequalities, like the logarithmic Hardy-Littlewood-Sobolev inequality. Further applications followed, like transport models for species that avoid over-crowding, or cross-diffusion equations for the description of *biologic segregation*. The inclusion of dissipative cross-diffusion systems into the framework of gradient flows in OT-like metrics appears to be one of the main challenges for the future development of the theory. This extension is not only relevant for biological applications, but is clearly of interest to participants with primary interest in physics or chemistry as well.

Further applications include the connection of OT with game theory, following the idea that many selection processes are based on competition. The ansatz is quite universal and has been used in other areas of the *life sciences* as well, like for the modeling of personal income in economics.

4.6. Medical Imaging

Applications of variational methods are widespread in medical imaging and especially for diffeomorphic image matching. The formulation of large deformation by diffeomorphisms consists in finding geodesics on a group of diffeomorphisms. This can be seen as a non-convex and smoothed version of optimal transport where a correspondence is sought between objects that can be more general than densities. Whereas the diffeomorphic approach is well established, similarity measures between objects of interest are needed in order to drive the optimization. While being crucial for the final registration results, these similarity measures are often non geometric due to a need of fast computability and gradient computation. However, our team pioneered the use of entropic smoothing for optimal transport which gives fast and differentiable similarity measures that take into account the geometry. Therefore, we expect an important impact on this topic, work still in progress. This example of application belongs to the larger class of inverse problems where a geometric similarity measure such as optimal transport might enhance notably the results. Concerning this particular application, potential interactions with the Inria team ARAMIS and also the team ASCLEPIOS can leverage new proposed similarity measure towards a more applicative impact.

4.7. Economics

Recent years have seen intense cross-fertilization between OT and various problems arising in economics. The principal-agent problem with adverse selection is particularly important in modern microeconomics, mathematically it consists in minimizing a certain integral cost functional among the set of c-concave functions, this problem is convex under some conditions related to the MTW regularity theory for OT as shown in the important paper [124]. Other examples of fruitful interactions between mathematical economics concern multi-marginal OT and multi-populations matching [102], or games with a continuum of agents and Cournot-Nash equilibria [73]. The team has as strong expertise, both numerical and theoretical in the field of variational problems subject to a convexity constraint and their applications to the principal-agent problem. Our expertise in numerical OT and entropic regularization will also enable us to develop efficient solvers for realistic matching and hedonic pricing models.

5. Highlights of the Year

5.1. Highlights of the Year

New ERC Grant for G. Peyré

Gabriel Peyré is the recipient of a second ERC grand (consolidator), project NORIA (http://www.gpeyre.com/noria/) on Numerical Optimal tRansport for ImAging, that will start on Oct. 2017.

Pisa

Four members of Mokaplan : G. Peyré, G. Carlier, J-D. Benamou, Simone di Marino (starting 2017) have been invited speakers at the Pisa Scuola Normale Bi-Annual Optimal Transport Conference (November 7-11). This is considered as the most prestigious conference in the field.

6. New Software and Platforms

6.1. 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 headlightsâŠ 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. Different solvers

- Participants: Simon Legrand, Jean-David Benamou, Quentin Merigot and Boris Thibert
- Contact: Jean-David Benamou
- URL: https://project.inria.fr/mokabajour/

6.2. Wasserstein Fisher Rao

Lenaic Chizat

https://github.com/lchizat/optimal-transport

This Julia toolbox provides several tools for solving optimal transport, the unbalanced extensions and related problems.

6.3. OT and Sparse Scaling

Bernhard Schmitzer

http://wwwmath.uni-muenster.de/num/wirth/people/Schmitzer/index.html?content=code

7. New Results

7.1. Inverse problems with sparsity prior

G. Peyré, V. Duval, Q. Denoyelle,C. Poon

In [12], we have studied the stability of a classical image processing method, the Total Variation (TV) Denoising model introduced by Rudin, Osher and Fatemi [171]. While TV denoising is a well studied problem, our contribution is one of the first to address the impact of noise on the solutions. We have shown that the level lines of the denoised image (hence the edges and the gradient of shades) are located near an area called “extended support” which depends on the curvature of the image to recover. This yields a precise description of the so-called “staircasing” effect which is characteristic of the method, as well as the support stability of the method (see Figure 12). In particular, we have proved that indicator functions of calibrable sets are stable to noise, in the sense that the level lines of the denoised image will be close to the boundary of the original set.

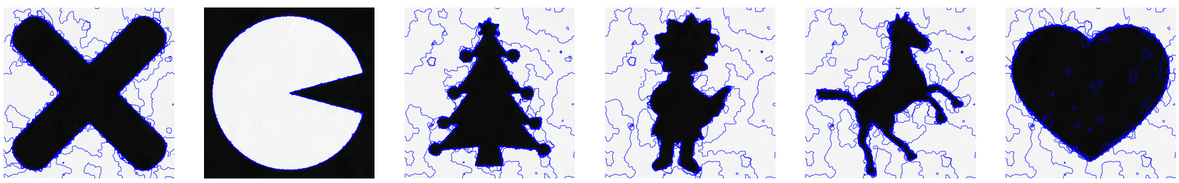

Figure 12. Level lines of denoised images with low regularization (TV denoising)

In [38], we have studied the problem of recovering a sparse signal (say, a sum of Dirac masses), from its blurred, partial, Radon transform, or equivalently by sampling the low frequency coefficients of its Fourier transform along a few radial lines. We have proved that, using a total variation (of measures) regularization approach in the spirit of [96], one may reconstruct exactly the signal under some geometric condition, or, in a compressed sensing approach, with high probability if one subsamples the coefficients. We propose a numerical algorithm to exactly solve this problem, by a converting it to a few low-dimensional Semi-Definite Programs.

7.2. Mean Field Games and augmented lagrangian methods for optimal transport

Roman Andreev

We apply the augmented Lagrangian method to the convex optimization problem of the instationary variational mean field games with diffusion. The system is first discretized with space-time tensor product piecewise polynomial bases. This leads to a sequence of linear problems posed on the space-time cylinder that are second order in the temporal variable and fourth order in the spatial variable. To solve these large linear problems with the preconditioned conjugate gradients method we propose a parameter-robust preconditioner that is based on a temporal transformation coupled with a spatial multigrid. Numerical examples illustrate the method. [27].

G. Carlier J-D. Benamou have written in collaboration with F. Santambrogio a review paper on variational MFG [31] both on theoretical and numerical aspects, the latter being addressed by augmented Lagrangian techniques developed by our team also in the context of optimal transport for an arbitrary Finsler metric cost [30] (the main advantage of our method being that we never have to evaluate the cost).

7.3. Gromov-Wasserstein methods in graphics and machine learning

G. Peyré, J. Solomon, M. Cuturi

A bottleneck of optimal transport (OT) methods for some applications in graphics and machine learning is that it requires the knowledge of an a priori fixed ground cost. This cost is often chosen as some power of a distance, which in turn requires that the data to compare or modify are pre-registered in a common embedding metric space (e.g. the 3-D or 2-D Euclidean space for shapes matching). For many applications (such a shape matching in vision or molecule comparison in quantum chemistry), this is simply not the case. We thus propose in [18], [21] to extend the computational machinery of OT to cope with an unknown cost by using the so-called Gromov-Wasserstein distance. This distance allows to compare probability distributions living in *different* and un-registered metric spaces, by coupling together pairs of points instead of single points. This allows to formulate a non-convex energy minimization, which is similar to the graph matching problem. We propose to use the entropic regularization scheme to solve it numerically, and we showed that it leads to a very effective Sinkhorn-like algorithm. In [18] (published in SIGGRAPH, the best computer graphics conference) we explore various application in computer graphics (such as shape matching or organization of collections of surfaces and images), while [21] (published in ICML, one of the two best machine learning conference) we extend this machinery to compute interpolation and barycenter of several metric space, with application to shape interpolation and supervised learning for quantum chemistry.

7.4. Optimal transport meets machine learning

A. Genevay, G. Peyré, M. Cuturi, F. Bach

Optimal transport has recently proved (in particular through the works of our team) to be very successful to solve various low dimensional problems, mostly in 2-D and 3-D. These successes are mainly due to the specicfic structure of these problems (the connections with PDE's and the use of entropic regularization), but these approaches do not scale to high dimensional and large scale problems that one encounters in machine learning. In these problems, it is not possible to discretize the space, and one does not have a direct access to the density to compare. One can rather only *sample* from these distributions. To address these difficulties,

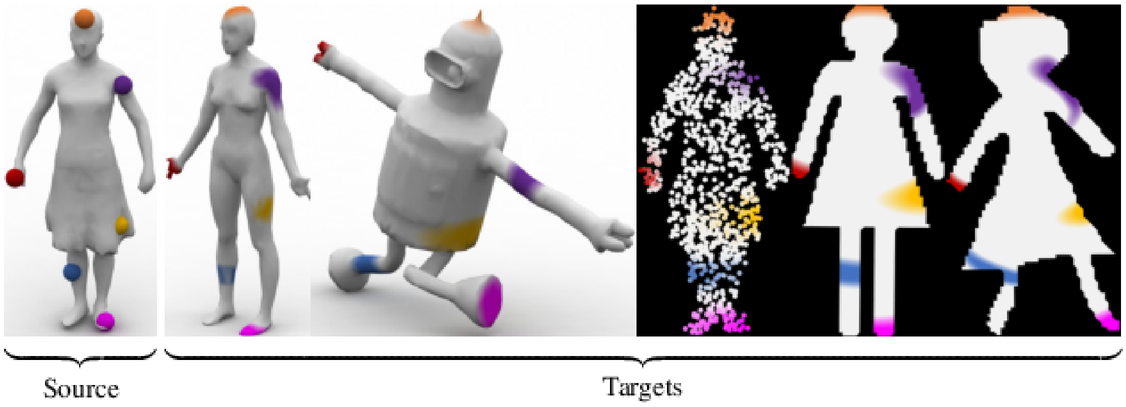

Figure 13. Example of matching induced between an input 3-D shape (on the left) and 3-D or 2-D shapes using the transport coupling computed using our entropy-regularized Gromov-Wasserstein problem. From [18].

we propose in [20] (published in NIPS, one of the best two machine learning conferences), the first provably convergent algorithm that can cope with high dimensional OT problems, with both discrete and continuous input measures. This approach leverage both the structure of the dual problem, and the smoothness induced by an entropic regularization. We show application of this method for classification of high dimensional bag of features histograms.

7.5. Optimal Transportation numerical methods for Fluid models

F-X. Vialard Q. Mérigot L. Nenna G. Carlier J-D. Benamou

Several new algorithms based on Optimal Transport have applied to Generalized Euler Geodesics and the Cauchy problem for the Euler equation. The methods rely on the generalized polar decomposition of Brenier, numerically implemented whether through semi-discrete optimal transport or through entropic regularization. It is robust enough to extract non-classical, multi-valued solutions of Euler's equations predicted by Brenier and Schnirelman. The semi-discrete approach also leads to a numerical scheme able to approximate regular solutions to the Cauchy problem for Euler equations. See Luca Nenna Thesis and [15].
A new link between optimal transport and fluid dynamic was discovered in [42]. Since the work of Brenier, optimal transport is tightly linked with the incompressible Euler equation and can be seen as a nonlinear extension of the pressure. Recently, a new optimal transport model between unbalanced measures has been proposed by some of the members of Mokaplan. In [41], it is shown that the corresponding fluid dynamic equation is the Camassa-Holm equation, well known to model waves in shallow water and wave breaking. On the theoretical side, we prove that the solutions to the Camassa-Holm equation can be seen as particular solutions of the incompressible Euler equation. This work paves the way for the study of the generalized Camassa-Holm geodesics and numerical methods based on unbalanced optimal transport scaling algorithms to solve it.

7.6. Scaling Algorithms and OT

G. Peyré F-X. Vialard L. Chizat B. Schmitzer S. Di Marino

B. Schmitzer has developed a sparse solver based on entropic regularization and numerical methods to solve unbalanced optimal transport (developed by our team in 2015) have been proposed in [37]. The core of the method consists in using the entropy functional as a reguflarizer and a barrier method. This is a generalization of the Sinkhorn method that has been introduced recently by M. Cuturi in numerical optimal transport. One important contribution of this work is to give a unified formulation of unbalanced optimal transport that can address a whole range of possible metrics and encompasses different applications such as Karcher-Fréchet averages, gradient flows, multimarginal unbalanced optimal transport. These two works are essentially based on a log-domain stabilized formulation, an adaptive truncation of the kernel and a coarse-to-fine scheme. This allows to solve large problems where the regularization is almost negligible.
In particular, this scaling algorithm is applied in its gradient flow formulation in the unbalanced case to obtain accurate simulations of the Hele-Shaw model, which models the cancer tumor growth.

7.7. Optimal transport meets economics

G. Carlier J-D. Benamou L. Nenna, G. De Bie

G. Carlier and L. Nenna in collaboration with Adrien Blanchet [32] developed an entropic-regularization scheme to compute Cournot Nash equilibria (i.e. equilibria in games with a continuum of players) for generic costs. With Lina Mallozzi, G. Carlier [36] introduced a partial optimal mass transport approach for spatial monopoly pricing both in the deterministic and stochastic cases. G. Carlier, J-D. Benamou and X. Dupuis developed various numerical strategies for solving the principal-agent problem in the framework of optimal pricing. Carlier, Chernozhukov and Galichon [34] studied multivariate quantile regression by optimal transport and duality techniques beyond the specified case, Gwendoline de Bie implemented these ideas by entropic regularization.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Optimal Transport applied to altimetric CTH dynamic interpolation

(S. Legrand V. Duval L. Chizat J-D. Benamou).

This collaboration between CLS and and funded by CNES intends to test on Column of Tropospheric Humidity data Optimal transportation interpolation techniques for balanced and unbalanced data.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

J-D. Benamou is the coordinator of the ANR ISOTACE (Interacting Systems and Optimal Transportation, Applications to Computational Economics) ANR-12-MONU-0013 (2012-2016). The consortium explores new numerical methods in Optimal Transportation AND Mean Field Game theory with applications in Economics and congested crowd motion. Check https://project.inria.fr/isotace/.
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 and F-X. Vialard are members of ANR MAGA 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

9.1.2. CNRS Mission pour l'interdisciplinarité (Défi Imag'In)

V. Duval and F-X. Vialard are members of the CAVALIERI project (CAlcul des VAriations pour L'Imagerie, l'Edition et la Recherche d'Images). This project, coordinated by V. Duval, aims at proposing new methods for comparing and reconstructing images relying on recent progress in the calculus of variations. Typical applications are co-segmentation, statistics transfer and interpolation, as well as tomographic reconstruction. A major emphasis is given on methods derived from (generalized) Optimal Transportation. See http://image.math.u-bordeaux1.fr/cavalieri/

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

Gabriel Peyré is the principal investigator of the ERC project SIGMA-Vision (http://gpeyre.github.io/sigma-vision/), running in 2011-2016. This project tackles theory, numerics and applications at the interface between imaging sciences, optimization and neurosciences. It features in particular several contributions on sparse regularization techniques for inverse problems, and optimal transport approaches for color and texture image processing. This theoretical and numerical contributions are applied to compute vision, computer graphics and neurosciences of the visual brain. Gabriel Peyré is the recipient of a second ERC grand (consolidator), project NORIA (http://www.gpeyre.com/noria/) on Numerical Optimal tRansport for ImAging, that will start on Oct. 2017.

9.3. International Initiatives

9.3.1. MOKALIEN

Title: Numerical Optimal Transportation in (Mathematical) Economics

International Partner (Institution - Laboratory - Researcher):

McGill University (Canada) - mathematics - Oberman Adam

Start year: 2014

See also: https://team.inria.fr/mokaplan/mokalien/

The team investigates new modelling and numerical resolution methods in Mathematical Economics using the theory of Optimal Transportation.

9.3.2. Participation in International Programs

F-X. Vialard was invited to participate in Mathematics of Shapes and Applications (4 - 31 July 2016) held in Singapore.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

The following people visited MOKAPLAN during 2016.

- Lina Mallozzi (Professor, Napoli): Feb. 28-March 5
- Andrei Sobolevski (Research Associate, Moscow) and Aleksei Kroshnin (PhD Student, Moscow): Oct 17-Oct 21
- Teresa Radice (Research Associate, Napoli): Jan. 25-Jan. 31, Apr. 7-Apr. 15 and Jul. 25-Aug. 10
- Giuseppe Buttazzo (Professor, Pisa): Nov. 29-Dec. 2

9.4.2. Visits to International Teams

9.4.2.1. Research Stays Abroad

Carlier stayed three weeks in Canada in July, one week in Victoria for a collaboration with Agueh (and a master committee) and two weeks in Montreal for the mokalien meeting and then duscussions with Oberman, he visited Naples twice (one week each time, to work with Mallozzi and Radice), Pisa twice (one week each time, to work with Buttazzo), NYU (3 days).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

V. Duval and F-X. Vialard have organized the CAVALIERI workshop which was held in the Inria Paris research center (October, 11th and 12th).

The team has organized the “Journées MokaTAO” together with the McTAO team on October 3rd and 4th. G. Peyré co-organized the SIGMA 2016 conference at the CIRM in Nov. 2016.

J-D. Benamou has co-organized Computational Optimal Transportation Workshop at CRM montreal (July 18-22).

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

Guillaume Carlier is in the board of Journal de l’école Polytechnique, Applied Mathematics and optimization (since 2016) and Mathematics and financial economics, with Filippo Santambrogio and Thierry Champion he co-edited a special issues of RICAM Series devoted to optimal transport. G. Peyré is editor for SIAM Journal of Imaging Sciences and Springer Journal of Mathematical Imaging and Vision. He co-edited a special issues of RICAM Series devoted to inverse problems.

10.1.2.2. Reviewer - Reviewing Activities

The members of the team are frequently reviewing papers in SIIMS (SIAM Journal on Imaging Sciences), JMAA (Journal of Mathematical Analysis and Applications), IPol (Image Processing Online), JVCI (Journal of Visual Communication and Image Representation), COCV, M2AN ... Discrete and computational geometry, Journal of the London Math Society, JOTA, JCP, “Information and Inference: A Journal of the IMA”, JMIV, Optimization Letters, PAMI, SIAM optimization and control, IPMI and MICCAI (leading conferences in medical imaging).

10.1.3. Invited Talks

V. Duval was invited to give talks at: the Oberwolfach workshop on Mathematical Imaging and Surface processing (January 2016), the Séminaire Parisien de Statistique (SEMSTAT, March 2016), the McGill-Mokalien Workshop on Numerical Optimal transport (July 2016) and the Demi-heure de Science at Inria Paris (October 2016).

G. Carlier gave seminars in Grenoble, Liège, Bielefeld and NYU, was plenary speaker at Smai-Mode conference (Toulouse), gave talks at the workshop Nonlinear problems from materials science and shape optimization (Pisa), workshop computational optimal transport (Montreal), workshop New Developments in Econometrics and Time Series (Madrid), MAFE Meeting (Bielefeld) and OTT16 (Pisa).

G. Peyré was plenary speaker at : Oxford Summer School on inverse problems (Jul 2016) ; Optimization without Border, (les Houches, Feb 2016) ; UCL workshop on sparse signal processing (Sep. 2016) ; workshop computational optimal transport (Montreal, Jul 2016) ; OTT16 (Pisa, Dec 2016).

J-D. Benamou talked at the Calculus of Variation Seminar in U. Paris Diderot (Nov.), he was invited speaker at OTT16 (Pisa, Dec 2016).

F-X. Vialard gave talks in séminaire de mathématiques appliquées, Cermics, ENPC, janvier; Geometric analysis theory in vision and control conference, Voss - Bergen, May; SIAM Imaging Science, Large Scale Inverse Problems in Medical Imaging, Albuquerque, May; Mathematics of Shapes and Applications, Singapore, July; Geometric Measure Theory: Analysis and non-smooth objects, CIMI analysis semester, Toulouse, September; MokaTAO meeting, Inria Paris, October; séminaire calcul des variations, Orsay, October; Journée transport optimal, équation de Monge-Ampère et applications, IHES, December.

10.1.4. Leadership within the Scientific Community

G. Peyré is in the scientific boards of Fondation Sciences Mathématiques de Paris (since 2013) ; Chaire CFM-ENS on Data Sciences (since 2016) ; Chaire Havas-Dauphine END (since 2013); Ceremade Paris-Dauphine (2013-2016)

10.1.5. Research Administration

J-D. Benamou is an elected member of the "Conseil Académique" of the PSL COMUE.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence : V. Duval, Analyse fonctionnelle, 27h équivalent TD, niveau L3, INSA Rouen-Normandie

Master : G. Peyré, Master 2 MVA on imaging and machine learning (Cachan) $30h$.

Licence : F-X. Vialard, Algèbre linéaire 2, Université Paris-Dauphine, $30h$.

10.2.2. Supervision

Internship : Gwendoline de Bie did her master internship (2nd year at ENSAE Paris-Tech) with G. Carlier on *Entropic regularization for multivariate quantile regresssion*,

Internship : Christine Durok *Une méthode itérative pour le probléme inverse du réflecteur discret* J-D. Benamou Q. Mérigot

PhD in progress: Julien André, CIFRE PhD thesis with with the company OPTIS Grenoble-INP (co-supervision D. Attali, B. Thibert, Q. Mérigot)

PhD in progress : Jocelyn Meyron, *Extension of semi-discrete Optimal transport to other costs*, lED de Grenoble, Q. Mérigot, D. Attali and B. Thibert.

PhD in progress : Miao Yu, *Optimal Transport distances and Geophyscial imaging* J-D. Benamou (co-direction J.-P. Vilotte, IPGP).

PhD in progress : Paul Catala, *Low-rank approaches for off-the-grid superresolution*, October 2016, G. Peyré and V. Duval.

PhD in progress : Lenaic Chizat, *Unbalanced Optimal Transport* , october 2014, F-X. Vialard and G. Peyré.

PhD in progress : Aude Genevay, *Optimal Transport for Machine Learning* , october 2015, J-D. Benamou and G. Peyré.

PhD in progress : Quentin Denoyelle, *Off-the-grid super-resolution: theory, algorithms and applications in fluorescence imaging*, October 2014, G. Peyré and V. Duval.

Postdoc completed: Dario Prandi, sub-Riemannian model for imaging, Oct. 2015 – Aug. 2016, G. Peyré and J-M Mirebeau

Postdoc completed: Thomas Gallouët, Fluid model and optimal transport, Oct. 2015 – Aug. 2016, Q. Mérigot and Yann Brenier.

Postdoc in progress: Roman Andreev, Numerical Methods for Mean Field Games , Mai 2015, Yves Achdou and J-D. Benamou.

PhD completed: Luca Nenna, Numerical Methods for Multi-Marginal Optimal Transportation, October 2013-December 2016, J-D. Benamou and G. Carlier.

PhD completed: Maxime Laborde, Systemes de particules en interaction, approche par flot de gradient dans l'espace de Wasserstein, september 2013-December 2016, G. Carlier.

PhD in progress: Jonathan Vacher, Machine learning approaches for neurosciences of the visual brain, October 2013-Jan 2017, G. Peyré and C. Monier.

Postdoc completed: Bernhard Schmitzer, fast algorithms for optimal transport, Oct. 2014-August 2016, G. Peyré.

Postdoc completed: Clarice Poon, *Support recovery using total variation and others sparse priors*, September 2015-August 2016, G. Peyré and V. Duval.

10.2.3. Juries

Vincent Duval was in the PhD comittee of Romain Hug (Grenoble, December, 9).
J-D. Benamou and G. Carlier were in the PhD comittee of Mathieu Laurière (Paris-Diderot, November 21), G. Carlier was on the Ph.D comittee of Guo (Ecole Polytechnique) and the HDR of Silva (Limoges, referee) and Lamboley (Dauphine, coordinateur). G. Peyré was in the PhD comittees of: Mitra Fatemi (EPFL, Feb. 2016), Morgane Henry (Grenoble, Mars 2016), Antoine Bonnefoy (Marseille, Mars 2016), Sébastien Combrexelle (Toulouse, Oct. 2016), Augustin Cosse (UCL, Aout 2016), Irène Kaltenmark (ENS Cachan, Oct. 2016), Olivia Miraucourt (Reims, Oct. 2016), Fred Maurice Ngole (CEA, Oct. 2016), Lara Raad (ENS Cachan, Oct. 2016), Emmanuel Soubies (Nice, Oct. 2016), Luc Le Magoarou (Rennes, Dec. 2016), Chen Da (Paris, Dec. 2016).

10.3. Popularization

G. Peyré wrote the large audience articles : "Claude Shannon et la compression de données" on Image des Mathematiques [0] ; "Parcimonie, problemes inverse et échantillonnage compressé" in La Gazette des Mathematiciens (SMF). G. Peyré organized the conference in the honnor of Claude Shannon in CIRM [0].

11. Bibliography

Major publications by the team in recent years

[1] M. AGUEH, G. CARLIER.*Barycenters in the Wasserstein space*, in "SIAM J. Math. Anal.", 2011, vol. 43, n° 2, p. 904–924, http://dx.doi.org/10.1137/100805741.

[2] J.-D. BENAMOU, Y. BRENIER.*A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, in "Numer. Math.", 2000, vol. 84, n° 3, p. 375–393, http://dx.doi.org/10.1007/s002110050002.

[3] J.-D. BENAMOU, G. CARLIER, M. CUTURI, L. NENNA, G. PEYRÉ.*Iterative Bregman Projections for Regularized Transportation Problems*, in "SIAM Journal on Scientific Computing", 2015, vol. 37, n° 2, p. A1111-A1138 [*DOI :* 10.1137/141000439], http://hal.archives-ouvertes.fr/hal-01096124.

[4] 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.

[0] http://images.math.cnrs.fr/Claude-Shannon-et-la-compression-des-donnees
[0] http://www.fr-cirm-math.fr/hommage-claude-shannon.html

[5] M. BRUVERIS, F.-X. VIALARD.*On Completeness of Groups of Diffeomorphisms*, in "ArXiv e-prints", March 2014.

[6] 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.

[7] F. GAY-BALMAZ, D. D. HOLM, D. M. MEIER, T. S. RATIU, F.-X. VIALARD.*Invariant Higher-Order Variational Problems*, in "Communications in Mathematical Physics", January 2012, vol. 309, p. 413-458, http://dx.doi.org/10.1007/s00220-011-1313-y.

[8] P. MACHADO MANHÃES DE CASTRO, Q. MÉRIGOT, B. THIBERT.*Intersection of paraboloids and application to Minkowski-type problems*, in "Numerische Mathematik", November 2015 [*DOI :* 10.1007/S00211-015-0780-Z], https://hal.archives-ouvertes.fr/hal-00952720.

[9] Q. MÉRIGOT.*A multiscale approach to optimal transport*, in "Computer Graphics Forum", 2011, vol. 30, n° 5, p. 1583–1592.

Publications of the year

Articles in International Peer-Reviewed Journal

[10] N. BONNEEL, G. PEYRÉ, M. CUTURI.*Wasserstein Barycentric Coordinates: Histogram Regression Using Optimal Transport*, in "ACM Transactions on Graphics", April 2016, vol. 35, n° 4 [*DOI :* 10.1145/2897824.2925918], https://hal.archives-ouvertes.fr/hal-01303148.

[11] G. CARLIER, G. PEYRÉ, J.-M. MIREBEAU, V. DUVAL.*A Γ-Convergence Result for the Upper Bound Limit Analysis of Plates*, in "ESAIM: Mathematical Modelling and Numerical Analysis", June 2016, vol. 50, n° 1, p. 215–235, https://hal.inria.fr/hal-01112226.

[12] A. CHAMBOLLE, V. DUVAL, G. PEYRÉ, C. POON.*Geometric properties of solutions to the total variation denoising problem*, in "Inverse Problems", October 2016, https://hal.archives-ouvertes.fr/hal-01323720.

[13] Q. DENOYELLE, V. DUVAL, G. PEYRÉ.*Support Recovery for Sparse Super-Resolution of Positive Measures*, in "Journal of Fourier Analysis and Applications", September 2016 [*DOI :* 10.1007/S00041-016-9502-X], https://hal.archives-ouvertes.fr/hal-01270184.

[14] P. MACHADO MANHÃES DE CASTRO, Q. MÉRIGOT, B. THIBERT.*Far-field reflector problem and intersection of paraboloids*, in "Numerische Mathematik", October 2016, vol. 134, n° 2, p. 389–411 [*DOI :* 10.1007/S00211-015-0780-Z], https://hal.archives-ouvertes.fr/hal-00952720.

[15] Q. MÉRIGOT, J.-M. MIREBEAU.*Minimal geodesics along volume preserving maps, through semi-discrete optimal transport*, in "SIAM Journal on Numerical Analysis", November 2016, vol. 54, n° 6, p. 3465–3492 [*DOI :* 10.1137/15M1017235], https://hal.archives-ouvertes.fr/hal-01152168.

[16] H. RAGUET, C. MONIER, L. FOUBERT, I. FEREZOU, Y. FREGNAC, G. PEYRÉ.*Spatially Structured Sparse Morphological Component Separation for Voltage-Sensitive Dye Optical Imaging*, in "Journal of Neuroscience Methods", 2016, vol. 257, p. 76-96, https://hal.archives-ouvertes.fr/hal-01200646.

[17] B. SCHMITZER.*A Sparse Multiscale Algorithm for Dense Optimal Transport*, in "Journal of Mathematical Imaging and Vision", April 2016 [*DOI :* 10.1007/s10851-016-0653-9], https://hal.archives-ouvertes.fr/hal-01385274.

[18] J. SOLOMON, G. PEYRÉ, V. G. KIM, S. SRA.*Entropic Metric Alignment for Correspondence Problems*, in "ACM Transactions on Graphics", June 2016, vol. 35, n^{o} 4, p. 72:1–72:13 [*DOI :* 10.1145/2897824.2925903], https://hal.archives-ouvertes.fr/hal-01305808.

[19] G. TARTAVEL, G. PEYRÉ, Y. GOUSSEAU.*Wasserstein Loss for Image Synthesis and Restoration*, in "SIAM Journal on Imaging Sciences", October 2016, vol. 9, n^{o} 4, p. 1726-1755, https://hal.archives-ouvertes.fr/hal-01292843.

International Conferences with Proceedings

[20] A. GENEVAY, M. CUTURI, G. PEYRÉ, F. BACH.*Stochastic Optimization for Large-scale Optimal Transport*, in "NIPS 2016 - Thirtieth Annual Conference on Neural Information Processing System", Barcelona, Spain, NIPS (editor), Proc. NIPS 2016, December 2016, https://hal.archives-ouvertes.fr/hal-01321664.

[21] G. PEYRÉ, M. CUTURI, J. SOLOMON.*Gromov-Wasserstein Averaging of Kernel and Distance Matrices*, in "ICML 2016", New-York, United States, Proc. 33rd International Conference on Machine Learning, June 2016, https://hal.archives-ouvertes.fr/hal-01322992.

Books or Proceedings Editing

[22] M. BERGOUNIOUX, J.-B. CAILLAU, T. HABERKORN, G. PEYRÉ, C. SCHNÖRR (editors). *Variational methods in imaging and geometric control*, Radon Series on Comput. and Applied Math., de Gruyter, December 2016, n^{o} 18, https://hal.archives-ouvertes.fr/hal-01315508.

Other Publications

[23] M. AGUEH, G. CARLIER.*Vers un théorème de la limite centrale dans l'espace de Wasserstein ?*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01422107.

[24] R. ANDREEV.*Algorithm based fault tolerance with wavelets*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01359755.

[25] R. ANDREEV.*Jumplets*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01338101.

[26] R. ANDREEV.*Learning stochastic eigenvalues*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01313404.

[27] R. ANDREEV.*Preconditioning the augmented Lagrangian method for instationary mean field games with diffusion*, April 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01301282.

[28] R. ANDREEV.*Quasi-optimality of approximate solutions in normed vector spaces*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01338040.

[29] R. ANDREEV, K. KIRCHNER.*Numerical methods for the 2nd moment of stochastic ODEs*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01394195.

[30] J.-D. BENAMOU, G. CARLIER, R. HATCHI.*A numerical solution to Monge's problem with a Finsler distance as cost*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01261094.

[31] J.-D. BENAMOU, G. CARLIER, F. SANTAMBROGIO.*Variational Mean Field Games*, March 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01295299.

[32] A. BLANCHET, G. CARLIER, L. NENNA.*Computation of Cournot-Nash equilibria by entropic regularization*, September 2016, working paper or preprint, https://hal.inria.fr/hal-01363468.

[33] M. BRUVERIS, F.-X. VIALARD.*On Completeness of Groups of Diffeomorphisms*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01253261.

[34] G. CARLIER, A. GALICHON, V. CHERNOZHUKOV.*Vector quantile regression beyond correct specification*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01420712.

[35] G. CARLIER, M. LABORDE.*A splitting method for nonlinear diffusions with nonlocal, nonpotential drifts*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01332356.

[36] G. CARLIER, L. MALLOZZI.*Optimal monopoly pricing with congestion and random utility via partial mass transport*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01420707.

[37] L. CHIZAT, G. PEYRÉ, B. SCHMITZER, F.-X. VIALARD.*Scaling Algorithms for Unbalanced Transport Problems*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01434914.

[38] C. DOSSAL, V. DUVAL, C. POON.*Sampling the Fourier transform along radial lines*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01421265.

[39] V. DUVAL, G. PEYRÉ.*Sparse Spikes Super-resolution on Thin Grids I: the LASSO*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01135200.

[40] V. DUVAL, G. PEYRÉ.*Sparse Spikes Super-resolution on Thin Grids II: the Continuous Basis Pursuit*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01389956.

[41] T. O. GALLOUËT, Q. MÉRIGOT.*A Lagrangian scheme for the incompressible Euler equation using optimal transport*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01425826.

[42] T. GALLOUËT, F.-X. VIALARD.*From unbalanced optimal transport to the Camassa-Holm equation*, December 2016, Comments welcome, 28 pages, https://hal.archives-ouvertes.fr/hal-01363647.

[43] J. KITAGAWA, Q. MÉRIGOT, B. THIBERT.*Convergence of a Newton algorithm for semi-discrete optimal transport*, March 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01290496.

[44] J. LOUET, A. PRATELLI, F. ZEISLER.*On the continuity of the total cost in the mass transport problem with relativistic cost functions*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01419526.

[45] G. PEYRÉ.*Claude Shannon et la compression des données*, July 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01343890.

[46] G. PEYRÉ.*Parcimonie, problèmes inverses et échantillonnage compressé*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01331580.

[47] B. SCHMITZER.*Stabilized Sparse Scaling Algorithms for Entropy Regularized Transport Problems*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01385251.

[48] R. TAHRAOUI, F.-X. VIALARD.*Riemannian cubics on the group of diffeomorphisms and the Fisher-Rao metric*, September 2016, 34 pages, comments welcome, https://hal.archives-ouvertes.fr/hal-01331110.

References in notes

[49] 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.

[50] 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.

[51] M. AGUEH, G. CARLIER.*Barycenters in the Wasserstein space*, in "SIAM J. Math. Anal.", 2011, vol. 43, n^o 2, p. 904–924, http://dx.doi.org/10.1137/100805741.

[52] 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.

[53] 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.

[54] 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.

[55] M. BATES.*Models of natural language understanding*, in "Proceedings of the National Academy of Sciences", 1995, vol. 92, n^o 22, p. 9977-9982.

[56] 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.

[57] 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.

[58] M. BEIGLBOCK, P. HENRY-LABORDÈRRE, 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.

[59] 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.

[60] 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.

[61] 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, http://dx.doi.org/10.1137/S0036139995294111.

[62] J.-D. BENAMOU, G. CARLIER. *Augmented Lagrangian algorithms for variational problems with divergence constraints*, in "JOTA", 2015.

[63] 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.

[64] 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.

[65] J.-D. BENAMOU, G. CARLIER, Q. MÉRIGOT, E. OUDET. *Discretization of functionals involving the Monge-Ampère operator*, HAL, July 2014, https://hal.archives-ouvertes.fr/hal-01056452.

[66] 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.

[67] 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, http://dx.doi.org/10.1051/m2an/2010017.

[68] 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.

[69] 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, http://dx.doi.org/10.3934/nhm.2009.4.605.

[70] M. BENNING, M. BURGER. *Ground states and singular vectors of convex variational regularization methods*, in "Meth. Appl. Analysis", 2013, vol. 20, p. 295–334.

[71] B. BERKELS, A. EFFLAND, M. RUMPF. *Time discrete geodesic paths in the space of images*, in "Arxiv preprint", 2014.

[72] J. BIGOT, T. KLEIN. *Consistent estimation of a population barycenter in the Wasserstein space*, in "Preprint arXiv:1212.2562", 2012.

[73] A. BLANCHET, G. CARLIER. *Optimal Transport and Cournot-Nash Equilibria*, in "Mathematics of Operations Resarch", 2015, to appear.

[74] 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^o 4, p. 658–686, http://dx.doi.org/10.1080/03605302.2012.757705.

[75] J. BLEYER, G. CARLIER, V. DUVAL, J.-M. MIREBEAU, G. PEYRÉ.*A Γ-Convergence Result for the Upper Bound Limit Analysis of Plates*, in "arXiv preprint arXiv:1410.0326", 2014.

[76] 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° 1, p. 22–45, http://hal.archives-ouvertes.fr/hal-00881872/.

[77] 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/.

[78] G. BOUCHITTÉ, G. BUTTAZZO.*Characterization of optimal shapes and masses through Monge-Kantorovich equation*, in "J. Eur. Math. Soc. (JEMS)", 2001, vol. 3, n° 2, p. 139–168, http://dx.doi.org/10.1007/s100970000027.

[79] 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° 6, p. 652–671, http://dx.doi.org/10.1016/j.matpur.2010.03.010.

[80] 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° 3, p. 200–217.

[81] Y. BRENIER.*Generalized solutions and hydrostatic approximation of the Euler equations*, in "Phys. D", 2008, vol. 237, n° 14-17, p. 1982–1988, http://dx.doi.org/10.1016/j.physd.2008.02.026.

[82] Y. BRENIER.*Décomposition polaire et réarrangement monotone des champs de vecteurs*, in "C. R. Acad. Sci. Paris Sér. I Math.", 1987, vol. 305, n° 19, p. 805–808.

[83] Y. BRENIER.*Polar factorization and monotone rearrangement of vector-valued functions*, in "Comm. Pure Appl. Math.", 1991, vol. 44, n° 4, p. 375–417, http://dx.doi.org/10.1002/cpa.3160440402.

[84] Y. BRENIER, U. FRISCH, M. HENON, G. LOEPER, S. MATARRESE, R. MOHAYAEE, A. SOBOLEVSKII.*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.

[85] 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° 4, p. 1344-1368, http://dx.doi.org/10.1137/110846324.

[86] M. BURGER, M. DIFRANCESCO, P. MARKOWICH, M. T. WOLFRAM.*Mean field games with nonlinear mobilities in pedestrian dynamics*, in "DCDS B", 2014, vol. 19.

[87] 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.

[88] M. BURGER, S. OSHER.*A guide to the TV zoo*, in "Level-Set and PDE-based Reconstruction Methods, Springer", 2013.

[89] 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, http://dx.doi.org/10.1137/07070543X.

[90] 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.

[91] 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.

[92] L. CAFFARELLI, S. 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.

[93] C. CANCERITOGLU.*Computational Analysis of LDDMM for Brain Mapping*, in "Frontiers in Neuroscience", 2013, vol. 7.

[94] E. CANDES, M. WAKIN.*An Introduction to Compressive Sensing*, in "IEEE Signal Processing Magazine", 2008, vol. 25, n^o 2, p. 21–30.

[95] 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.

[96] 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.

[97] 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, http://dx.doi.org/10.1007/s00526-012-0555-7.

[98] 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, http://dx.doi.org/10.1016/S0304-4068(00)00057-4.

[99] G. CARLIER, V. CHERNOZHUKOV, A. GALICHON.*Vector Quantile Regression*, Arxiv 1406.4643, 2014.

[100] 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/.

[101] G. CARLIER, X. DUPUIS.*An iterated projection approach to variational problems under generalized convexity constraints and applications*, In preparation, 2015.

[102] G. CARLIER, I. EKELAND.*Matching for teams*, in "Econom. Theory", 2010, vol. 42, n^o 2, p. 397–418, http://dx.doi.org/10.1007/s00199-008-0415-z.

[103] 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, http://dx.doi.org/10.1137/060672832.

[104] 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.

[105] G. CARLIER, A. OBERMAN, É. OUDET.*Numerical methods for matching for teams and Wasserstein barycenters*, in "M2AN", 2015, to appear.

[106] G. CARLIER, F. SANTAMBROGIO.*A continuous theory of traffic congestion and Wardrop equilibria*, in "Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)", 2011, vol. 390, n^{o} Teoriya Predstavlenii, Dinamicheskie Sistemy, Kombinatornye Metody. XX, p. 69–91, 307–308, http://dx.doi.org/10.1007/s10958-012-0715-5.

[107] 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.

[108] 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.

[109] 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.

[110] A. CHAMBOLLE, T. POCK.*On the ergodic convergence rates of a first-order primal-dual algorithm*, in "Preprint OO/2014/09/4532", 2014.

[111] 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/.

[112] 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.

[113] 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.

[114] 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.

[115] 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.

[116] 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, http://dx.doi.org/10.1137/0151002.

[117] M. CUTURI, D. AVIS.*Ground Metric Learning*, in "J. Mach. Learn. Res.", January 2014, vol. 15, n^{o} 1, p. 533–564, http://dl.acm.org/citation.cfm?id=2627435.2627452.

[118] 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.

[119] 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.

[120] 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.

[121] V. DUVAL, G. PEYRÉ.*Sparse Spikes Deconvolution on Thin Grids*, HAL, 2015, n^{o} 01135200, http://hal.archives-ouvertes.fr/hal-01135200.

[122] 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.

[123] 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.

[124] A. FIGALLI, R. MC CANN, Y. KIM.*When is multi-dimensional screening a convex program?*, in "Journal of Economic Theory", 2011.

[125] 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.

[126] 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.

[127] 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.

[128] 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, http://dx.doi.org/10.1137/120875065.

[129] 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.

[130] 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.

[131] 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.

[132] 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.

[133] T. HASTIE, R. TIBSHIRANI, J. FRIEDMAN.*The Elements of Statistical Learning*, Springer Series in Statistics, Springer New York Inc., New York, NY, USA, 2001.

[134] G. HERMAN.*Image reconstruction from projections: the fundamentals of computerized tomography*, Academic Press, 1980.

[135] D. D. HOLM, J. T. RATNANATHER, A. TROUVÉ, L. YOUNES.*Soliton dynamics in computational anatomy*, in "NeuroImage", 2004, vol. 23, p. S170–S178.

[136] 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.

[137] 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, http://dx.doi.org/10.1137/S0036141096303359.

[138] 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.

[139] L. KANTOROVITCH.*On the translocation of masses*, in "C. R. (Doklady) Acad. Sci. URSS (N.S.)", 1942, vol. 37, p. 199–201.

[140] 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.

[141] 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.

[142] 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.

[143] 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, http://dx.doi.org/10.1137/140975528.

[144] A. S. LEWIS.*Active sets, nonsmoothness, and sensitivity*, in "SIAM Journal on Optimization", 2003, vol. 13, n^{o} 3, p. 702–725.

[145] 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.

[146] 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), http://dx.doi.org/10.1137/050629070.

[147] 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.

[148] D. LOMBARDI, E. MAITRE.*Eulerian models and algorithms for unbalanced optimal transport*, in "Preprint hal-00976501", 2013.

[149] 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.

[150] 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.

[151] S. G. MALLAT.*A wavelet tour of signal processing*, Third, Elsevier/Academic Press, Amsterdam, 2009.

[152] 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.

[153] 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.

[154] J.-M. MIREBEAU.*Adaptive, Anisotropic and Hierarchical cones of Discrete Convex functions*, in "Preprint", 2014.

[155] 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.

[156] Q. MÉRIGOT.*A multiscale approach to optimal transport*, in "Computer Graphics Forum", 2011, vol. 30, n^o^ 5, p. 1583–1592.

[157] 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, http://dx.doi.org/10.1137/130938359.

[158] N. PAPADAKIS, G. PEYRÉ, E. 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/.

[159] B. PASS, N. GHOUSSOUB.*Optimal transport: From moving soil to same-sex marriage*, in "CMS Notes", 2013, vol. 45, p. 14–15.

[160] 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, http://dx.doi.org/10.1137/100804917.

[161] J. PENNINGTON, R. SOCHER, C. MANNING.*Glove: Global Vectors for Word Representation*, in "Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)", Association for Computational Linguistics, 2014, p. 1532–1543.

[162] 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.

[163] J. PETITOT.*The neurogeometry of pinwheels as a sub-riemannian contact structure*, in "Journal of Physiology-Paris", 2003, vol. 97, n^{o} 23, p. 265–309.

[164] G. PEYRÉ.*Texture Synthesis with Grouplets*, in "Pattern Analysis and Machine Intelligence, IEEE Transactions on", April 2010, vol. 32, n^{o} 4, p. 733–746.

[165] 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^{o} 1, p. 335–358.

[166] C. POON.*Structure dependent sampling in compressed sensing: theoretical guarantees for tight frames*, in "Applied and Computational Harmonic Analysis", 2015.

[167] C. PRINS, J.H.M. TEN. THIJE BOONKKAMP, J. VAN . ROOSMALEN, W.L. IJZERMAN, T.W. TUKKER.*A numerical method for the design of free-form reflectors for lighting applications*, in "External Report, CASA Report, No. 13-22", 2013, http://www.win.tue.nl/analysis/reports/rana13-22.pdf.

[168] H. RAGUET, J. FADILI, G. PEYRÉ.*A Generalized Forward-Backward Splitting*, in "SIAM Journal on Imaging Sciences", 2013, vol. 6, n^{o} 3, p. 1199–1226 [*DOI :* 10.1137/120872802], http://hal.archives-ouvertes.fr/hal-00613637/.

[169] J.-C. ROCHET, P. CHONÉ.*Ironing, Sweeping and multi-dimensional screening*, in "Econometrica", 1998.

[170] J. RUBINSTEIN, G. WOLANSKY.*Intensity control with a free-form lens*, in "J Opt Soc Am A Opt Image Sci Vis.", 2007, vol. 24.

[171] L. RUDIN, S. OSHER, E. FATEMI.*Nonlinear total variation based noise removal algorithms*, in "Physica D: Nonlinear Phenomena", 1992, vol. 60, n^{o} 1, p. 259–268, http://dx.doi.org/10.1016/0167-2789(92)90242-F.

[172] O. SCHERZER, M. GRASMAIR, H. GROSSAUER, M. HALTMEIER, F. LENZEN.*Variational Methods in Imaging*, Springer, 2008.

[173] 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.

[174] 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.

[175] B. SCHÖLKOPF, A. J. SMOLA.*Learning with kernels : support vector machines, regularization, optimization, and beyond*, Adaptive computation and machine learning, MIT Press, 2002, http://www.worldcat.org/oclc/48970254.

[176] 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.

[177] R. TIBSHIRANI.*Regression shrinkage and selection via the Lasso*, in "Journal of the Royal Statistical Society. Series B. Methodological", 1996, vol. 58, n^{o} 1, p. 267–288.

[178] A. TROUVÉ, F.-X. VIALARD.*Shape splines and stochastic shape evolutions: A second order point of view*, in "Quarterly of Applied Mathematics", 2012.

[179] 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/.

[180] 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^{o} 2, p. 229-241, http://dx.doi.org/10.1007/s11263-011-0481-8.

[181] 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, http://dx.doi.org/10.1007/978-3-319-10404-1_29.

[182] C. VILLANI.*Topics in optimal transportation*, Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003, vol. 58, xvi+370.

[183] 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.

[184] X.-J. WANG.*On the design of a reflector antenna. II*, in "Calc. Var. Partial Differential Equations", 2004, vol. 20, n^{o} 3, p. 329–341, http://dx.doi.org/10.1007/s00526-003-0239-4.

[185] 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^{o} 3, p. 293–318.

[186] 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^{o} 6, p. 1031–1044.

Team MUSE

Measuring networks for enhancing USer Experience

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
Networks and Telecommunications

Table of contents

Team MUSE

Creation of the Team: 2013 October 03

Keywords:

Computer Science and Digital Science:

1.2.3. - Routing
1.2.4. - QoS, performance evaluation
1.2.5. - Internet of things
3.1.7. - Open data
3.1.8. - Big data (production, storage, transfer)
3.3. - Data and knowledge analysis
3.5. - Social networks

Other Research Topics and Application Domains:

6.3. - Network functions
6.4. - Internet of things
6.5. - Information systems

1. Members

Research Scientists

Renata Cruz Teixeira [Team leader, Inria, Senior Researcher, HDR]
Vassilis Christophides [Inria, Advanced Research position]
Anna-Kaisa Pietilainen [Inria, Starting Research position, until Sep 2016]

Faculty Member

Timur Friedman [Univ. Paris VI, Associate Professor, until Aug 2016]

Technical Staff

Michail Katsarakis [Inria, University of Crete, until May 2016]

PhD Students

Maximilian Bachl [Univ. Paris VI, from Oct 2016]
Sara El Aouad [Inria/Technicolor, granted by CIFRE, from May 2014]
Diego Neves Da Hora [Inria/Technicolor, granted by CIFRE, from February 2014]
Giuseppe Scavo [Inria, granted by Alcatel-Lucent Bell Labs, from November 2013]

Post-Doctoral Fellow

Francesco Bronzino [Inria, from Dec 2016]

Administrative Assistant

Anna Bednarik [Inria]

Others

Arash Molavi Kakhki [Inria, Ph.D Student, Northeastern University, from Nov 2016]
Julio Adriazola Soto [Inria, Inria Chile, from Jan 2016 until Mar 2016]
Michele Pittoni [Inria, Univ. Paris VI, M2 Student, from Feb 2016 until Jul 2016]
Adhir Chakravarti [Inria, Univ. Paris VI, M2 Student, from Feb 2016 until Jul 2016]
Zied Ben Houidi [Bell Labs (Alcatel), from Apr 2016]

2. Overall Objectives

2.1. Overall Objectives

Muse's research is broadly in the area of network measurements. We focus on developing new algorithms and systems to improve user experience online. In particular, we are addressing two main problems of today's Internet users:

1. 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.
2. 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.

3. Research Program

3.1. Active 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 outside. Our goal is to develop adaptative 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).

3.2. Passive monitoring methods

This part our research develops 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.

3.3. 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. We will 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.

3.4. Filtering real-time Web streams

The Web has become a large-scale real-time information system forcing us to revise both how to effectively assess relevance of information for a user and how to efficiently implement information retrieval and dissemination functionality. To increase information relevance, Real-time Web applications such as Twitter and Facebook, extend content and social-graph relevance scores with "real-time" user generated events (e.g. re-tweets, replies, likes). To accommodate high arrival rates of information items and user events we explore a publish/subscribe paradigm in which we index queries and update on the fly their results each time a new item and relevant events arrive. In this setting, we need to process continuous top-k text queries combining both static and dynamic scores. To the best of our knowledge, this is the first work addressing how non-predictable, dynamic scores can be handled in a continuous top-k query setting.

3.5. Flexible online drift detection

Monitoring streaming content is a challenging big data analytics problem, given that very large datasets are rarely (if ever) stationary. In several real world monitoring applications (e.g., newsgroup discussions, network connections, etc.) we need to detect significant change points in the underlying data distribution (e.g., frequency of words, sessions, etc.) and track the evolution of those changes over time. These change points, depending on the research community, are referred to as temporal evolution, non-stationarity, or concept drift and provide valuable insights on real world events (e.g. a discussion topic, an intrusion) to take a timely action. In our work, we adopt a query-based approach to drift detection and address the question of processing drift queries over very large datasets. To the best of our knowledge, our work is the first to formalize flexible drift queries on streaming datasets with varying change rates.

4. Application Domains

4.1. 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 "Interne 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 causes of potential problems and—ideally—ways to fix them. We intend to develop a set of easy-to-use home network diagnosis tools that can reliably identify performance and functionality shortcomings rooted in the home. The development of home network 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, 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. Third, 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.

We are conceiving methods for two application scenarios: (i) when the end user in the home deploys our diagnostic tools either on the home gateway (the gateway often combines a DSL/cable modem and an access point; it connects the home network to the ISP) or on devices connected to the home network and (ii) when ISPs collect measurements from homes of subscribers and then correlate these measurements to help identify problems.

Assisting end users. We are developing algorithms to determine whether network performance problems lie inside or outside the home network. Given that the home gateway connects the home with the rest of the Internet, we are designing an algorithm (called *HoA*) that analyzes traffic that traverses the gateway to distinguish access link and home network bottlenecks. A measurement vantage point on the gateway is key for determining if the performance bottleneck lies within the home network or the access ISP, but we also need to deploy diagnosis tools in end-devices. First, some users may not want (or not know how) to deploy a new home gateway in their homes. Second, some problems will be hard to diagnose with only the vantage point of the gateway (for example, when a device cannot send traffic or when the wireless is poor in certain locations of a home). We can obtain more complete visibility by leveraging *multiple* measurement nodes around the home, potentially including the home gateway, all participating jointly in the measurement task. We have an ongoing project to realize a home network analyzer as a web-based measurement application built on top of our team's recently developed browser-based measurement platform, *Fathom*. To integrate the home gateway in the analyzer, we plan to engage the BISmark Project. BISmark already provides a web server as well as extensive configurability, allowing us to experiment freely with both passive as well as active measurements. We must develop a home network analyzer that can first discover the set of devices connected to the home network that can collaborate on the diagnosis task. We will then develop tomography algorithms to infer where performance problems lie given measurements taken from the set of available vantage points.

Assisting Internet Service Providers (ISPs). Our discussions with several large access ISPs reveal that service calls are costly, ranging from $9–25 per call, and as many as 75% of service calls from customers are usually caused by problems that have nothing to do with the ISP. Therefore, ISPs are eager to deploy techniques to assist in home network diagnosis. In many countries ISPs control the home gateway and set-top-boxes in the home. We plan to develop more efficient mechanisms for home users to report trouble to their home ISP and consequently reduce the cost of service calls. This project is in collaboration with Technicolor and Portugal Telecom. Technicolor is a large manufacturer of home gateways and set-top-boxes. Portugal Telecom is the largest broadband access provider in Portugal. Technicolor already collects data from 200 homes in Portugal. We are working with the data collected in this deployment together with controlled experiments to develop methods to diagnose problems in the home wireless.

4.2. Quality of Experience

An increasing number of residential users consume online services (e.g., VoD, Web browsing, or Skype) in their everyday activities (e.g., for education or entertainment purposes), using a variety of devices (e.g., tablets, smartphones, laptops). A high Quality of Service (QoS) is essential for sustaining the revenue of service providers, carriers, and device manufactures. Yet, the perceived Quality of Experience (QoE) of users is far from perfect e.g., videos that get stalled or that take a long time to load. Dissatisfied users may change Internet Service Providers (ISPs) or the online services. Hence, the incentives for measuring and improving QoE in home networks are high while mapping network and application QoS to QoE is a challenging problem. In this work we have focused in measuring several network Quality-of-Service (QoS) metrics, such as latency and bandwidth, both in residential Wi-Fi as well as broadband networks, homes are using for connecting to the Internet.

The WiFi Context. Residential Wi-Fi performance, however, is highly variable. Competing Wi-Fi networks can cause contention and interference while poor channel conditions between the station and the access point (AP) can cause frame losses and low bandwidth. In some cases, the home Wi-Fi network can bottleneck Internet access. While problems in the Wi-Fi network may affect several network QoS metrics, users will typically only notice a problem when poor Wi-Fi affects the QoE of Internet applications. For example, a Wi-Fi network with low bandwidth may go unnoticed unless the time to load Web pages increases significantly. A user observing degraded QoE due to Wi-Fi problems may mistakenly assume there is a problem with the Internet Service Provider (ISP) network. Our discussions with residential ISPs confirm that often customers call to complain about problems in the home Wi-Fi and not the ISP network.

Prior work has focused on QoS metrics for some applications (e.g., on-line video, Web browsing, or Skype) with no attempt to identify when Wi-Fi quality affects QoE. We are particularly interested in assisting ISPs to predict when home Wi-Fi quality degrades QoE. ISPs can use this system to detect customers experiencing poor QoE to proactively trigger Wi-Fi troubleshooting. ISPs often control the home AP, so we leverage Wi-Fi metrics that are available on commercial APs. Detecting when Wi-Fi quality degrades QoE using these metrics is challenging. First, we have no information about the applications customers are running at any given time. ISPs avoid capturing per-packet traffic traces from customers, because of privacy considerations and the overload of per-packet capture. Thus, we must estimate the effect of Wi-Fi quality on QoE of popular applications, which most customers are likely to run. In this context, we study Web as a proof of concept, as a large fraction of home traffic corresponds to Web. Second, application QoE may be degraded by factors other than the Wi-Fi quality (e.g., poor Internet performance or an overloaded server). Although a general system to explain any QoE degradation would be extremely helpful, our monitoring at the AP prevents us from having the end-to-end view necessary for such general task. Instead, we focus on identifying when Wi-Fi quality degrades QoE. Finally, Wi-Fi metrics available in APs are coarse aggregates such as the average PHY rate or the fraction of busy times. It is open how to effectively map these coarse metrics into QoE.

Predicting QoE. Clearly, different actors in the online service chain (e.g., video streaming services, ISPs) have different incentives and means to measure and affect the user QoE. Uncovering statistically equivalent subsets of QoS metrics across and within levels provides actionable knowledge for building QoE predictors. To achieve this goal, we leverage recent advances on feature selection algorithms to exploit available experimental evidence of the joint probability distributions of QoE/QoS metrics. This type of statistical reasoning will enable us to determine local causal relationships between a target QoE variable, seen as effect, and multiple QoS metrics across or within levels, seen as causes. Such data-driven analysis is justified by the multiplicity of dependencies that exist between network or application QoS metrics as different adaptation mechanisms (e.g., TCP congestion avoidance, HTTP bitrate adaptation) are activated at each level in real life. Building optimal predictors based on (eventually several) probabilistically minimal subsets of features opens the way for a principled comparison of the predictors.

4.3. Data Analytics for the Internet of Things

The Internet of Things (IoT) is rapidly transforming the physical world into a large scale information system. A wave of smart "things" smoothly disappear in our environment (aka *Pervasive Computing*), or be embodied in humans (aka *Wearable Computing*, and continuously produce valuable information regarding almost every living context and process. *Making sense of the data streams "things" produce and share* is crucial for disruptive IoT applications. From smart devices and homes, to smart roads and cities, IoT data analytics is expected to enable a resource-conscious automation of our everyday life in terms of operational efficiency, security, safety as well as of a lower energy footprint.

Multi-dimensional Usage Patterns. We have initially investigated how data analytics for Machine-to-Machine (M2M) data (connectivity, performance, usage) produced by connected devices in residential Intranet of Things, could support novel *home automation services* that enrich the living experience in smart homes. We have investigated new data mining techniques that go beyond binary association rule mining for traditional market basket analysis, considered by previous works. We design a multidimensional pattern mining framework, which collects raw data from operational home gateways, it discretizes and annotates the raw data, it produces traffic usage logs which are fed in a multidimensional association rule miner, and finally it extracts home residents habits. Using our analysis engine, we extract complex device co-usage patterns of 201 residential broadband users of an ISP, subscribed to a n-play service. Such fine-grained device usage patterns provide valuable insights for emerging use cases, such as adaptive usage of home devices (aka horizontal integration of things). Such use cases fall within the wider area of human-cognizant Machine-to-Machine communication aiming to predict user needs and complete tasks without users initiating the action or interfering with the service. While this is not a new concept, according to Gartner cognizant computing is a natural evolution of a world driven not by devices but collections of applications and services that span across multiple devices, in which human intervention becomes as little as possible, by analyzing past human habits. To realize this vision, we are interested in co-usage patterns featuring spatio-temporal information regarding the context under which devices have been actually used in homes. For example, a network extender which is currently turned off, could be turned on at a certain day period (e.g., evening) when it has been observed to be highly used along with other devices (e.g., a laptop or a tablet). Alternatively, the identification of frequent co-usage of particular devices at a home (say iPhone with media player), could be used by a things recommender to advertise the same set of devices at another home (say another iPhone user could be interested in a media player).

Time Series Motif. Furthermore, we are interested in extracting previously unknown recurring patterns (aka motifs) directly from traffic time series reported by residential gateways. Such motifs could help ISPs to reduce the cost for *serving and diagnosing remotely home networks*, or even help assist in *defining home-specific bandwidth sharing and prioritization policies*. More precisely, traffic motifs enriched with detailed home device information is a valuable input for root cause diagnosis and can be contrasted to the trouble description reported by users to the ISP. Moreover, in their majority, ISPs typically broadcast firmware and software updates to all gateways at nights (some operators even on a daily basis). This may cause service outages, given that some gateways may exhibit an active network usage during night time. A fine-grained temporal characterization of residential bandwidth consumption will enable ISPs to differentiate RGWs firmware update policies according to the least cumbersome time window per home, thus, improving the overall QoE of residential users. Finally, home network resources (bandwidth) are shared not only among residents using an increasing number of on-line applications (e.g., social networking, gaming, uploading/downloading, etc.) and real time services (TV on-demand, teleconferencing), but also with guests, neighbors, or even the occasional passes by. Existing methods for bandwidth sharing and traffic prioritization are static and coarse. ISPs usually allocate a fixed percentage of home bandwidth to non-residential users, while traffic prioritization in commodity gateways is at best based on the network port on which traffic is sent or received. We believe that behavioural patterns extracted by gateway traffic time series can be used to support dynamic policies for sharing home bandwidth that consider the online habits of residential users. For example, in-home traffic congestion can be avoided by ordering the traffic patterns of different devices observed especially during afternoon and weekends. These patterns reveal the bandwidth consumption behavior of different groups of residential users (adults and children employ different devices during the same time-slots) while the comparison of traffic domination help us to distinguish between residents and guests (pattern-specific vs global traffic dominant devices).

4.4. Crowd-sourced Information Filtering and Summarization

With the explosion of the People-centric Web, there is a proliferation of crowd-sourced content either under the form of qualitative reviews (mainly textual) and quantitative ratings (as 5 star ratings) regarding diverse products or services or under the form of various "real-time" feedback events (e.g., re-tweets, replies, likes, clicks, etc.) on published web content (ranging from traditional news, TV series, and movies to specialized blogs and posts shared over social networks). Such content captures the wisdom of the crowd and is valuable information source for building collaborative filtering systems and text summarization tools coping with information overload. For example, they can assist users to pick the most interesting web pages (e.g. Delicious) or to choose which movie to watch next (e.g. Netflix).

Implicit Feedback in Communities of a Place. We are initially interested in addressing one of the main limitation of collaborative filtering systems namely, the strong user engagement required to provide the necessary input (e.g., regarding their friends, tags or sites of preference) which is usual platform specific (i.e., for a particular social network, tagging, or bookmark system). The lack of user engagement translates into cold start and data sparsity. To cope with this limitation, we are developing a system called WeBrowse that passively observes network traffic to extract user clicks (i.e., the URLs users visit) for group of people who live, study, or work in the same place. Examples of such communities of a place are: (i) the students of a campus, (ii) the people living in a neighbourhood or (iii) researchers working in the same site. WeBrowse then promotes the hottest and most popular content to the community members sharing common interests.

Personalized Review Summarization. Finally, we are interested in helping people to take informed decisions regarding their shopping or entertainment activities. The automated summarization of a review corpus (for example, movie reviews from Rotten Tomatoes or IMDB; or restaurant reviews from Yelp) aims to assist people to form an opinion regarding a product/service of interest, by producing a coherent summary that is helpful and can be easily assimilated by humans. We are working on review summarisation methods that combine both objective (i.e., related to the review corpus) and subjective (i.e., related to the end-user interests) interestingness criteria of the produced reviews. In this respect we are exploiting domain models (e.g., Oscar's merit categories for movies) to elicit user preferences and mine the aspects of products/services actually commented in the textual sentences of reviews. For example, different summaries should be produced when a user is more interested in the actors performance rather than the movie story. We are particularly interested in extracting automatically the signatures of aspects (based on a set of seed terms) and rank review sentences on their importance and relevance w.r.t. the aspects they comment. Last but not least we are optimizing the automatically constructed summary w.r.t. to a number of criteria such as the number of the length of included sentences from the original reviews, the polarity of sentiments in the described aspects, etc.

5. New Software and Platforms

5.1. Fathom

Fathom - browser-based network measurement platform
KEYWORDS: Internet access - Performance measure - Network monitoring
FUNCTIONAL DESCRIPTION

Fathom is a Firefox browser extension that explores the browser as a platform for network measurement and troubleshooting. It provides a wide range of networking primitives directly to in-page JavaScript including raw TCP/UDP sockets, higher-level protocol APIs such as DNS, HTTP, and UPnP, and ready-made functionality such as pings and traceroutes.

- Participants: Anna-Kaisa Pietilainen and Stephane Archer
- Contact: Anna-Kaisa Pietilainen
- URL: https://muse.inria.fr/fathom/

5.2. HostView

FUNCTIONAL DESCRIPTION

End-host performance monitoring and user feedback reporting

- Participants: George Rosca, Anna-Kaisa Pietilainen and Renata Cruz Teixeira
- Contact: Renata Cruz Teixeira
- URL: https://team.inria.fr/muse/

5.3. Online HoA

Online implementation of home and access throughput bottleneck detection algorithm 'HoA'
FUNCTIONAL DESCRIPTION

"Home or Access" (HoA) is a system that localizes performance problems in home and access networks. Originally, we implement HoA as custom firmware that collect traces from off-the-shelf home routers. HoA uses timing and buffering information from passively monitored traffic at home routers to detect both access link and wireless network bottlenecks. HoA runs offline on a server to locate last-mile downstream throughput bottlenecks based on the analysis of packet traces collected from home routers. Our attempts to run HoA online on commodity home routers, however, revealed the challenges with performing per-packet analysis on such resource-constrained devices. The online HoA resolves this issue. We design an access bottleneck detector based on lightweight pings of the access link, and a wireless bottleneck detector based on a model of wireless capacity using metrics that are easily available in commodity home routers such as the wireless physical rate and the count of packets/bytes transmitted.

- Contact: Renata Cruz Teixeira
- URL: https://github.com/inria-muse/browserlab

5.4. SimilarityExplanation

Prototype implementation for explaining a set of similar and recommended movies.
FUNCTIONAL DESCRIPTION

In this web-based prototype for similar movies explanation, we propose two types of browsing for: personalized browsing and non personalized browsing. In the non personalized browsing we suppose that we don't have the user profile. Similar movie sublists are ordered only according to their similarity to the selected movie. For the personalized browsing , we select users that have different profiles from our dataset. We give these users names of actors, according to the types of movies they watch. For each user, we compute the predicted ratings using the matrix factorization model. We select pairs of genres to display to each user based on the preferred genres for the user. In our prototype we identify the preferred genres per user based on the most frequent movie genre pairs that the user has already seen. We then organize the recommended movies with a high rating prediction in sublists, according to the user most preferred genre pairs. When a user selects a movie from the sublists of recommended movies, our application suggests the similar movies presented under four sublists with the added list of words. The sublists are personalized for each user by reordering the movies according to the users predicted ratings.

- Contact: Sara El Aouad
- URL: https://team.inria.fr/muse/

5.5. UCN

User-Centric Networking
FUNCTIONAL DESCRIPTION

The User-Centric Networking (UCN) project is seeking to understand how people consume various kinds of content when using computer networks. Within this project we are undertaking a detailed user study across a range of environments in order to understand the practices involved in consuming media and other content according to context.

- Participants: Renata Cruz Teixeira and Anna-Kaisa Pietilainen
- Contact: Anna-Kaisa Pietilainen
- URL: https://team.inria.fr/muse/

5.6. WeBrowse

FUNCTIONAL DESCRIPTION

WeBrowse is the first passive crowdsource-based content curation system. Content curation is the act of assisting users to identify relevant and interesting content in the Internet. WeBrowse requires no active user engagement to promote content. Instead, it extracts the URLs users visit from traffic traversing an ISP network to identify popular content. WeBrowse contains a set of heuristics to identify the set of URLs users visit and to select the subset that are interesting to users.

- Contact: Giuseppe Scavo
- URL: http://webrowse.polito.it/

6. New Results

6.1. Home Network or Access Link? Locating Last-mile Downstream Throughput Bottlenecks

Participants: Srikanth Sundaresan (ICSI), Nick Feamster (Princeton), Renata Teixeira

As home networks see increasingly faster downstream throughput speeds, a natural question is whether users are benefiting from these faster speeds or simply facing performance bottlenecks in their own home networks. We studied the problem whether downstream throughput bottlenecks occur more frequently in their home networks or in their access ISPs. We identified lightweight metrics that can accurately identify whether a throughput bottleneck lies inside or outside a user's home network and developed a detection algorithm that locates these bottlenecks. We validated this algorithm in controlled settings and characterized bottlenecks on two deployments, one of which included 2,652 homes across the United States. We found that wireless bottlenecks are more common than access- link bottlenecks—particularly for home networks with downstream throughput greater than 20 Mbps, where access-link bottlenecks are relatively rare.

6.2. Characterizing Home Device Usage From Wireless Traffic Time Series

Participants: Katsiaryna Mirylenka (IBM), Vassilis Christophides, Themis Palpanas (University Rene Descartes), Ioannis Pefkianakis (HP Labs), Martin May (Technicolor)

We conducted a thorough analysis of traffic dynamics of heterogeneous wireless (WiFi) devices connected to 196 real RGWs, which are subscribers of a major European ISP. We focus on a time-oriented analysis of continuous traffic data to extract previously unknown patterns recurring of internet consumption that happen within, or across homes. We also assess the impact of different types of devices, such as laptops, desktops (classiffied as fixed devices), and tablets, smartphones (classiffied as portables), on these patterns. Unsupervised learning techniques are used for patterns discovery as the ground truth data regarding home activities are not available. Rather than partitioning homes or devices into distinct behavioral clusters, we are looking to extract informative motifs of bandwidth consumption within or across homes. The main contributions of this work are:

- We propose a novel analysis framework for wireless home traffic data, namely: (a) a correlation-based similarity measure, which exploits the evolution characteristics, rather than the absolute traffic values, and is invariant to scaling;(b) a notion of strong stationarity that in addition to the similarity of data distributions imposes a correlation similarity across non-overlapping time windows; and (d) a definition of dominant devices based on the correlation similarity, that enables an intuitive and statistically grounded interpretation of the results.
- We evaluate the effectiveness of the proposed framework using real data of wireless traffic observations and report the main findings: (a) there are many repetitive patterns within and across RGWs which describe the intrinsic user behavior of users and valuable to ISPs; (b) as networking time series are not stationary certain aggregation should be performed in order to find statistically significant patterns. The best time windows to aggregate home traffic data is found to be 8 hours for weekly patterns and 3 hours for daily patterns; (c) frequent weekly patterns correspond to heavy bandwidth usage both during weekdays and weekends, and frequent daily patterns correspond to (mostly) evening usage, (d) weekend usage tends to rely on portable devices, weekday usage relies more on fixed devices, while discontinuous usage within a day (mostly active in the evening or the morning) is still due to portable devices; and (e) almost every RGW involves a device that dominates its overall traffic, thus the behavior of this device should be mainly considered by ISPs while planning the updates.

6.3. Towards a Causal Analysis of Video QoE from Network and Application QoS

Participants: Michalis Katsarakis, Renata Teixeira, Maria Papadopouli (Univeristy of Crete), Vassilis Christophides

We have exploited an original framework for mining causal relationships among a 5-start rating of user QoE and various QoS metrics at network and application level. In particular, we have analysed QoE scores provided by a set of users for YouTube video streaming applications under different network conditions. We found that optimal QoE predictors we can be build using a minimal signature of only three features from application or network QoS metrics compared to four when features from both layers are considered. A thorough comparative analysis of the prediction accuracy of three models build using minimal signatures composed of (i) only network QoS, (ii) only application QoS, and (iii) both QoS features demonstrated that we can predict the QoE using only network QoS metrics and more surprisingly, predicting the QoE from network QoS metrics is as accurate as when using application QoS metrics. This work is the first step towards our ambition to assess QoE directly from network QoS metrics obtained via passive measurements of real traffic generated by online users. We will rely on the extracted minimal QoE/QoS signatures to build real-time predictors and compare their accuracy when using only network, only application or both QoS metrics. Last but not least, we plan to extend our experimental setting for other online applications such as teleconferencing services.

6.4. Predicting the effect of home Wi-Fi quality on Web QoE

Participants: Diego Neves da Hora, Renata Teixeira, Karel Van Doorselaer (Technicolor), Koen Van Oost (Technicolor)

We developed a model that predicts the effect of Wi-Fi quality on Web QoE, using solely Wi-Fi metrics commonly available in commercial APs. We trained our predictor during controlled experiments on a Wi-Fi testbed and assess its accuracy through cross-validation, obtaining an RMSE of 0.6432 MO, and by applying it on a separate validation dataset, obtained on an uncontrolled environment, finding an RMSE of 0.9283. Finally, we apply our predictor on Wi-Fi metrics collected in the wild from 4,880 APs over a period of 40 days. We find that Wi-Fi quality is mostly good for Web—in more than 60% of samples Wi-Fi quality does not degrade Web QoE. When we consider average complexity Web pages, however, Wi-Fi quality degrades Web QoE in 11% of samples. Moreover, we saw that 21% of devices present more than 20% of poor Web QoE samples, with 5% of these showing highly intermittent QoE degradations, which are particularly hard to diagnose, indicating the need for a long-term monitoring approach to detect and fix problems.

6.5. Passive Wi-Fi Link Capacity Estimation on Commodity Access Points

Participants: Diego Neves da Hora, Karel Van Doorselaer (Technicolor), Koen Van Oost (Technicolor), Renata Teixeira, Christophe Diot (Safran)

We propose an algorithm to estimate the link capacity based on passive metrics from APs, which is ready to be deployed at scale. We show that it is possible to estimate the link capacity per PHY rate based on a limited set of parameters related to the particular AP instance. Then, we extend the initial model to estimate the link capacity when the PHY rate varies. We measured the link capacity in different link quality conditions and found that more than 90% of the estimations present error below 15% without prior parameter tuning, and more than 95% present estimation error below 5% with appropriate parameter tuning using fixed PHY rate tests.

6.6. Content-Based Publish/Subscribe System for Web Syndication

Participants: Zeinab Hmedeh CNAM, Harry Kourdounakis (FORTH-ICS, Vassilis Christophides, Cedric du Mouza (CNAM), Michel Scholl (CNAM), and Nicolas Travers (CNAM)

Content syndication has become a popular way for timely delivery of frequently updated information on the Web. Today, web syndication technologies such as RSS or Atom are used in a wide variety of applications spreading from large-scale news broadcasting to medium-scale information sharing in scientific and professional communities. However, they exhibit serious limitations for dealing with information overload in Web 2.0. There is a vital need for efficient real-time filtering methods across feeds, to allow users to effectively follow personally interesting information.

To efficiently check whether all keywords of a subscription also appear in an incoming item (i.e., broad match semantics), we need to index the subscriptions. Count-based (CI) and tree-based (TI) are two main indexing schemes proposed in the literature for counting explicitly and implicitly the number of contained key-words. The majority of related data structures cannot be employed for conjunctions of keywords (rather than attribute-value pairs) due to the space high-dimensionality. In this paper, we are interested in efficient implementations of both indexing schemes using inverted lists (IL) for CI and a variant for distinct terms of ordered tries (OT) for TI and study their behavior for critical parameters of realistic web syndication workloads. Although these data structures have been employed to evaluate broad match queries in the context of selective information dissemination and sponsored search or for mining frequent item sets, their memory and matching time requirements appear to be quite different in our setting. This is due to the peculiarities of web syndication systems which are characterized 1) by information items of average length (25?36 distinct terms) which are greater than advertisement bids (4?5 terms) and smaller than documents of Web collections (12K terms) and 2) by very large vocabularies of terms (up to 1.5M terms). Note also that due to broad match semantics, information retrieval techniques for optimizing ILs (e.g., early pruning) are not suited in our setting.

We present analytical models for memory requirements and matching time and we conduct a thorough experimental evaluation to exhibit the impact of critical parameters of realistic web syndication workloads. We found that for small vocabularies, POT matching time is one order of magnitude faster than the best IL (RIL), while for large vocabularies (like the one used on the Web), RIL outperforms the matching POT, which uses almost four times more memory space. The actual distribution of term occurrences has almost no impact on the size of the three indexing structures while it significantly affects the number of nodes that need to be visited upon matching something that justifies OT performance gains. The smaller the subscription length, the larger the OT factorization gain w.r.t. IL and the larger the rank of the term from which the OT substructure degenerates to an IL.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. ANR

7.1.1.1. BottleNet: Understanding and Diagosing end-to-end communication problems

Type: ANR Project (N ANR-15-CE25-0013-01)

Instrument: PRCE

Duration: February 2016 - 36 mois

Coordinator: Renata Teixeira

Other partners: Inria Paris, LORIA, Ip-label SaS, ORANGE SA , TSP (SAMOVAR) Telecom SudParis, University of Lille 1 (CRIStAL)

Inria contact: Renata Teixeira

Abstract: As our lives become more dependent on the Internet, it is easy to understand peoples frustration when poor Internet performance prevents them from accomplishing ever-more important online activities. The Quality of Experience (QoE) when accessing the Internet is thus a key factor for todays society. When users experience poor Internet QoE, they are often helpless. The complexity of Internet services and of users local connectivity has grown dramatically in the last years with the proliferation of proxies and caches at the core and of home wireless and 3G/4G access. However, it is hard even for experts to diagnose the sources of performance bottlenecks. We argue that the new rules for regulating Internet access and plans to improve it should focus on the end-to-end Internet QoE - i.e., the user experience when performing her usual online activities on the Internet - and should provide means for all different players to accurately diagnose the sources of poor Internet QoE. The objective of BottleNet is to deliver methods, algorithms, and software systems to measure Internet QoE and diagnose the root cause of poor Internet QoE. Our goal calls for tools that run directly at users devices. We plan to collect network and application performance metrics directly at users devices and correlate it with user perception to model Internet QoE, and to correlate measurements across users and devices to diagnose poor Internet QoE. This data-driven approach is essential to address the challenging problem of modeling user perception and of diagnosing sources of bottlenecks in complex Internet services. BottleNet will lead to new solutions to assist users, network and service operators as well as regulators in understanding Internet QoE and the sources of performance bottleneck.

7.1.2. CNRS Big Data Projects

7.1.2.1. BigGeoQUP: Big Geospatial Data Quality and User Privacy

Type: CNRS Mission Interdisciplinarité (MI) DEFI MASTODONS "La qualité des données dans les Big Data"

Instrument: AAP 2016

Duration: Mars 2016 - 12 mois

Coordinator: Dimitris Kotzinos (University of Cergy Pointoise)

Other partners: Inria Paris, IGN-COGIT

Inria contact: Vassilis Christophides

Abstract: Big Geo Data represents an important type of the crowd sourced data that are available today at a global scale. This kind of data refers to locations, i.e., Points of Interest (POIs), and is usually published in social media (e.g., Facebook, Google+) or in specialized platforms (e.g., Open Street Maps, Yelp). The quality (e.g., precision, accuracy, consistency) of geo-referenced crowd sourced content depends on the origin (machine vs. human generated), the level of detail of the extraction methods, as well as the obfuscation techniques used to protect users privacy. There is clearly a tradeoff between enhancing the quality of published geo data and the privacy risks entailed for the individuals, also known as geoprivacy, to uncover places visited, trajectories pursuit etc. Understanding the different aspects of geographic/geometric/geospatial quality involved in crowd-sourced geo data and assessing the privacy risks introduced by enhancing its quality in personal, social, and urban applications is a challenging topic. In this project we are interested in studying the following questions in concrete use-cases:

- How can we measure the quality of geographic/geometric/geospatial data involved in crowd-sourced content?
- How can we assess the privacy risks introduced by enhancing geospatial quality in personal, social and urban applications especially in the context of social media platforms?

In this project Inria (MUSE) is interested in the opportunities and privacy concerns of mobile location analytics supporting customers or travellers experience in venues of various types and sizes (e.g., retail stores, shopping malls, airports, theme parks, etc.). Such (indoor) location-based services in smart spaces presents new privacy risks as data continuously flows between visitors mobile devices, networked sensors embedded into the environment as well as the backend analytics platform in order to track users and anticipate their needs.

7.2. European Initiatives

7.2.1. FP7 & H2020 Projects

7.2.1.1. User-Centric Networking (UCN)

Type: FP7

Instrument: Specific Targeted Research Project

Duration: October 2013 - September 2016

Coordinator: Technicolor

Other partners: Eurecom, Fraunhofer FOKUS, Intamac, University of Cambridge, University of Nottingham, Martel, NICTA, Portugal Telecom

Inria contact: Renata Teixeira

Abstract: This project introduces the concept of User Centric Networking (UCN), which is a new paradigm leveraging user information at large to deliver novel content recommendation systems and content delivery frameworks. UCN recommendation and content delivery systems will leverage in-depth knowledge about users to help them find relevant content, identify nearby network resources and plan how to deliver the actual content to the appropriate device at the desired time. These systems will additionally account for influences from users' social networks on their content consumption. The goal of this project is to design a UCN system architecture for user-centric connected media services. We will build UCN upon three complementary research pillars:

1. understanding user context: This data can be broadly categorized into three groups. First, the physical and environmental context A second category of data is that which can be extracted from social network interactions. The third category of data is behavioural
2. profiling and predicting user interests: By gaining a deep understanding of the user, we may be able to cast a much wider net in the content ocean and locate a richer catalogue of interesting content for the user
3. personalizing content delivery: Rather than the user (or the service provider) having to worry about the mode of connectivity, device, service, location, etc., the network intelligently directs and adapts the transport stream, or perhaps pre-fetches and replicates content chunks, to the particular and immediate needs of the user.

See also: http://usercentricnetworking.eu/

7.3. International Initiatives

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

- BetterNet (Inria Project Lab): An observatory to measure and Improve Internet Service Access from User Experience. Project Coordinator: Isabelle Chrisment, Partners: Inria teams Spirals, Diana, Muse, Dionysos and Madynes, Inria Chile

BetterNet intends to build and deliver a scientific and technical collaborative observatory to measure and improve the Internet service access as perceived by users. 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:

- tools, models and algorithms/heuristics will be provided to collect data,
- acquired data will be analyzed, and shared appropriately with scientists, stakeholders and civil society, and
- new value-added services will be proposed to end-users.

This project will also allow Inria to become a key reference in the digital field, not only for scientist researchers but also for policy makers, rulers, and, for citizens in general, by giving them a more accurate and reliable basis for decisions making at an individual scale (privacy strategies) or at a collective scale (legal norms). An originality of our proposal is to ensure that researchers from different disciplines (exact sciences, applied sciences and social sciences) will collaborate in the design of this observatory, in the dissemination of results for the research community, executives and public at large.

7.3.2. Inria International Partners

7.3.2.1. Informal International Partners

- Princeton (Prof. Nick Feamster): We have a long-term collaboration on measuring the performance of residential broadband Internet access networks and more recently on home network diagnosis.
- ICSI (Dr. Srikanth Sundaresan, Dr. Christian Kreibich, Dr. Robin Sommer): With C. Kreibich, we have been developing Fathom, a browser-based network measurement platform. We are now adding home network diagnosis capabilities to Fathom. We are collaborating with S. Sundaresan on detecting last-mile bottlenecks. In addition, with Robin Sommer we are working on the potential of matching the profiles of a user across multiple online social networks.
- Northwestern University (Prof. Fabian Bustamante and his doctoral student Zachary Bischof): we are working on identifying user activity from network traffic.

7.4. International Research Visitors

7.4.1. Internships

- Arash Molavi Kakhki, Ph.D Student, Northeastern University, from Nov 2016
- Julio Adriazola Soto, M2 Student, Inria Chile, from Jan 2016 until Mar 2016
- Michele Pittoni, M2 Student, Univ. Paris VI, from Feb 2016 until Jul 2016
- Adhir Chakravarti, M2 Student, Univ. Paris VI, from Feb 2016 until Jul 2016
- Zied Ben Houidi Bell Labs (Alcatel), from Apr 2016

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. Chair of Conference Program Committees

- R. Teixeira, Co-chair of the program committee of ACM IMC 2017,
- R. Teixeira, Co-chair of the program committee of ACM/ISOC ANRW 2017,
- V. Christophides, Tutorial co-chair of the 22nd IEEE Symposium on Computers and Communications (ISCC17) 03 - 06 July 2017, Heraklion, Crete, Greece (http://www.ics.forth.gr/iscc2017/)
- V. Christophides, Area Chair for "Semi-structured, Web, and Linked Data Management" in the 18th International Conference on Data Engineering (ICDE'16), May 16-20, 2016, Helsinki, Finland (http://icde2016.fi)
- V. Christophides, Workshop co-chair of the 1st International Workshop on Big Geo Data Quality and Privacy (BIGQP) with D. Kotzinos, C. Nikolaou, and Y. Theodoridis, March 21, 2017, Venice, Italy (http://www-etis.ensea.fr/BigGeoQ-UP/BIGQP2017/).
- V. Christophides, Workshop co-chair of the 2nd International Workshop on Preservation of Evolving Big Data, co-located with EDBT 2016 with G. Papastefanatos and S. Viglas March 15, 2016 Bordeaux, France (http://www.diachron-fp7.eu/2nd-diachron-workshop.html)

8.1.1.2. Member of the Conference Program Committees

- V. Christophides, PC member of the 2nd IFIP/IEEE International Workshop on Analytics for Network and Service Management (AnNet17) May 8, 2017 in Lisbon, Portugal.
- R. Teixeira, PC Member of 2016 Internet Measurement Conference (IMC) November 14-16, 2016 Santa Monica, California, USA.

- V. Christophides, PC member of the 1st International Workshop on Keyword-based Access and Ranking at Scale (KARS), collocated with EDBT2017, Venice, Italy.
- V. Christophides, PC member of the 25th ACM International Conference on Information and Knowledge Management (CIKM16), 24-28 October, Indianapolis, USA.

8.1.2. Invited Talks

- V. Christophides, "Web-scale Blocking, Iterative and Progressive Entity Resolution", ICDE 2017 Tutorial, with K. Stefanidis, and V. Efthymiou, April 19-22, 2017 in San Diego, California, USA.
- R. Teixeira, "HostView: Measuring Internet quality of experience on end-hosts", Dagsthul Seminar 16012, January 4 – 7, 2016.
- R. Teixeira, "The Challenges of Measuring Internet Quality of Experience", Invited Keynote at the Internet-QoE workshop co-located with SIGCOMM'16.
- R. Teixeira, "Residential Internet Performance: A view from the Gateway", Invited talk at the 1st Stakeholder Consultation Workshop on Mapping of Broadband Services in Europe (SMART 2014/0016).
- T. Friedman, A.-K. Pietilainen, R. Teixeira, "Internet Measurements: A Hands-on Introduction", Invited Tutorial at the ACM SIGCOMM 2016, Florianapolis, Brazil, August 22-26 2016.
- V. Christophides, "Resolving Entities in the Web of Data: Concepts, Algorithms and Systems", Invited Tutorial at the International Summer School on Web Science and Technology (WebST 2016) Bilbao, Spain, 18-22 July 2016, co-organized by Deusto and Rovira i Virgili University.

8.1.3. Leadership within the Scientific Community

- R. Teixeira, Vice-chair of ACM SIGCOMM (Since July 2013).
- R. Teixeira, Chair of the ACM SIGCOMM Industrial Liaison Board.
- R. Teixeira, Member of the steering committee of the ACM Internet Measurement Conference (Nov. 2009–Jun. 2016).
- V. Christophides, Member of the EDBT Association (since 2014).

8.1.4. Scientific Expertise

- Reviewer of H2020 and FP7 projects: CogNet, SUPERFLUIDITY, MONROE, MAMI.
- Technical advisory board of the project "Mapping of Broadband Services in Europe".

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master: Renata Teixeira, "Methodology for research in networking", 10h eqTD, M2, UPMC, France.

Master: Timur Friedman, Renata Teixeira, "Network Metrology", 24h CM, M2, UPMC, France. Created this new master's level class.

Master: Vassilis Christophides, "Big Data Processing and Analytics", 25h M2, Ecole CentraleSupelec, France. Created this new master's level class in Spring 2016.

E-learning

Mooc: Timur Friedman, Renata Teixeira, "Network Metrology", preparation of a five-week Mooc (May 2016) in the platform FUN, supported by Inria.

8.2.2. Supervision

PhD: Giuseppe Scavo (December 2016), "Content Curation and Characterization in Communities of a Place", advisors: Z. Ben-Houidi and R. Teixeira

PhD: Sofia Kleisarchaki (November 2016): "Difference Analysis in Big Data: Exploration, Explanation, Evolution" advisors: S. Amer-Yahia and V. Christophides

PhD in progress:

- Maximiliam Bachl (Oct 2016-now), "Home network troubleshooting made easy feedback", advisor: R. Teixeira
- Sara el Aouad (May 2014-now): "Improving the quality of recommendation using semi-structured user feedback", advisors: C. Diot and R. Teixeira
- Diego da Hora (Feb 2014-now): "Crowdsourced Home Network Diagnosis", advisors: C. Diot and R. Teixeira
- Vassilis Efthimiou (Feb 2013-now): "Entity resolution in the Web of Data", advisor: V. Christophides

8.2.3. Juries

R. Teixeira: Member of Doctoral examining committee of Sheharbano Khattak (University of Cambridge, 2016),

R. Teixeira: Member of Doctoral examining committee of Zachary Scott Bischoff (Northwestern University, 2016).

V. Christophides: External evaluator of the Ph.D. Thesis of Maria Koutraki, University PARIS-SACLAY, September 2016

V. Christophides: Member of the appointment committee for the MCF 4234 position on Big Data at the University of Cergy-Pontoise, Mai 2016.

9. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] F. GEERTS, T. UNGER, G. KARVOUNARAKIS, I. FUNDULAKI, V. CHRISTOPHIDES.*Algebraic Structures for Capturing the Provenance of SPARQL Queries*, in "Journal of the ACM", March 2016, vol. 63, n^{o} 1, 7 [*DOI :* 10.1145/2810037], https://hal.inria.fr/hal-01411827.

[2] Z. HMEDEH, H. KOURDOUNAKIS, V. CHRISTOPHIDES, C. DU MOUZA, M. SCHOLL, N. TRAVERS.*Content-Based Publish/Subscribe System for Web Syndication*, in "Journal of Computer Science and Technology", March 2016, vol. 31, p. 359 - 380 [*DOI :* 10.1007/s11390-016-1632-8], https://hal.inria.fr/hal-01411848.

International Conferences with Proceedings

[3] M. KATSARAKIS, R. TEIXEIRA, M. PAPADOPOULI, V. CHRISTOPHIDES.*Towards a Causal Analysis of Video QoE from Network and Application QoS*, in "ACM SIGCOMM Workshop on QoE-based Analysis and Management of Data Communication Networks (Internet-QoE 2016)", Florianopolis, Brazil, August 2016 [*DOI :* 10.1145/2940136.2940142], https://hal.inria.fr/hal-01338726.

[4] 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.

[5] G. POGHOSYAN, I. PEFKIANAKIS, P. LE GUYADEC, V. CHRISTOPHIDES.*Mining usage patterns in residential intranet of things*, in "The 7th International Conference on Ambient Systems, Networks and Technologies", Madrid , Spain, International Workshop on Big Data and Data Mining Challenges on IoT and Pervasive Systems (BigD2M 2016), May 2016, vol. 00, 6, https://hal.inria.fr/hal-01411676.

[6] S. SUNDARESAN, N. FEAMSTER, R. TEIXEIRA.*Home Network or Access Link? Locating Last-Mile Downstream Throughput Bottlenecks*, in "PAM 2016 - Passive and Active Measurement Conference", Heraklion, Greece, March 2016, p. 111-123 [*DOI :* 10.1007/978-3-319-30505-9_9], https://hal.inria.fr/hal-01294924.

Conferences without Proceedings

[7] D. NEVES DA HORA, R. TEIXEIRA, K. VAN DOORSELAER, K. VAN OOST.*Predicting the effect of home Wi-Fi quality on Web QoE*, in "Proceedings SIGCOMM Workshop on QoE-based Analysis and Management of Data Communication Networks (Internet-QoE)", Florianópolis, Brazil, ACM, August 2016 [*DOI :* 10.1145/2940136.2940146], https://hal.inria.fr/hal-01339522.

[8] D. NEVES DA HORA, K. VAN DOORSELAER, K. VAN OOST, R. TEIXEIRA, C. DIOT.*Passive Wi-Fi Link Capacity Estimation on Commodity Access Points*, in "Traffic Monitoring and Analysis Workshop (TMA) 2016", Louvain-la-Neuve, Belgium, April 2016, https://hal.inria.fr/hal-01292633.

Other Publications

[9] M. BACHL.*Collaborative Home Network Troubleshooting*, Université Pierre & Marie Curie - Paris 6 ; Inria, September 2016, https://hal.inria.fr/hal-01415767.

[10] V. EFTHYMIOU, K. STEFANIDIS, V. CHRISTOPHIDES.*Minoan ER: Progressive Entity Resolution in the Web of Data*, March 2016, 19th International Conference on Extending Database Technology, EDBT 2016, Poster [*DOI :* 10.5441/002/EDBT.2016.79], https://hal.inria.fr/hal-01411910.

[11] M. PITTONI.*Online Identification of Last-Mile Throughput Bottlenecks on Home Routers*, UPMC, September 2016, Supervisors: Renata Teixeira (Inria) , Promethee Spathis (UPMC)Advisors: Anna-Kaisa Pietilainen (Inria) , Srikanth Sundaresan (Samsara Networks/ICSI), Nick Feamster (Princeton University), https://hal.inria.fr/hal-01401856.

[12] N. VOUZOUKIDOU, B. AMANN, V. CHRISTOPHIDES.*Continuous Top-k Queries over Real-Time Web Streams*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01411893.

Project-Team MUTANT

Synchronous Realtime Processing and Programming of Music Signals

IN COLLABORATION WITH: Sciences et technologies de la musique et du son

IN PARTNERSHIP WITH:
CNRS

Institut de Recherche et Coordination Acoustique/Musique (IRCAM)

Université Pierre et Marie Curie (Paris 6)

RESEARCH CENTER
Paris

THEME
Embedded and Real-time Systems

Table of contents

Project-Team MUTANT

Creation of the Team: 2012 January 01, updated into Project-Team: 2013 January 01, end of the Project-Team: 2016 December 31

Keywords:

Computer Science and Digital Science:

2.1.8. - Synchronous languages
2.1.10. - Domain-specific languages
2.3. - Embedded and cyber-physical systems
2.4. - Verification, reliability, certification
3.3.2. - Data mining
3.4. - Machine learning and statistics
5.1.1. - Engineering of interactive systems
5.7. - Audio modeling and processing
5.7.2. - Music
7.4. - Logic in Computer Science
8.3. - Signal analysis

Other Research Topics and Application Domains:

9.2. - Art
9.2.1. - Music, sound

1. Members

Research Scientists

Jean-Louis Giavitto [CNRS, Senior Researcher (Team Leader from April 2016), HDR]
Florent Jacquemard [Inria, Researcher, HDR]

Technical Staff

Arshia Cont [Team leader until March 2016, IRCAM, HDR]
Adrien Ycart [Ircam, September 2015 until June 2016]

PhD Students

Julia Blondeau [CNRS]
Philippe Cuvillier [Univ. Paris VI]
Pierre Donat-Bouillud [Univ. Paris VI]
Clément Poncelet [Univ. Paris VI]
Maxime Sirbu [Univ. Paris VI]

Administrative Assistants

Chantal Chazelas [Inria]
Virginie Collette [Inria]
Cindy Crossouard [CNRS]

Others

Émile Enguehard [ENS Paris, Internship, Apr 2016 until Aug 2016]
Florent Mercier [Inria, Internship, from Apr 2016 until Sep 2016]
Samuel Wiseman [Inria, Internship, from Jun 2016 until Jul 2016]

2. Overall Objectives

2.1. Overall Objectives

The research conducted in *MuTant* is devoted both to leveraging capabilities of musical interactions between humans and computers, and to the development of tools to foster the authoring of interaction and time in computer music. Our research program belongs to the field of *Interactive music systems* for computer music composition and performance introduced in mid-1980s at Ircam. Within this paradigm, the computer is brought into the workflow of musical creation as an *intelligent performer* [64] and equipped with a *listening machine* [62] capable of analyzing, coordinating and anticipating its own and other musicians' actions within a musically coherent and synchronous context. Figure 1 illustrates this paradigm.

The use of Interactive Music Systems have become universal ever since and their practice has not ceased to nourish multidisciplinary research. From a research perspective, an interactive music systems deals with two problems: *Real-time Machine Listening* [61], [62] and *Synchronous and Timed Real-time Programming* [28] in *Computer Music*. The strong coupling and union (as opposed to an intersection) of the two field has become a necessity in music practices to provide temporal scenarios describing real-time interactions between computer environments and human musicians (in forms of programs or augmented music scores), and employ them in real-time on stage with a high degree of musical autonomy and competence, whilst ensuring the major issues of fault-tolerance and time-correctness.

Whereas each field has generated subsequent literature, few attempts have been made to address the global problem by putting the two domains in direct interaction.

Figure 1. General scheme of Interactive Music Systems

MuTant's research program has developed new *Real-time Machine Listening* mechanisms (see Section 3.1), new reactive and strongly timed real-time software architectures (see Section 3.2), as well as contributions to the field of verification and test on dynamic setups and work-flows such as those observed in Music (see Section 3.3). The major incarnation of our research is the award winning *Antescofo* language and real-time system, deployed since our inception in major international festivals with more than 100 known repertoire pieces regularly played throughout the world.

3. Research Program

3.1. Machine Listening

Participants: Arshia Cont, Philippe Cuvillier, Florent Jacquemard, Maxime Sirbu, Adrien Ycart.

When human listeners are confronted with musical sounds, they rapidly and automatically find their way in the music. Even musically untrained listeners have an exceptional ability to make rapid judgments about music from short examples, such as determining music style, performer, beating, and specific events such as instruments or pitches. Making computer systems capable of similar capabilities requires advances in both music cognition, and analysis and retrieval systems employing signal processing and machine learning.

Machine listening in our context refers to the capacity of our computers to understand “non-speech sound” by analyzing the content of music and audio signals and combining advanced signal processing and machine learning. The major focus of MuTant has been on Real-time Machine listening algorithms spanning *Real-time Recognition Systems* (such as event detection) and also *Information Retrieval* (such as structure discovery and qualitative parameter estimation). Our major achievement lies in our unique Real-time Score Following (aka Audio-to-Score Alignment) system that are featured in the Antescofo system (cf. Section 5.1). We also contributed to the field of On-line Music Structure Discovery in Audio Processing, and lately to the problem of off-line rhythmic quantization on Symbolic Data.

3.1.1. Real-time Audio-to-Score Alignment.

This is a continuation of prior work of team-founder [1] which proved the utility of strongly-timed probabilistic models in form of Semi-Markov Hidden States. Our most important theoretical contribution is reported in [37], [38] that introduced Time-coherency criteria for probabilistic models and led to general robustness of the Antescofo listening machine, and allowed its deployment for all music instruments and all setups around the world. We further studied the integration of other recognition algorithms in the algorithm in form of *Information Fusion* and for singing voice based on Lyric data in [49]. Collaboration with our japanese counterparts led to extensions of our model to the symbolic domain reported in [56]. Collaboration with the SIERRA team created a joint research momentum for fostering such applications to weakly-supervised discriminative models reported in [54]. Our Real-time Audio-to-Score alignment is a major component of the Antescofo software described in Section 5.1.

3.1.2. Online Methods for Audio Segmentation and Clustering.

To extend our listening approach to general sound, we envisioned dropping the prior information provided by music scores and replacing it by the inherent structure in general audio signals. Early attempts by the team leader employed [2] Methods of Information Geometry, an attempt to join Information Theory, Differential Geometry and Signal Processing. We were among the first teams in the world advocating the use of such approaches for audio signal processing and we participated in the growth of the community. A major breakthrough of this approach is reported in [39] and the PhD Thesis [40] that outline a general real-time change detection mechanism. Automatic structure discovery was further pursued in a MS thesis project in 2013 [55]. By that time we realized that Information Manifolds do not necessarily provide the invariance needed for automatic structure discovery of audio signals, especially for natural sounds. Following this report, we pursued an alternative approach in 2014 and in collaboration with the Inria SIERRA Team [30]. The result of this joint work was published in IEEE ICASSP 2015 and won the best student paper award [29]. We are currently studying massive applications of this approach to natural sounds and in robotics applications in the framework of Maxime Sirbu’s PhD project.

3.1.3. Symbolic Music Information Retrieval and Rhythm Transcription.

Rhythmic data are commonly represented by tree structures (rhythms trees) due to the theoretical proximity of such structures with the proportional representation of time values in traditional musical notation. We are studying the application to rhythm notation of techniques and tools for symbolic processing of tree structures, in particular tree automata and term rewriting.

Our main contribution in that context is the development of a new framework for rhythm transcription [23], [22], [65], [31] addressing the problem of converting a sequence of timestamped notes, *e.g.* a file in MIDI format, into a score in traditional music notation. This problem is crucial in the context assisted music composition environments and music score editors. It arises immediately as insoluble unequivocally: in order to fit the musical context, the system has to balance constraints of precision and readability of the generated scores. Our approach is based on algorithms for the exploration and lazy enumeration of large sets of weighted trees (tree series), representing possible solutions to a problem of transcription. A side problem concerns the equivalent notations of the same rhythm, for which we have developed a term rewrite approach, based on a new equational theory of rhythm notation [42], [51], [52].

3.2. Synchronous and realtime programming for computer music

Participants: Julia Blondeau, Arshia Cont, Jean-Louis Giavitto.

The research presented here aims at the development of a programming model dedicated to authoring of time and interaction for the next generation of interactive music systems. Study, formalization and implementation of such programming paradigm, strongly coupled to recognition systems discussed in the previous section, constitutes the second objective of the MuTant project.

The tangible result of this research is the development of the Antescofo system (cf. Section 5.1) for the design and implementation of musical scenarios in which the human and computer actions are in constant real-time interaction. Through such development, Antescofo has already made itself into the community; it serves as the backbone of temporal organization of more than 100 performances since 2012 and used both for preexisting pieces and new creations by music ensembles such as Berliner Philharmoniker, Los Angeles Philharmonic, Ensemble Intercontemporain or Orchestre de Paris to name a few.

Compared to programmable sequencers or interactive music systems (like Max or PureData) the Antescofo DSL offers a rich notion of time reference and provides explicit time frame for the environment with a comprehensive list of musical synchronization strategies and proposes and predictable mechanisms for controlling time at various timescales (temporal determinism) and across concurrent code modules (time-mediated concurrency).

3.2.1. Multiple Times.

Audio and music often involve the presence and cooperation of multiple notions of *time*: an ideal time authored by the composer in a score and also a performance time produced jointly by the performers and the real-time electronics; where instant and duration are expressed both in physical time (milliseconds), in relative time (relative to an unknown dynamic tempo) or through logical events and relations ("at the peak of intensity", "at the end of the musical phrase", "twice faster").

Antescofo is the first languages that addresses this variety of temporal notions, relying on the synchronous approach for the handling of atomic and logical events and an anticipative notion of tempo for the handling of relative duration [35], [45]. A first partial model of time at work in Antescofo (single time, static activities) has been formalized relying on parametric timed automata [43] and constitutes the reference semantics for tests (cf. section 3.3). A denotational semantics of the complete language (multiple times and dynamic constructions including anticipative synchronization strategies) has been published in [44].

3.2.2. Human-Computer Synchronizations.

Antescofo introduces the notion of *temporal scope* to formalize relationships between temporal information specified in the score and their realization during a performance [36]. A temporal scope is attached to a sequence of actions, can be inherited or dynamically changed as a result of a computation. A synchronization strategy is part of a temporal scope definition. They use the performer's position information and its tempo estimation from the listening module, to drive the passing of time in a sequence of atomic and durative actions.

Synchronization strategies have been systematically studied to evaluate their musical relevance in collaboration with Orchestre de Paris and composer Marco Stroppa. Anticipative strategies enable handling of uncertainties inherent in musical event occurrence, exhibiting a smooth musical rendering whilst preserving articulation points and target events [63].

3.2.3. Temporal Organization.

Several constructions dedicated to the expression of the temporal organization of musical entities and their control have enriched the language from the start of the project. These construction have been motivated by composer's research residences in our team: representation of open scores (J. Freeman); anticipative synchronization strategies (C. Trapani); adaptive sampling of continuous curve in relative time for the dynamic control of sound synthesis (J.-M. Fernandez); musical gesture (J. Blondeau); first class processes, actors and continuation combinators for the development of libraries of reusable parametric temporal behaviors (M. Stroppa, Y. Maresz); *etc.*

The reaction to a logical event is a unique feature in the computer music system community [57]. It extend the well known `when` operator in synchronous languages with process creation. Elaborating on this low-level mechanism, *temporal patterns* [48] enable expression of complex temporal constraints mixing instant and duration. The problem of online matching where the event are presented in real time and the matching is computed incrementally as well, has received a recent attention from the model-checking community, but with less constrained causal constraints.

3.2.4. Visualization and Monitoring of Event-driven and Time-driven Computations.

The authoring of complex temporal organization can be greatly improved through adapted visual interfaces, and has led to the development of *AscoGraph*, a dedicated user interface to Antescofo. Ascograph is used both for edition and monitoring interface of the system during performances [34]. This project was held from end 2012 to end 2014 thanks to Inria ADT and ANR support.

An information visualisation perspective has been taken for the design of timeline-based representation of action items, looking for information coherence and clarity, facility of seeking and navigation, hierarchical distinction and explicit linking [33] while minimizing the information overload for the presentation of the nested structure of complex concurrent activities [32].

3.3. Semantics, Verification and Test of Mixed Scores

Participants: Jean-Louis Giavitto, Florent Jacquemard, Clément Poncelet.

We address the questions of *functional reliability* and *temporal predictability* in score-based interactive music systems such as Antescofo. On the one hand, checking these properties is difficult for these systems involving an amount of human interactions as well as timing constraints (for audio computations) beyond those of many other real-time applications such as embedded control. On the other hand, although they are expected to behave properly during public concerts, these systems are not safety critical, and therefore a complete formal certification is not strictly necessary in our case.

Our objective in this context is to provide techniques and tools to assist both programmers of scores (*i.e.* composers) and the developers of the system itself. [47], [46]. It should be outlined that the former are generally not experts in real-time programming, and we aim at giving them a clear view of what will be the outcome of the score that they are writing, and what are the limits of what is playable by the system. To help the development of Antescofo, we have built a framework for automated timed conformance testing. [14], [18], [58], [60], [59].

In both cases, it is important to be able to predict statically the behavior of the system in response to every possible musician input. This cannot be done manually and requires first a formal definition of the semantics of scores, and second using advanced symbolic state exploration techniques (model checking) [43].

4. Highlights of the Year

4.1. Highlights of the Year

Startup Creation

Arshia Cont with José Echeveste and Philippe Cuvillier (former PhD students) are creating a Startup around Antescofo to bring the product to greater public starting March 2016 http://antescofo.com. The project is hosted by the French Incubator AgoraNov.

It was awarded the “Emergence Award” in 2015 that help emerging new technology companies to study the project, and an i-LAB prize in 2016, supported by the French Ministry of Culture and Bpifrance, and it has been a finalist of the Midemlab 2016.

5. New Software and Platforms

5.1. Antescofo

Anticipatory Score Following and Real-time Language

FUNCTIONAL DESCRIPTION. Antescofo is a modular polyphonic Score Following system as well as a Synchronous Programming language for musical composition. The first module allows for automatic recognition of music score position and tempo from a realtime audio Stream coming from performer(s), making it possible to synchronize an instrumental performance with computer realized elements. The synchronous language (DSL) within Antescofo allows flexible writing of time and interaction in computer music.

- Participants: Arshia Cont, Jean-Louis Giavitto, Florent Jacquemard and José Echeveste
- Contact: Arshia Cont
- URL: http://forumnet.ircam.fr/product/antescofo/

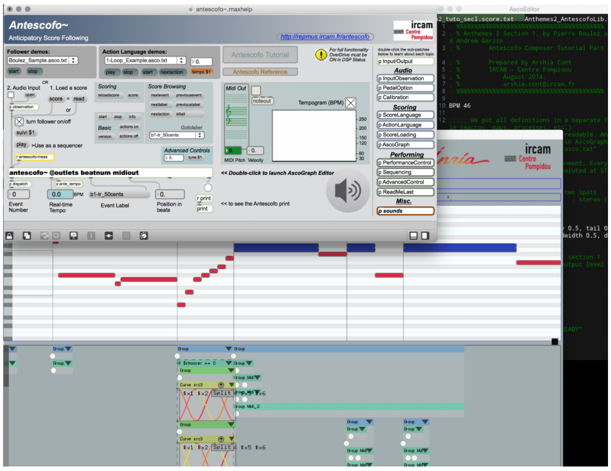

Figure 2. Antescofo and AscoGraph Screenshots

The design of the Antescofo DSL clearly benefits of a strong and continuous involvement in the production of world-class composer pieces and their continuous recreation throughout the world. These interactions motivate new developments, challenge the state of the art and in return, opens new creative dimensions for composers and musicians. The maturity of the system is assessed by the generalization of its use in a large proportion of Ircam new productions, and its use outside Ircam all around the world (Brasil, Chile, Cuba, Italy, China, US, etc.). Antescofo enjoys an active community of 150 active users: http://forumnet.ircam.fr/user-groups/antescofo/

5.2. OMRQ

Library for rhythm transcription integrated in the assisted composition environment OpenMusic.

FUNCTIONAL DESCRIPTION. Rhythm transcription is the conversion of sequence of timed events into the structured representations of conventional Western music notation. Available as a graphical component of OpenMusisc, the library OMRQ privileges user interactions in order to search for an appropriate balance between different criteria, in particular the precision of the transcription and the readability of the musical scores produced.

This system follows a uniform approach, using hierarchical representations of timing notations in the form of rhythm trees, and efficient parsing algorithms for the lazy enumeration of solutions of transcription.

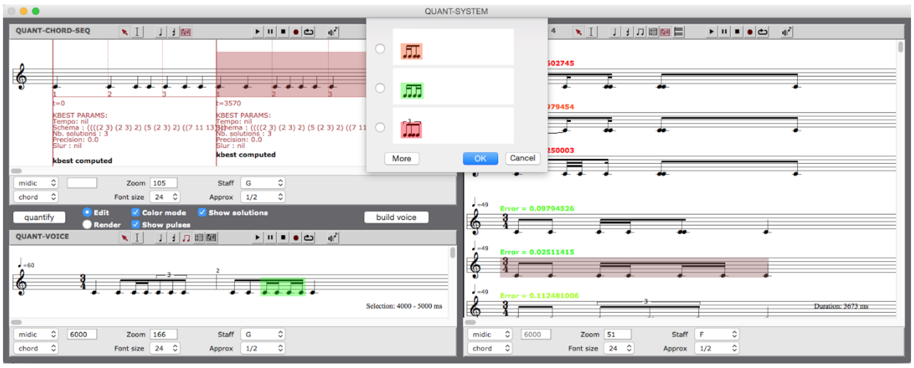

Figure 3. Screenshot of the Open Music Rhythm Quantization library

Its implementation is carried out via a dedicated interface allowing interactive exploration of the solutions space, their visualization and local editing, with particular attention to the processing of grace notes and rests.

- Participants: Florent Jacquemard and Adrien Ycart
- Contact: Florent Jacquemard
- URL: http://repmus.ircam.fr/cao/rq, https://bil.inria.fr/fr/software/view/2904/tab

5.3. Antescofo Timed Test Platform

Timed testing plateform for Antescofo.

FUNCTIONAL DESCRIPTION. The frequent use of Antescofo in live and public performances with human musicians implies strong requirements of temporal reliability and robustness to unforeseen errors in input. To address these requirements and help the development of the system and authoring of pieces by users, we are developing a platform for the automation of testing the behavior of Antescofo on a given score, with of focus on timed behavior. It is based on state of the art techniques and tools for *model-based testing* of embedded systems [50], and makes it possible to automate the following main tasks:

1. offline and on-the-fly generation of relevant input data for testing (i.e. fake performances of musicians, including timing values), with the sake of exhaustiveness,
2. computation of the corresponding expected output, according to a formal specification of the expected behavior of the system on a given mixed score,
3. black-box execution of the input test data on the System Under Test,
4. comparison of expected and real output and production of a test verdict.

The input and output data are timed traces (sequences of discrete events together with inter-event durations). Our method is based on formal models (specifications) in an ad hoc medium-level intermediate representation (IR). We have developed a compiler for producing automatically such IR models from Antescofo high level mixed scores.

Then, in the offline approach, the IR is passed, after conversion to Timed Automata, to the model-checker Uppaal, to which is delegated the above task (1), following coverage criteria, and the task (2), by simulation. In the online approach, tasks (1) and (2) are realized during the execution of the IR by a Virtual Machine developed on purpose. Moreover, we have implemented several tools for Tasks (3) and (4), corresponding to different boundaries for the implementation under test (black box): e.g. the interpreter of Antescofo's synchronous language alone, or with tempo detection, or the whole system.

- Participants: Clément Poncelet, Florent Jacquemard, Pierre Donat-Bouillud
- Contact: Clément Poncelet

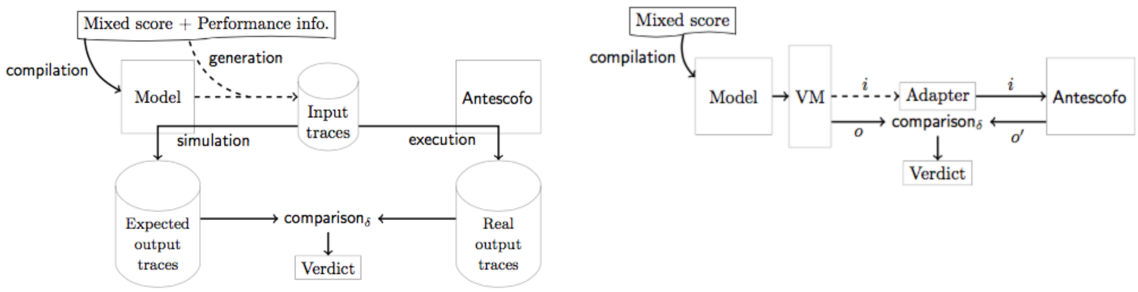

Figure 4. Offline and Online workflows for Antescofo Model Based Testing

These implementations have been conducted as a part of Clément Poncelet's PhD Thesis.

5.4. Ascograph

The Antescofo graphical score editor.

FUNCTIONAL DESCRIPTION. AscoGraph, released in 2013, provides a autonomous Integrated Development Environment (IDE) for the authoring of Antescofo scores. Antescofo listening machine, when going forward in the score during recognition, uses the message passing paradigm to perform tasks such as automatic accompaniment, spatialization, etc. The Antescofo score is a text file containing notes (chord, notes, trills, ...) to follow, synchronization strategies on how to trigger actions, and electronic actions (the reactive language).

This editor shares the same score parsing routines with Antescofo core, so the validity of the score is checked on saving while editing in AscoGraph, with proper parsing errors handling.

Graphically, the application is divided in two parts (Figure 5). On the left side, a graphical representation of the score, using a timeline with tracks view. On the right side, a text editor with syntax coloring of the score is displayed. Both views can be edited and are synchronized on saving. Special objects such as "curves", are graphically editable: they are used to provide high-level variable automation facilities like breakpoints functions (BPF) with more than 30 interpolations possible types between points, graphically editable.

6. New Results

6.1. Embedding Audio Processing

Participants: Jean-Louis Giavitto, Pierre Donat-Bouillud.

Audio processing has been integrated in the Antescofo language. This experimental extension aims at providing sample-accurate control and dynamic audio graphs directly in Antescofo. Currently, FAUST

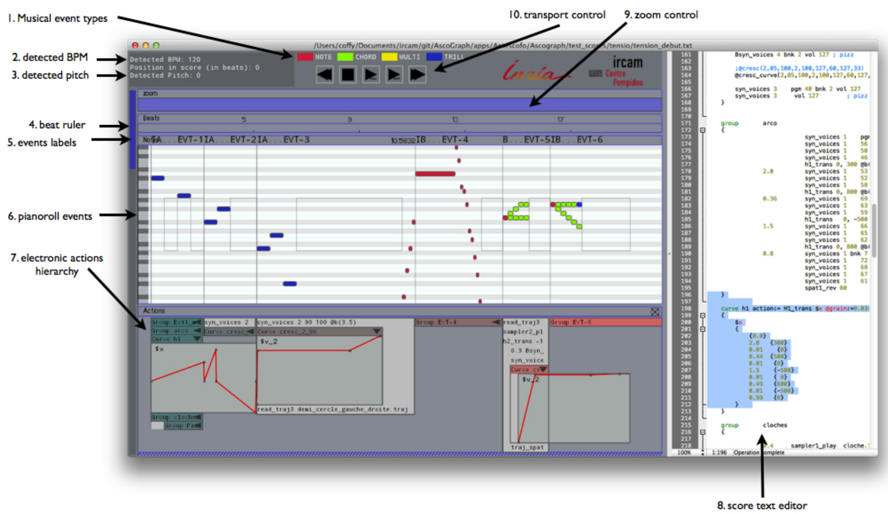

Figure 5. Screenshot of Ascograph, the Antescofo graphical score editor

(through a native embedding of the in-core compiler) and a few specific signal processors (notably FFT) can be defined. The tight integration enable specification of multiple-timed signal processing in conjunction with control programs. One example of this integration is the use of symbolic curve specification to specify variations of control parameters at sample rate, a task whose correctness in real-time is not at the scope of competing systems. Our approach has proven to provide such mechanisms at a lower computational cost; for example a factor of two in the *remaking* of Boulez' piece *Antheme 2* compared to the original version with the audio effects managed in Max. We will further pursue such optimizations while extending sample accuracy, by developing a type-system to preserve block computations in case of preemptive audio processing [41].

The reduced footprint enable the embedding of an *Antescofo* engine with internal audio processing on Raspberry PI and UDOO nano-computers (early results are reported in [26]).

6.2. Representation of Rhythm and Quantization

Participants: Florent Jacquemard, Adrien Ycart, Pierre Donat-Bouillud.

Rhythmic data are commonly represented by tree structures (rhythms trees) in assisted music composition environments, such as OpenMusic, due to the theoretical proximity of such structures with traditional musical notation. We are studying the application in this context of techniques and tools for processing tree structures, which were originally used in natural language processing. We are particularly interested in two well established formalisms with solid theoretical foundations: weighted automata for trees and dags and term rewriting.

Our main contribution in that context is the development of a new framework for rhythm transcription, the problem of the generation, from a sequence of timestamped notes, *e.g.* a file in MIDI format, of a score in traditional music notation) – see Section 5.2. This problem arises immediately as insoluble unequivocally: we shall calibrate the system to fit the musical context, balancing constraints of precision, or of simplicity / readability of the generated scores. In collaboration with Jean Bresson (Ircam) and Slawek Staworko (LINKS), we are developing an approach based on algorithms for the enumeration of large sets of weighted trees (tree series), representing possible solutions to a problem of transcription. The implementation work is performed by Adrien Ycart, under a research engineer contract with Ircam. This work has been presented in [22], [23].

Moreover, in collaboration with Prof. Masahiko Sakai (Nagoya University), we are working on symbolic processing of music notation, based on the above models. We proposed a structural theory (equational system on rhythm trees) defining equivalence on rhythm notations [42], [51], and use this approach, for instance, to generate, by transformation, different possible notations of the same rhythm, with the ability to select either alternative notation in accordance with certain constraints, *e.g.* in the context of transcription.

Related results on the property of confluence of term rewriting systems were presented in [19] (invited talk), and other work on data tree processing, in collaboration with Luc Segoufin and Jeremie Dimino, have published in [16].

6.3. Model-based Testing an Interactive Music System

Participants: Clément Poncelet, Florent Jacquemard, Pierre Donat-Bouillud.

We have been pursuing in 2016 our applications of model-based timed testing techniques to the interactive music system Antescofo, in the context of the Phd of Clément Poncelet and in relation with the developments presented in Section 5.3.

Several formal methods have been developed for automatic conformance testing of critical embedded software, with the execution of a real implementation under test (IUT, or black-box) in a testing framework, where carefully selected inputs are sent to the IUT and then the outputs are observed and analyzed. In conformance model-based testing (MBT), the input and corresponding expected outputs are generated according to formal models of the IUT and the environment. The case of IMS presents important originalities compared to other applications of MBT to realtime systems. On the one hand, the time model of IMS comprises several time units, including the wall clock time, measured in seconds, and the time of music scores, measured in number of beats relatively to a tempo. This situation raises several new problems for the generation of test suites and their execution. On the other hand, we can reasonably assume that a given mixed score of Antescofo specifies completely the expected timed behavior of the IMS, and compile automatically the given score into a formal model of the IUT's expected behavior, using an intermediate representation. This give a fully automatic test method, which is in contrast with other approaches which generally require experts to write the specification manually.

We have developed online and offline approaches to MBT for Antescofo. The offline approach relies on tools of the Uppaal suite [53], [50], using a translation of our models into timed automata. The online approach is based on a virtual machine executing the models of score in Intermediate Representation (IR).

To this respect, the transformation of Antescofo's mixed scores (in DSL) into IR, described in Section 5.3, can be seen as the premise of a compiled approach for Antescofo.

These results have been published this year in the Journal of New Music Research [14], the journal Science of Computer Programming [18], and in the PhD of Clément Poncelet, defended in November 2016.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. ANR

Mutant was the PI of the ANR INEDIT project, ended in october 2015. The INEDIT project aims to provide a scientific view of the interoperability between common tools for music and audio productions, in order to open new creative dimensions coupling *authoring of time* and *authoring of interaction*.

Mutant participates also actively in the Efficace ANR Project. This project explores the relations between computation, time and interactions in computer-aided music composition, using OpenMusic and other technologies developed at IRCAM and at CNMAT (UC Berkeley).

The MuTant team is also an active member of the ANR CHRONOS Network by Gérard Berry, Collège de France).

7.2. European Initiatives

7.2.1. Collaborations in European Programs, Except FP7 & H2020

Program: PHC Amadeus (France-Austria)

Project acronym: LETITBE

Project title: Logical Execution Time for Interactive And Composition Assistance Music Systems

Duration: 01/2015 - 01/2017

Coordinator: Florent Jacquemard, Christoph Kirsch

Other partners: Department of Computer Sciences University of Salzburg, Austria

Abstract: The objective of the LETITBE project is to contribute to the development of computer music systems supporting advanced temporal structure in music and advanced dynamics in interactivity. For this purpose we are proposing to re-design and re-engineer computer music systems (from IRCAM at Paris) using advanced notions of time and their software counterparts developed for safety-critical embedded systems (from University of Salzburg). In particular, we are applying the so-called logical execution time paradigm as well as its accompanying time safety analysis, real-time code generation, and portable code execution to computer music systems. Timing in music is obviously very important. Advanced treatment of time in safety-critical embedded systems has helped address extremely challenging problems such as predictability and portability of real-time code. We believe similar progress can be made in computer music systems potentially enabling new application areas. The objective of the project is ideally suited for a collaboration of partners with complementary expertise in computer music and real-time systems.

This year, Pierre Donat-Bouillud has spent 5 months in the University of Salzburg and one month in the University of California Berkeley, in the context of the LETITBE project, before starting his PhD in Mutant. Several other student exchanges and scientists visits between Salzburg and Paris have been funded this year by the LETITBE projetc

7.3. International Initiatives

7.3.1. Inria International Partners

7.3.1.1. Informal International Partners

- We are collaborating with Slawek Staworko (LINKS and Algomus, Lille – on leave at U. Edinburgh in 2016), and the Algomus group at Lille, in the context of our projects on rhythm transcription described at Sections 5.2 and 6.2. This collaboration led this year to the following publications: [23], [22].
- We are pursuing a long term collaboration with Masahiko Sakai (U. Nagoya) on term rewriting techniques and applications (in particular applications related to rhythm notation) [19], [27].
- MuTant team collaborates with *Bucharest Polytechnic University*, in the framework of Grig Burloiu's PhD Thesis on *AscoGraph* UIX design which has resulted in a the new design of AscoGraph (see 5.4) and publications [13], [32], [33].
- MuTant team collaborated with researchers at National Institute of Informatics of Tokyo on real-time Symbolic Alignment of music data [56].

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Selection

8.1.1.1. Member of the Conference Program Committees

Jean-Louis Giavitto has participated in the program committee of the 42st International Computer Music Conference (ICMC), the 10th IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO 2016), the Digital Entertainment Technologies and Arts(DETA) of GECCO-2016, the 15th International Conference on the Synthesis and Simulation of Living Systems (ALIFE XV) and the 2nd International Conference on Technologies for Music Notation and Representation (TENOR 2016).

Florent Jacquemard has been involved in the Program Committees of the 2d International Conference on Technologies for Music Notation and Representation (TENOR 2016), the 8th International Symposium on Symbolic Computation in Software (SCSS 2017), the National Conference *Journées d'Informatique Musicale* (JIM 2016), and the special issue of the journal Information and Computation for the 10th International Conference on Language and Automata Theory and Applications (LATA 2016).

8.1.1.2. Member of the Editorial Boards

Jean-Louis Giavitto is associate redactor (former redactor-in-chief) of TSI (Technique et Science Informatiques) published by Lavoisier. He has coorganized with Antoine Spicher (Univ. Paris Est), Stefan Dulman (Univ. Twente) and Mirko Viroli (Univ. of Milano) a special issue of *The Knowledge Engineering Review* on Spatial Computing published in november 2016.

8.1.1.3. Reviewer - Reviewing Activities

The members of the team contributed as reviewers for the journal Information and Computation, IEEE Transactions on Multimedia, IEEE Transactions on Audio and Speech Signal Processing, ACM Transactions on Intelligent Systems, Theoretical Computer Science, IEEE ICASSP, ICMC, SMC, Formal Methods, LATA and more...

8.1.2. Invited Talks

Jean-Louis Giavitto was invited to the seminar @SystemX at Saclay.

Florent Jacquemard gave an invited talk at the 5th International Workshop on Confluence (IWC 2016), hosted by Innsbruck University Center at Obergurgl, Austria [19].

8.1.3. Leadership within the Scientific Community

Jean-Louis Giavitto was a member of the Prix de thèse du GDR GPL as well as a member of the Faust Award 2016 (the Faust Open-Source Software Competition is intended to promote innovative high-quality free audio software developed with the Faust programming language, as well as development tools build around the Faust compiler itself).

8.1.4. Scientific Expertise

Jean-Louis Giavitto is in scientific board of the GDR GPL (Genie de la programmation et du logiciel). He is also a reviewer for FET projects for the UC.

8.2. Teaching - Supervision - Juries

8.2.1. Supervision

- PhD defended: José Echeveste, *Accorder le temps de la machine et celui du musicien*, started in October 2011, supervisor: Arshia Cont and Jean-Louis Giavitto.
- PhD defended (November 2016): Clément Poncelet, Formal methods for analyzing human-machine interaction in complex timed scenario. Started in October 2013, supervisor: Florent Jacquemard.
- PhD defended (December 2016): Philippe Cuvillier, Probabilistic Decoding of strongly-timed events in realtime, supervisor: Arshia Cont.
- PhD in progress: Julia Blondeau, *Espaces compositionnels et temps multiples : de la relation forme/matériauq (thèse en art)*, supervisor: Jean-Louis Giavitto, co-director Dominique Pradelle (Philosophy, Sorbonne), started October 2015.
- PhD in progress: Maxim Sirbu, Online Interaction via Machine Listening. Supervisors: Arshia Cont (MuTant) and Mathieu Lagrange (IrCyNN), started October 2015.
- PhD in progress: Pierre Donat-Bouillud, Modeling, analysis and execution of cyber-temporal systems. Supervisor: Florent Jacquemard, co-director: Jean-Louis Giavitto, started October 2016.

8.2.2. Juries

Jean-Louis Giavitto was Chairman of the jury of Clément Poncelet. He was reviewer of the PhD thesis of Jaime Arias (University of Bordeaux, Sémantique Formelle et Vérification Automatique de Scénarios Hiérarchiques Multimédia avec des Choix Interactifs), and examiner of the PhD of Mattia Bergomi (Università di Milano and UPMC, Dynamical and Topological Tools for (Modern) Music Analysis).

Florent Jacquemard was reviewer of the PhD thesis of Etienne Dubourg (University of Bordeaux, Contributions to the theory of tile languages). He is reviewer of the PhD thesis of Nicolas Guiomard-Kagan (Université de Picardie Jules Verne, *Traitement de la polyphonie pour l'analyse informatique de partitions musicales*). He has been examiner of the PhD of Emil-Mircea Andriescu (UPMC, MiMove, Dynamic Data Adaptation for the Synthesis and Deployment of Protocol Mediators) and examiner of the PhD of Carles Creus Lo´pez (UPC Barcelona, Tree Automata with Constraints and Tree Homomorphisms).

9. Bibliography

Major publications by the team in recent years

[1] A. CONT.*A coupled duration-focused architecture for realtime music to score alignment*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2010, vol. 32, n^{o} 6, p. 974-987, http://articles.ircam.fr/textes/Cont09a/.

[2] A. CONT, S. DUBNOV, G. ASSAYAG.*On the Information Geometry of Audio Streams with Applications to Similarity Computing*, in "IEEE Transactions on Audio, Speech and Language Processing", May 2011, vol. 19, n^o^ 4.

[3] A. CONT, J. ECHEVESTE, J.-L. GIAVITTO, F. JACQUEMARD.*Correct Automatic Accompaniment Despite Machine Listening or Human Errors in Antescofo*, in "ICMC 2012 - International Computer Music Conference", Ljubljana, Slovenia, IRZU - the Institute for Sonic Arts Research, September 2012, http://hal.inria.fr/hal-00718854.

[4] P. DONAT-BOUILLUD, J.-L. GIAVITTO, A. CONT, N. SCHMIDT, Y. ORLAREY.*Embedding native audio-processing in a score following system with quasi sample accuracy*, in "ICMC 2016 - 42th International Computer Music Conference", Utrecht, Netherlands, September 2016, https://hal.inria.fr/hal-01349524.

[5] J. ECHEVESTE, A. CONT, J.-L. GIAVITTO, F. JACQUEMARD.*Operational semantics of a domain specific language for real time musician-computer interaction*, in "Discrete Event Dynamic Systems", August 2013, vol. 23, n^o^ 4, p. 343-383 [*DOI :* 10.1007/s10626-013-0166-2], http://hal.inria.fr/hal-00854719.

[6] L. FANCHON, F. JACQUEMARD.*Formal Timing Analysis Of Mixed Music Scores*, in "2013 ICMC - International Computer Music Conference", Perth, Australia, August 2013, https://hal.inria.fr/hal-00829821.

[7] F. JACQUEMARD, C. PONCELET.*An Automatic Test Framework for Interactive Music Systems*, in "Journal of New Music Research", 2016, vol. 45, n^o^ 2, 18, https://hal.inria.fr/hal-01274035.

[8] C. PONCELET, F. JACQUEMARD.*Model-Based Testing for Building Reliable Realtime Interactive Music Systems*, in "Science of Computer Programming", 2016, Journal: Science of Computer Programming special issue for ACM SAC (SVT) 2015, https://hal.archives-ouvertes.fr/hal-01314969.

[9] C. PONCELET SANCHEZ, F. JACQUEMARD.*Test Methods for Score-Based Interactive Music Systems*, in "ICMC SMC 2014", Athen, Greece, September 2014, https://hal.inria.fr/hal-01021617.

[10] A. YCART, F. JACQUEMARD, J. BRESSON, S. STAWORKO.*A Supervised Approach for Rhythm Transcription Based on Tree Series Enumeration*, in "International Computer Music Conference (ICMC)", Utrecht, Netherlands, Proceedings of the 42nd International Computer Music Conference (ICMC), September 2016, https://hal.inria.fr/hal-01315689.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] P. CUVILLIER.*On temporal coherency of probabilistic models for audio-to-score alignment*, UPMC - Paris 6 Sorbonne Universités, December 2016, https://hal.inria.fr/tel-01448687.

[12] C. PONCELET SANCHEZ.*Model-Based Testing Real-Time and Interactive Music Systems*, EDITE, November 2016, https://hal.archives-ouvertes.fr/tel-01443327.

Articles in International Peer-Reviewed Journal

[13] G. BURLOIU, A. CONT, C. PONCELET.*A visual framework for dynamic mixed music notation*, in "Journal of New Music Research", 2016 [*DOI :* 10.1080/09298215.2016.1245345], https://hal.inria.fr/hal-01390502.

[14] F. JACQUEMARD, C. PONCELET.*An Automatic Test Framework for Interactive Music Systems*, in "Journal of New Music Research", 2016, vol. 45, n^{o} 2, 18, https://hal.inria.fr/hal-01274035.

[15] F. JACQUEMARD, M. RUSINOWITCH.*One-variable context-free hedge automata*, in "Journal of Computer and System Sciences", 2016 [*DOI :* 10.1016/J.JCSS.2016.10.006], https://hal.inria.fr/hal-01426626.

[16] 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^{o} 2, 32, https://hal.inria.fr/hal-00769249.

[17] J. PASCALIE, M. POTIER, T. KOWALIW, J.-L. GIAVITTO, O. MICHEL, A. SPICHER, R. DOURSAT.*Developmental Design of Synthetic Bacterial Architectures by Morphogenetic Engineering*, in "ACS Synthetic Biology", May 2016, vol. 5, n^{o} 8, p. 842-861 [*DOI :* 10.1021/ACSSYNBIO.5B00246], https://hal.inria.fr/hal-01354757.

[18] C. PONCELET, F. JACQUEMARD.*Model-Based Testing for Building Reliable Realtime Interactive Music Systems*, in "Science of Computer Programming", 2016, Journal: Science of Computer Programming special issue: SAC SVT 2015, https://hal.archives-ouvertes.fr/hal-01314969.

Invited Conferences

[19] F. JACQUEMARD.*Some results on confluence: decision and what to do without*, in "5th International Workshop on Confluence", Obergurgl, Austria, Proceedings of the 5th International Workshop on Confluence, Beniamino Accattoli, Ashish Tiwari , September 2016, https://hal.inria.fr/hal-01367535.

International Conferences with Proceedings

[20] P. DONAT-BOUILLUD, J.-L. GIAVITTO, A. CONT, N. SCHMIDT, Y. ORLAREY.*Embedding native audio-processing in a score following system with quasi sample accuracy*, in "ICMC 2016 - 42th International Computer Music Conference", Utrecht, Netherlands, September 2016, https://hal.inria.fr/hal-01349524.

[21] R. LAJUGIE, P. BOJANOWSKI, P. CUVILLIER, S. ARLOT, F. BACH.*A weakly-supervised discriminative model for audio-to-score alignment*, in "41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP)", Shanghai, China, Proceedings of the 41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP), March 2016, https://hal.archives-ouvertes.fr/hal-01251018.

[22] A. YCART, F. JACQUEMARD, J. BRESSON, S. STAWORKO.*A Supervised Approach for Rhythm Transcription Based on Tree Series Enumeration*, in "International Computer Music Conference (ICMC)", Utrecht, Netherlands, Proceedings of the 42nd International Computer Music Conference (ICMC), September 2016, https://hal.inria.fr/hal-01315689.

National Conferences with Proceeding

[23] A. YCART, J. BRESSON, F. JACQUEMARD, S. STAWORKO.*An interactive approach to rhythm transcription in OpenMusic*, in "Journées d'Informatique Musicale 2016", Albi, France, Journées d'Informatique Musicale 2016, AFIM, March 2016, https://hal.inria.fr/hal-01298806.

Scientific Books (or Scientific Book chapters)

[24] J.-L. GIAVITTO, A. SPICHER.*A Topological Approach of Musical Relationships*, in "Mathemusical Conversations - Mathematics and Computation in Music Performance and Composition", J. B. L. SMITH, E. CHEW, G.

ASSAYAG (editors), World Scientific, September 2016, vol. Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, n° 32, p. 283–298, https://hal.archives-ouvertes.fr/hal-01377659.

[25] A. SPICHER, J.-L. GIAVITTO.*Interaction-based Programming in MGS*, in "Advances in Unconventional Computing - Volume 1: Theory", A. ADAMATZKY (editor), Series: Emergence, Complexity and Computation, Springer, 2017, vol. 22, p. 305-342 [*DOI :* 10.1007/978-3-319-33924-5], https://hal.archives-ouvertes.fr/hal-01377667.

Scientific Popularization

[26] N. SCHMIDT GUBBINS, A. CONT, J.-L. GIAVITTO.*First steps toward embedding real-time audio computing in Antescofo*, in "Journal de Investigacio´n de Pregado (Investigacio´n, Interdisciplina, Innovacio´n)", 2016, vol. 6, https://hal.archives-ouvertes.fr/hal-01257524.

Other Publications

[27] F. JACQUEMARD, A. YCART, M. SAKAI.*Generating equivalent rhythmic notations based on rhythm tree languages*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01403982.

References in notes

[28] G. ASSAYAG, A. GERZSO (editors). *New Computational Paradigms for Computer Music*, Sciences de la musique, Editions Delatour, 6 2009.

[29] A. BIETTI, F. BACH, A. CONT.*An online EM algorithm in hidden (semi-)Markov models for audio segmentation and clustering*, in "ICASSP 2015 - 40th IEEE International Conference on Acoustics, Speech and Signal Processing", Brisbane, Australia, April 2015, https://hal.inria.fr/hal-01115826.

[30] A. BIETTI.*Online learning for audio clustering and segmentation*, ENS Cachan, September 2014, https://hal.inria.fr/hal-01064672.

[31] J. BRESSON, D. BOUCHE, J. GARCIA, T. CARPENTIER, F. JACQUEMARD, J. MACCALLUM, D. SCHWARZ.*Projet EFFICACE : Développements et perspectives en composition assistée par ordinateur*, in "Journées d'Informatique Musicale", Montréal, Canada, 2015, https://hal.archives-ouvertes.fr/hal-01142022.

[32] G. BURLOIU, A. CONT.*Non-overlapping, Time-coherent Visualisation of Action Commands in the AscoGraph Interactive Music User Interface*, in "First International Conference on Technologies for Music Notation and Representation", Paris, France, May 2015, https://hal.inria.fr/hal-01135727.

[33] G. BURLOIU, A. CONT.*Visualizing Timed, Hierarchical Code Structures in AscoGraph*, in "International Conference on Information Visualisation", Barcelona, Spain, University of Barcelona, July 2015, https://hal.inria.fr/hal-01155618.

[34] T. COFFY, J.-L. GIAVITTO, A. CONT.*AscoGraph: A User Interface for Sequencing and Score Following for Interactive Music*, in "ICMC 2014 - 40th International Computer Music Conference", Athens, Greece, September 2014, https://hal.inria.fr/hal-01024865.

[35] A. CONT, J. ECHEVESTE, J.-L. GIAVITTO.*The Cyber-Physical System Approach for Automatic Music Accompaniment in Antescofo*, in "Acoustical Society Of America", Providence, Rhode Island, United States, May 2014, Best Paper Award for Students and Young Presenters, https://hal.inria.fr/hal-00997842.

[36] A. CONT, J. ECHEVESTE, J.-L. GIAVITTO, F. JACQUEMARD.*Correct Automatic Accompaniment Despite Machine Listening or Human Errors in Antescofo*, in "ICMC 2012 - International Computer Music Conference", Ljubljana, Slovenia, IRZU - the Institute for Sonic Arts Research, September 2012, https://hal.inria.fr/hal-00718854.

[37] P. CUVILLIER, A. CONT.*Coherent Time Modeling of semi-Markov Models with Application to Real-Time Audio-to-Score Alignment*, in "MLSP 2014 - IEEE International Workshop on Machine Learning for Signal Processing (2014)", Reims, France, J. LARSEN, K. GUELTON (editors), IEEE, September 2014, https://hal.inria.fr/hal-01058366.

[38] P. CUVILLIER.*Time-coherency of Bayesian priors on transient semi-Markov chains for audio-to-score alignment*, in "MaxEnt 2014", Amboise, France, SEE, September 2014, https://hal.inria.fr/hal-01080235.

[39] A. DESSEIN, A. CONT.*An information-geometric approach to real-time audio segmentation*, in "IEEE Signal Processing Letters", April 2013, vol. 20, n^o 4, p. 331-334 [*DOI :* 10.1109/LSP.2013.2247039], https://hal.inria.fr/hal-00793999.

[40] A. DESSEIN.*Computational Methods of Information Geometry with Real-Time Applications in Audio Signal Processing*, Université Pierre et Marie Curie - Paris VI, December 2012, https://tel.archives-ouvertes.fr/tel-00768524.

[41] P. DONAT-BOUILLUD.*Multimedia scheduling for interactive multimedia systems*, ENS Rennes ; Universite Rennes 1, June 2015, 47, https://hal.inria.fr/hal-01168098.

[42] P. DONAT-BOUILLUD, F. JACQUEMARD, M. SAKAI.*Towards an Equational Theory of Rhythm Notation*, in "Music Encoding Conference 2015", Florence, Italy, May 2015, https://hal.inria.fr/hal-01105418.

[43] J. ECHEVESTE, A. CONT, J.-L. GIAVITTO, F. JACQUEMARD.*Operational semantics of a domain specific language for real time musician-computer interaction*, in "Discrete Event Dynamic Systems", August 2013, vol. 23, n^o 4, p. 343-383 [*DOI :* 10.1007/s10626-013-0166-2], https://hal.inria.fr/hal-00854719.

[44] J.-M. ECHEVESTE.*A programming language for Computer-Human Musical Interaction*, Université Pierre et Marie Curie - Paris VI, May 2015, https://tel.archives-ouvertes.fr/tel-01196248.

[45] J. ECHEVESTE, J.-L. GIAVITTO, A. CONT.*A Dynamic Timed-Language for Computer-Human Musical Interaction*, Inria, December 2013, n^o RR-8422, https://hal.inria.fr/hal-00917469.

[46] L. FANCHON.*Temporal Analysis of Mixed Intrumental/Electronic Music Scores*, Ecole Centrale Paris, November 2012, 51, https://hal.inria.fr/hal-00762004.

[47] L. FANCHON, F. JACQUEMARD.*Formal Timing Analysis Of Mixed Music Scores*, in "2013 ICMC - International Computer Music Conference", Perth, Australia, August 2013, https://hal.inria.fr/hal-00829821.

[48] J.-L. GIAVITTO, J. ECHEVESTE.*Real-Time Matching of Antescofo Temporal Patterns*, in "PPDP 2014 - 16th International Symposium on Principles and Practice of Declarative Programming", Canterbury, United Kingdom, ACM, September 2014 [*DOI :* 10.1145/2643135.2643158], https://hal.archives-ouvertes.fr/hal-01054667.

[49] R. GONG, P. CUVILLIER, N. OBIN, A. CONT.*Real-Time Audio-to-Score Alignment of Singing Voice Based on Melody and Lyric Information*, in "Interspeech", Dresde, Germany, September 2015, https://hal.archives-ouvertes.fr/hal-01164550.

[50] A. HESSEL, K. G. LARSEN, M. MIKUCIONIS, B. NIELSEN, P. PETTERSSON, A. SKOU.*Testing real-time systems using UPPAAL*, in "Formal methods and testing", Berlin, Heidelberg, R. M. HIERONS, J. P. BOWEN, M. HARMAN (editors), Springer-Verlag, 2008, p. 77–117, http://dl.acm.org/citation.cfm?id=1806209.1806212.

[51] F. JACQUEMARD, P. DONAT-BOUILLUD, J. BRESSON.*A Structural Theory of Rhythm Notation based on Tree Representations and Term Rewriting*, in "Mathematics and Computation in Music: 5th International Conference, MCM 2015", London, United Kingdom, D. M. TOM COLLINS, A. VOLK (editors), Lecture Notes in Artificial Intelligence, Springer, June 2015, vol. 9110, 12, https://hal.inria.fr/hal-01138642.

[52] F. JACQUEMARD, Y. KOJIMA, M. SAKAI.*Term Rewriting with Prefix Context Constraints and Bottom-Up Strategies*, in "25th International Conference on Automated Deduction (CADE'15)", Berlin, Germany, A. P. FELTY, A. MIDDELDORP (editors), LNCS, Springer, August 2015, https://hal.inria.fr/hal-01149319.

[53] M. KRICHEN, S. TRIPAKIS.*Black-box conformance testing for real-time systems*, in "In 11th International SPIN Workshop on Model Checking of Software (SPIN'04), volume 2989 of LNCS", Springer, 2004, p. 109–126.

[54] R. LAJUGIE, P. BOJANOWSKI, P. CUVILLIER, S. ARLOT, F. BACH.*A weakly-supervised discriminative model for audio-to-score alignment*, in "41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP)", Shanghai, China, Proceedings of the 41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP), March 2016, https://hal.archives-ouvertes.fr/hal-01251018.

[55] V. LOSTANLEN.*Decouverte automatique de structures musicales en temps reel par la geometrie de l'information*, ATIAM, University of Paris 6 (UPMC), September 2013, https://hal.inria.fr/hal-00849736.

[56] E. NAKAMURA, P. CUVILLIER, A. CONT, N. ONO, S. SAGAYAMA.*Autoregressive hidden semi-Markov model of symbolic music performance for score following*, in "16th International Society for Music Information Retrieval Conference (ISMIR)", Malaga, Spain, October 2015, https://hal.inria.fr/hal-01183820.

[57] J. NIKA, J. ECHEVESTE, M. CHEMILLIER, J.-L. GIAVITTO.*Planning Human-Computer Improvisation*, in "International Computer Music Conference", Athens, Greece, September 2014, 330, https://hal.archives-ouvertes.fr/hal-01053834.

[58] C. PONCELET, F. JACQUEMARD.*Model Based Testing of an Interactive Music System*, in "Proceedings of the 30th ACM/SIGAPP Symposium On Applied Computing (ACM SAC)", Salamanca, Spain, ACM, April 2015 [*DOI :* 10.1145/2695664.2695804], https://hal.archives-ouvertes.fr/hal-01097345.

[59] C. PONCELET.*Génération de tests de conformité pour un système musical interactif temps-réel*, November 2013, MSR'13, Poster, https://hal.inria.fr/hal-01133592.

[60] C. PONCELET SANCHEZ, F. JACQUEMARD.*Test Methods for Score-Based Interactive Music Systems*, in "ICMC SMC 2014", Athen, Greece, September 2014, https://hal.inria.fr/hal-01021617.

[61] R. ROWE.*Interactive music systems: machine listening and composing*, MIT Press, Cambridge, MA, USA, 1992.

[62] E. D. SCHEIRER.*Music listening systems*, MIT Media Lab, 2000, http://web.media.mit.edu/~tristan/Classes/MAS.945/Papers/Technical/Scheirer_Thesis.pdf.

[63] C. TRAPANI, J. ECHEVESTE.*Real Time Tempo Canons with Antescofo*, in "International Computer Music Conference", Athens, Greece, September 2014, 207, https://hal.archives-ouvertes.fr/hal-01053836.

[64] B. VERCOE.*The Synthetic Performer in the Context of Live Performance*, in "Proceedings of the ICMC", 1984, p. 199–200.

[65] A. YCART.*Quantification rythmique dans OpenMusic*, UPMC, September 2015, 51, https://hal.inria.fr/hal-01202257.

Project-Team MYCENAE

Multiscale dYnamiCs in neuroENdocrine AxEs

RESEARCH CENTER
Paris

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team MYCENAE

Creation of the Project-Team: 2014 January 01

Keywords:

Computer Science and Digital Science:

6.1.1. - Continuous Modeling (PDE, ODE)
6.1.2. - Stochastic Modeling (SPDE, SDE)
6.1.3. - Discrete Modeling (multi-agent, people centered)
6.1.4. - Multiscale modeling
6.2.1. - Numerical analysis of PDE and ODE
6.2.3. - Probabilistic methods
6.3.1. - Inverse problems
6.3.4. - Model reduction

Other Research Topics and Application Domains:

1.1.3. - Cellular biology
1.1.4. - Developmental biology
1.1.10. - Mathematical biology
1.3.1. - Understanding and simulation of the brain and the nervous system
2.2.2. - Nervous system and endocrinology

1. Members

Research Scientists

Frédérique Clément [Team leader, Inria, Senior Researcher, HDR]
Jonathan Touboul [Inria, Researcher, detached from Corps des Mines, HDR]

PhD Students

Richard Bailleul [CIRB]
Tanguy Cabana [UPMC, until Oct 2016]
Yi Cui [UPMC]
Elif Köksal Ersöz [Inria, until Nov 2016]
Lucile Megret [Univ. Paris VI, until Oct 2016]
Frédérique Robin [Inria, from Oct 2016]

Post-Doctoral Fellows

Soledad Fernández García [Inria, until March 2016]
Justyna Signerska-Rynkowska [Inria, until March 2016]

Administrative Assistant

Martine Verneuille [Inria, Assistant]

Others

Jean-Pierre Françoise [UPMC, Professor, HDR]
Marie Postel [UPMC, Associate Professor, HDR]
Alexandre Vidal [Univ. Évry Val d'Essonne, Associate Professor]

2. Overall Objectives

2.1. Overall Objectives

MYCENAE (Multiscale dYnamiCs in neuroENdocrine AxEs) is a project-team dedicated to mathematical neuroendocrinology and mathematical neuroscience. We are interested in the modeling, analysis and simulation of multiscale in time and/or space dynamics in the fields of neuroscience, endocrinology and physiology. Our main research topics are the followings:

- Numerical and theoretical studies of slow-fast systems with complex oscillations
- Non conservative transport equations for cell population dynamics
- Macroscopic limits of stochastic neural networks and neural fields

3. Research Program

3.1. Project team positioning

The main goal of MYCENAE is to address crucial questions arising from both Neuroendocrinology and Neuroscience from a mathematical perspective. The choice and subsequent study of appropriate mathematical formalisms to investigate these dynamics is at the core of MYCENAE's scientific foundations: slow-fast dynamical systems with multiple time scales, mean-field approaches subject to limit-size and stochastic effects, transport-like partial differential equations (PDE) and stochastic individual based models (SIBM).

The scientific positioning of MYCENAE is on the way between Mathematical Biology and Mathematics: we are involved both in the modeling of physiological processes and in the deep mathematical analysis of models, whether they be (i) models developed (or under development) within the team (ii) models developed by collaborating teams or (iii) benchmark models from the literature.

Our research program is grounded on previous results obtained in the framework of the REGATE (REgulation of the GonAdoTropE axis) Large Scale Initiative Action and the SISYPHE project team on the one hand, and the Mathematical Neuroscience Team in the Center for Interdisciplinary Research in Biology (Collège de France), on the other hand. Several of our research topics are related to the study and generalization of 2 master models: a 4D, multiscale in time, nonlinear model based on coupled FitzHugh-Nagumo dynamics that has proved to be a fruitful basis for the study of the complex oscillations in hypothalamic GnRH dynamics [34], [33], and a nD, multiscale in space, system of weakly-coupled non conservative transport equations that underlies our approach of gonadal cell dynamics [35],[7]. Most our topics in mathematical neuroscience deal with the study of complex oscillatory behaviors exhibited either by single neurons or as emergent macroscopic properties of neural networks, from both a deterministic and stochastic viewpoint.

3.2. Numerical and theoretical studies of slow-fast systems with complex oscillations

In dynamical systems with at least three state variables, the presence of different time scales favors the appearance of complex oscillatory solutions. In this context, with (at least) two slow variables MixedMode Oscillations (MMO) dynamics can arise. MMOs are small and large amplitude oscillations combined in a single time series. The last decade has witnessed a significant amount of research on this topic, including studies of folded singularities, construction of MMOs using folded singularities in combination with global dynamics, effects of additional time scales, onset of MMOs via singular Hopf bifurcations, as well as generalization to higher dimensions. In the same period, many applications to neuroscience emerged [8]. On the other hand, bursting oscillations, another prototype of complex oscillations can occur in systems with (at least) two fast variables. Bursting has been observed in many biological contexts, in particular in the dynamics of pancreatic cells, neurons, and other excitable cells. In neuronal dynamics a burst corresponds to a series of spikes, interspersed with periods of quiescent behavior, called inter-burst intervals. We are interested in systems combining bursting, MMOs and canards. One of the interesting directions is torus canards, which are canard-like structures occurring in systems combining canard explosion with fast rotation [4]. Torus canards help understand transitions from spiking or MMO dynamics to bursting. Another study on the boundary of bursting and MMOs is the work of [37] on the so-called plateau bursting. A major challenge in this direction is to gain a complete understanding of the transition from "3 time scales" to "2 fast/ 1 slow" (bursting) and then to "1 fast/ 2 slow (MMOs)". Also, a key challenge that we intend to tackle in the next few years is that of large dynamical systems with many fast and many slow variables, which additionally are changing in time and/or in phase space. We aim to pursue this research direction both at theoretical and computational level, using numerical continuation approaches based on the location of unstable trajectories by using fixed point methods, rather than simulation, to locate trajectories.

3.3. Non conservative transport equations for cell population dynamics

Models for physiologically-structured populations can be considered to derive from the so-called McKendrick-Von Foerster equation or renewal equation that has been applied and generalized in different applications of population dynamics, including ecology, epidemiology and cell biology. Renewal equations are PDE transport equations that are written so as to combine conservation laws (e.g. on the total number of individuals) with additional terms related to death or maturation, that blur the underlying overall balance law.

The development of ovarian follicles is a tightly-controlled physiological and morphogenetic process, that can be investigated from a middle-out approach starting at the cell level. To describe the terminal stages of follicular development on a cell kinetics basis and account for the selection process operated amongst follicles, we have developed a multiscale model describing the cell density in each follicle, that can be roughly considered as a system of weakly-coupled, non conservative transport equations with controlled velocities and source term. Even if, in some sense, this model belongs to the class of renewal equations for structured populations, it owns a number of specificities that render its theoretical and numerical analysis particularly challenging: 2 structuring variables (per follicle, leading as a whole to $2n$D system), control terms operating on the velocities and source term, and formulated from moments of the unknowns, discontinuities both in the velocities and density on internal boundaries of the domain representing the passage from one cell phase to another.

On the theoretical ground, the well-posedness (existence and uniqueness of weak solutions with bounded initial data) has been established in [11], while associated control problems have been studied in the framework of hybrid optimal control [5]. On the numerical ground, the formalism dedicated to the simulation of these hyperbolic-like PDEs is that of finite volume method. Part of the numerical strategy consists in combining in the most efficient way low resolution numerical schemes (such as the first-order Godunov scheme), that tend to be diffusive, with high resolution schemes (such as the Lax Wendroff second-order scheme), that may engender oscillations in the vicinity of discontinuities [2], with a critical choice of the limiter functions. The 2D finite volume schemes are combined with adaptive mesh refinement through a multi-resolution method [3] and implemented in a problem-specific way on parallel architecture [1].

3.4. Macroscopic limits of stochastic neural networks and neural fields

The coordinated activity of the cortex is the result of the interactions between a very large number of cells. Each cell is well described by a dynamical system, that receives non constant input which is the superposition of an external stimulus, noise and interactions with other cells. Most models describing the emergent behavior arising from the interaction of neurons in large-scale networks have relied on continuum limits ever since the seminal work of Wilson and Cowan and Amari [38], [32]. Such models tend to represent the activity of the network through a macroscopic variable, the population-averaged firing rate.

In order to rationally describe neural fields and more generally large cortical assemblies, one should yet base their approach on what is known of the microscopic neuronal dynamics. At this scale, the equation of the activity is a set of stochastic differential equations in interaction. Obtaining the equations of evolution of the effective mean-field from microscopic dynamics is a very complex problem which belongs to statistical physics. As in the case of the kinetic theory of gases, macroscopic states are defined by the limit of certain quantities as the network size tends to infinity. When such a limit theorem is proved, one can be ensured that large networks are well approximated by the obtained macroscopic system. Qualitative distinctions between the macroscopic limit and finite-sized networks (finite-size effects), occurs in such systems. We have been interested in the relevant mathematical approaches dealing with macroscopic limits of stochastic neuronal networks, that are expressed in the form of a complex integro-differential stochastic implicit equations of McKean-Vlasov type including a new mathematical object, the spatially chaotic Brownian motion [14].

The major question consists in establishing the fundamental laws of the collective behaviors cortical assemblies in a number of contexts motivated by neuroscience, such as communication delays between cells [13], [12] or spatially extended areas, which is the main topic of our current research. In that case additional difficulties arise, since the connection between different neurons, as well as delays in communications, depend on space in a correlated way, leading to the singular dependence of the solutions in space, which is not measurable.

4. Application Domains

4.1. Introduction

MYCENAE addresses rather “upstream” questions in neuroendocrinology and neuroscience. Nevertheless, MYCENAE’s expected results can contribute to more applied issues in these fields, mainly by helping understand the mechanisms underlying physiological and pathological processes and also by designing new concepts for biomedical data analysis. MYCENAE thematics are related to societal issues concerning endocrine disruptors, reproductive biotechnologies, and neurological diseases, especially in case of pathological synchronizations encountered in epilepsy and Parkinson’s disease.

4.2. Neuroendocrinology and Neuroscience

We are interested in the complex dynamical processes arising within neuroendocrine axes, with a special focus on the reproductive (hypothalamo-pituitary-gonadal) axis. This axis can be considered as the paragon of neuroendocrine axes, since it both concentrates all remarkable dynamics that can be exhibited by these axes and owns its unique specificities, as gonads are the only organs that host germ cells. Since, in neuroendocrine axes, neural systems are embedded within endocrine feedback loops and interact with peripheral organs, one also needs to get interested in the peripheral dynamics to be able to “close the loop” and account for the effect of peripheral inputs on neural dynamics. In the case of the HPG axis, these dynamics are especially complex, because they involve developmental processes that occur even in adult organisms and combine the glandular function of the gonads with their gametogenic function.

Neuroendocrinology is thus a scientific field at the interface between Neuroscience, Endocrinology and Physiology (and even of Developmental Biology in the case of the HPG axis). On a neuroscience ground, mathematical neuroendocrinology is specifically interested in endocrine neurons, which have the uncommon ability of secreting neurohormones into the blood stream. Neuroendocrine networks are characterized by the emergence of very slow rhythms (on the order of an hour), finite size effects due to their relative small number of neurons (on the order of a few thousands for the Gonadotropin-Releasing-Hormone network) and neuroanatomical particularities, that impact the way they can synchronize and desynchronize. On a physiological ground, gonadal cell biology raises specific cell biology issues on more than one account. First, the gonads are the only organs sheltering the germ cell lines (corresponding to oogenesis in ovaries and spermatogenesis in testes). Hence, the two modes of cell division, mitosis and meiosis are encountered in these tissues. Second, there are intricate interactions between the gonadal somatic cells (granulosa cells in the ovaries, sertoli cells in the testes) and the germ cells. Third, the control of gonadal cell populations is exerted within endocrine feedback loops involving both the hypothalamus and pituitary, which results naturally in multiscale population dynamics coupled with hormonally-controlled cell kinetics.

MYCENAE's research topics in mathematical neuroscience deal with complex oscillations, synchronization and plasticity.

We study (i) the emergence of network-level behaviors from individual dynamics of excitable cells (mainly neurons, but not exclusively, as the pituitary cells belong to the family of excitable cells): complete synchronization or synchronization of specific events, effect of the recruitment rate in the synchronization process, dependence on the neuro-anatomical and functional coupling properties; (ii) the control of the different possible configurations of the network depending on external (e.g. daylength) and/or internal inputs (e.g. metabolic status), at the source of plasticity processes in cognitive (vision learning) or neuroendocrine systems (differential sensitivity to gonadal steroids and peptides across the different steps of the reproductive life); (iii) the encoding of neuro-hormonal signals as complex oscillations, on the electrical, ionic (calcium dynamics) and secretory levels; and (iv) the decoding of these signals by their target neuronal or non-neuronal cells.

More recently, we have been interested into developmental biology issues in neurosciences: neurogenesis and brain development. The anatomical and functional organization of the nervous system, and especially the brain, is highly structured and tightly regulated. The surface of the cortex, its thickness, but also the size and shape of the brain areas associated to the different sensory or motor areas are very reliable quantities across different individuals. In collaboration with different teams of biologists, we develop and investigate models of the development of the brain, at different time and spatial scale.

The biological relevance of our modeling and model-based signal analysis approaches is grounded on our network of collaborations with teams of experimentalist biologists. In particular, we have long standing collaborations with the UMR 6175 (INRA-CNRS-Université François Rabelais-Haras Nationaux) "Physiologie de la Reproduction et des Comportements" that covers most our research topics in reproductive neuroendocrinology. We have especially close links with the Bingo (Integrative Biology of the ovary) and Bios (Biology and Bioinformatics of Signaling Systems) teams, which were partners of the REGATE LSIA. We have been jointly investigating issues relative to terminal or basal follicular development [6], [7], analysis of neurosecretory patterns [15] and modeling of GPCR (G-Protein Coupled Receptors) signaling networks [9]. We also have special links with the Center for Interdisciplinary Research in Biology (CIRB, Collège de France), headed by Alain Prochiantz, that help us get a better understanding of how the brain connectivity develops and how it is functionally organized. An instance of a recent collaborative work is the study of the organization of spatial frequencies in the primary visual cortex [36].

5. Highlights of the Year

5.1. Highlights of the Year

- PhD defense of Lucile Megret. Explosion of limit cycles : qualitative analysis, numerical simulations and models. Université Pierre & Marie Curie – Sorbonne Universités, November 25th 2016.
- PhD defense of Elif Köksal Ersöz. A mathematical study on coupled multiple timescale systems, synchronization of populations of endocrine neurons. Université Pierre & Marie Curie – Sorbonne Universités, December 13th 2016.
- PhD defense of Tanguy Cabana. Limits of randomly connected networks and their dynamics. Université Pierre & Marie Curie – Sorbonne Universités, December 14th 2016.
- Invited plenary conference at ICAR2016 http://www.icar2016.org 18th International Congress on Animal Reproduction. Multiscale mathematical modeling of the hypothalamo-pituitary-gonadal axis. Tours (France), June 26-30th 2016.

6. New Software and Platforms

6.1. DynPeak

KEYWORDS: Biology - Health - Physiology

SCIENTIFIC DESCRIPTION

DynPeak is an algorithm for pulse detection and frequency analysis in hormonal time series. A new release of the DynPeak Scilab atom toolbox has been delivered in 2016 https://atoms.scilab.org/toolboxes/Dynpeak/2.1.0

- Participants: Frédérique Clement, Serge Steer, Thierry Martinez
- Partner: INRA
- Contact: Frédérique Clement
- URL: https://team.inria.fr/mycenae/en/software/

7. New Results

7.1. Numerical and theoretical studies of slow-fast systems with complex oscillations

7.1.1. Coupled multiple timescale dynamics in populations of endocrine neurons: Pulsatile and surge patterns of GnRH secretion

Participants: Elif Köksal Ersöz, Alexandre Vidal, Frédérique Clément.

The gonadotropin releasing hormone (GnRH) is secreted by hypothalamic neurons into the pituitary portal blood in a pulsatile manner. The alternation between a frequency-modulated pulsatile regime and the ovulatory surge is the hallmark of the GnRH secretion pattern in ovarian cycles of female mammals. In this work, we aimed at modeling additional features of the GnRH secretion pattern: the possible occurrence of a two-bump surge ("camel surge") and an episode of partial desynchronization before the surge.

We have proposed a six-dimensional extension of a former four-dimensional model with three timescale and introduced two mutually-coupled, slightly heterogenous GnRH subpopulations (secretors) regulated by the same slow oscillator (regulator). We have considered two types of coupling functions between the secretors, including dynamic state-dependent coupling, and we have used numerical and analytic tools to characterize the coupling parameter values leading to the generation of a two-bump surge in both coupling cases. We have revealed the impact of the slowly varying control exerted by the regulator onto the pulsatile dynamics of the secretors, which leads to dynamic bifurcations and gives rise to desynchronization. To assess the occurrence time of desynchronization during the pulsatile phase, we have introduced asymptotic tools based on quasi-static and geometric approaches, as well as analytic tools based on the H-function derived from phase equation and numerical tracking of period-doubling bifurcations. We discuss the role of coupling parameters in the two-bump surge generation and the speed of desynchronization.

7.1.2. Symmetric coupling of multiple timescale systems with mixed-mode oscillations

Participants: Soledad Fernández García, Alexandre Vidal, Fabrizio de Vico Fallani [EPI Aramis], Frédérique Clément.

We have analyzed a six-dimensional slow-fast system consisting of two coupled identical oscillators. Each oscillator is a three-dimensional system consisting of a FitzHugh-Nagumo system with an additional variable representing the calcium concentration. Individually, each three-dimensional subsystem possesses an attractive Mixed-Mode oscillations limit cycle, displaying small oscillations due to the presence of a folded saddle-node type II singularity for a certain range of the parameters values. We have considered a linear coupling through the fast variable in the slow equation and study the synchronization patterns of two identical systems with identical coupling parameter. Apart from stable in-phase and stable anti-phase synchronization patterns, the system presents almost-in-phase synchronization, oscillation death of one of the oscillators and total oscillation death, intertwined with complex transitions involving period doubling cascade, period adding phenomena and chaos. We have pointed out the role of Mixed-Mode oscillations in the birth of the different patterns and the transitions from one regime to another.

Part of these results have been presented as a contributed talk to the SIAM conference on life science https://www.siam.org/meetings/ls16/: (A Study of the Synchronization Between Two Coupled Neuron Models Generating Mixed-Mode Oscillations. A. Vidal, S.Fernández García, F. Clément, F. De Vico Fallani). MS48 Applications of Multiple Time Scale Dynamics in Biological Systems.

7.1.3. 3D-Explosion of cycles and spike-adding in the Hindmarsh-Rose model

Participants: Lucile Megret, Mathieu Desroches [Sophia], Jean-Pierre Françoise, Maciej Krupa [Sophia].

We have considered slow-fast systems that feature bursting oscillations, the minimal configuration being two fast variables and one slow variable. In the Hindmarsh-Rose model, as the slow variable z evolves, the fast dynamics undergoes several bifurcations (two Hopf bifurcations, two homoclinic bifurcations, two focus-node and two saddle-node bifurcations). We have focused on the existence of a sequence of 3D-candidate limit periodic sets of a new type. Numerical simulations have shown that it generates for the full 3D-dynamics and (the small parameter) " small enough a 3D-explosion of cycles. We have discussed the relation between this 3D-explosion and the spike-adding. We have also emphasized another new phenomenon induced by the slow-crossing of a saddle-node bifurcation with solutions which after coming close to the fold point, continue to follow it along its non-hyperbolic center manifold. We have shown how this phenomenon is also involved in the spike-adding mechanism taking place in square-wave bursters such as the Hindmarsh-Rose system.
Part of these results have been presented at the “36e Séminaire de la Sociéte Francophone de Biologie théorique”, St-Flour (France), June 12-15 2016.

7.1.4. Wild oscillations in a nonlinear neuron model with resets

Participants: Jonathan Rubin [University of Pittsburgh], Justyna Signerska-Rynkowska, Jonathan Touboul, Alexandre Vidal.

In a series of two studies, we have investigated the mechanisms by which complex oscillations are generated in a class of nonlinear dynamical systems with resets modeling the voltage and adaptation of neurons.

The first study [30] presents a mathematical analysis showing that the system can support bursts of any period as a function of model parameters, and that are organized in a period-incrementing structure. In continuous dynamical systems with resets, such period-incrementing structures are complex to analyze. In the present context, we have used the fact that bursting patterns correspond to periodic orbits of the adaptation map that governs the sequence of values of the adaptation variable at the resets. Using a slow-fast approach, we have shown that this map converges towards a piecewise linear discontinuous map whose orbits are exactly characterized. That map shows a period-incrementing structure with instantaneous transitions. We have further shown that the period-incrementing structure persists for the full system with non-constant adaptation, yet the transitions are more complex. We have also established the presence of chaos at the transitions.

The second study [31] shows that these neuron models can generically display a form of mixed-mode oscillations (MMOs), which are trajectories featuring an alternation of small oscillations with spikes or bursts (multiple consecutive spikes). The mechanism by which these are generated relies fundamentally on the hybrid structure of the flow: invariant manifolds of the continuous dynamics govern small oscillations, while discrete resets govern the emission of spikes or bursts, contrasting with classical MMO mechanisms in ordinary differential equations involving more than three dimensions and generally relying on a timescale separation. The decomposition of mechanisms reveals the geometrical origin of MMOs, allowing a relatively simple classification of points on the reset manifold associated to specific numbers of small oscillations. We have shown that the MMO pattern can be described through the study of orbits of a discrete adaptation map, which is singular as it features discrete discontinuities with unbounded left- and right-derivatives. We have studied the orbits of the map via rotation theory for circle maps and elucidated in detail complex behaviors arising in the case where MMOs display a single small oscillation per cycle.

7.1.5. Canard Explosions in delay differential equations

Participants: Jonathan Touboul, Maciej Krupa [Sophia].

We have analyzed in [21] canard explosions in delayed differential equations with a one-dimensional slow manifold. This study is applied to explore the dynamics of the van der Pol slow-fast system with delayed self-coupling. In the absence of delays, this system provides a canonical example of a canard explosion. We have shown that as the delay is increased a family of "classical" canard explosions ends as a Bogdanov-Takens bifurcation occurs at the folds points of the S-shaped critical manifold.

7.2. Non conservative transport equations for cell population dynamics

7.2.1. Dimensional reduction of a multiscale model based on long time asymptotics

Participants: Frédérique Clément, Frédéric Coquel [CMAP], Marie Postel, Kim Long Tran.

We have considered a class of kinetic models for which a moment equation has a natural interpretation. We have shown that, depending on their velocity field, some models lead to moment equations that enable one to compute monokinetic solutions economically. We have detailed the example of a multiscale structured cell population model, consisting of a system of 2D transport equations. The reduced model, a system of 1D transport equations, is obtained from computing the moments of the 2D model with respect to one variable. The 1D solution is defined from the solution of the 2D model starting from an initial condition that is a Dirac mass in the direction removed by reduction. For arbitrary initial conditions, we have compared 1D and 2D model solutions in asymptotically large time. Finite volume numerical approximations of the 1D reduced model can be used to compute the moments of the 2D solution with proper accuracy, both in the conservative and non conservative framework. The numerical robustness is studied in the scalar case, and a full scale vector case is presented [29].

These results have been partly presented in a workshop on "Asymptotic behavior of systems of PDEs arising in physics and biology : theoretical and numerical points of view" (ABPDE II), Lille, June 15-17, 2016.

7.2.2. Analysis of the asymptotic behavior of a model for the morphogenesis in ovarian follicles

Participants: Frédérique Clément, Frédérique Robin, Romain Yvinec [INRA].

We have designed and analyzed a simplified version of our multiscale model for the morphogenesis of ovarian follicles [6]. We have formulated both a stochastic model, in the framework of branching processes, and a deterministic one, in the framework of nonconservative transport equations. The simplifications result in linear models, in which the oocyte growth is uncoupled from the proliferation of the surrounding follicular cells. The cell population is distributed into concentric layers around the oocyte, and structured according to the cell age. Cells are subject to the process of cell division, which resets their age and allow them to possibly move to the adjacent outer layer. Since there is no symmetry in the cell displacements (the only allowed cell motion is centrifugal), we have faced the problem of the model irreducibility. To study the asymptotic behavior, we thus had to adapt the classical results based on entropy or the computation of stochastic moments. We have proved that there is, as expected, an exponential asymptotic growth led by a Malthus parameter, which can be computed analytically in the simplest (Markovian) case, or numerically. Interestingly, the value of this global parameter merges with one of the local Malthus-like parameters defined on the layer level. In both the deterministic and stochastic cases, we could derive accurate information on the time-varying mean cell number per layer and we also got additional information on the asymptotic age distribution.

This work has been undergone in the framework of the master thesis of Frédérique Robin (M2 Mathématiques du Vivant, Université Paris-Saclay), and pursued as a PhD subject. Preliminary results have been the matter of a presentation during the "Journées INRA-Inria" held in Mallemort (France) on October 6-7th: F. Clément, F. Robin, R Yvinec. Dynamiques de populations cellulaires structurées individus-centrées : Morphogenèse des follicules ovariens.

7.2.3. Numerical study of a mathematical model for the dynamics of progenitor cell populations in the mouse cerebral cortex

Participants: Marie Postel, Alice Karam [IBPS], Frédérique Clément, Sylvie Schneider-Maunoury [IBPS].

We have studied numerically our multi-scale mathematical model of structured cell populations during the development of cerebral cortex. The model accounts for three main cell types: apical progenitors (APs), intermediate progenitors (IPs), and neurons. Each cell population is structured according to the cell age distribution. Since the model describes the different phases of the cell division cycle, we could derive the numeric equivalents of many of the experimental indexes measured in experimental setups, including classical mitotic or labeling indexes targeting the cells in phase S or mitosis, and more elaborated protocols based on double labeling with fluorescent dyes. We have formulated a multi-criterion objective function which enables us to combine experimental observations of different nature and to fit the data already acquired in the framework of the NeuroMathMod project (Sorbonne-Universités Émergence call with IBPS, Institut de Biologie Paris Seine). With the retrieved parameters, the model can provide useful information not supplied by the data, such as the cell origin of neurons (direct neurogenesis from AP or IPgenic neurogenesis) and the proportion of IPs cells undergoing several rounds of cell cycles.

7.3. Macroscopic limits of stochastic neural networks and neural fields

7.3.1. Limit theorems and effective dynamics

Participants: Jonathan Touboul, Philippe Robert [EPI RAP], Cristobal Quiñinao [IMT], Stéphane Mischler [CEREMADE].

We have pursued our investigations on the dynamics of large-scale neural networks modeling the brain, in two main directions:

We have studied in [26] the mean-field limit and stationary distributions of a pulse-coupled network modeling the dynamics of a large neuronal assemblies. Our model takes into account explicitly the intrinsic randomness of firing times, contrasting with the classical integrate-and-fire model. The ergodicity properties of the Markov process associated with finite networks have been investigated. We have derived the limit in distribution of the sample path of the state of a neuron of the network when its size gets large. The invariant distributions of this limiting stochastic process have been analyzed as well as their stability properties. We have shown that the system undergoes transitions as a function of the averaged connectivity parameter, and can support trivial states (where the network activity dies out, which is also the unique stationary state of finite networks in some cases) and self-sustained activity when connectivity level is sufficiently large, both being possibly stable.

We have investigated in [23] existence and uniqueness of solutions of a McKean-Vlasov evolution PDE representing the macroscopic behavior of interacting Fitzhugh-Nagumo neurons. This equation is hypoelliptic, nonlocal and has unbounded coefficients. We have proven the existence of a solution to the evolution equation and non trivial stationary solutions. Moreover, we have demonstrated the uniqueness of the stationary solution in the weakly nonlinear regime. Eventually, using a semigroup factorisation method, we have shown exponential nonlinear stability in the small connectivity regime.

7.3.2. Spectrum of random matrices

Participants: Jonathan Touboul, Gilles Wainrib [ENS], Luis Carlos Garcia Del Molino [New-York University], Khashayar Pakdaman [IJM].

We have considered in [20] the ensemble of Real Ginibre matrices with a positive fraction $\alpha > 0$ of real eigenvalues. We have demonstrated a large deviation principle for the joint eigenvalue density of such matrices and we have introduced a two phase log-gas whose stationary distribution coincides with the spectral measure of the ensemble. Using these tools we have provided an asymptotic expansion for the probability $p^n_{\alpha n}$ that an $n \times n$ Ginibre matrix has $k = \alpha n$ real eigenvalues and we have characterized the spectral measures of these matrices.

7.4. Modeling of brain development and brain functions

7.4.1. Organization of the visual cortex

Participants: Jonathan Touboul, Jérôme Ribot [CIRB], Alberto Romagnoni [ENS], Daniel Bennequin [IMG-PRG], Chantal Milleret [CIRB].

In the early visual cortex, information is processed within functional maps whose layout is thought to underlie visual perception. However, the precise organization of these functional maps as well as their interrelationships remains unresolved. We have investigated using new data acquisition and analysis as well as mathematical modeling, the inter-relationship between different visual maps in cat visual cortex.

We have shown in [25] that spatial frequency representation in cat areas 17 and 18 exhibits singularities around which the map organizes like an electric dipole potential. These singularities are precisely co-located with singularities of the orientation map: the pinwheel centers. We have first shown, using high resolution optical imaging, that a large majority (around 80%) of pinwheel centers exhibit in their neighborhood semi-global extrema in the spatial frequency map. These extrema create a sharp gradient that was confirmed with electrophysiological recordings. Based on an analogy with electromagnetism, a mathematical model of a dipolar structure has been proposed and accurately fitted to optical imaging data for two third of pinwheel centers with semi-global extrema. We have concluded that more than half of orientation pinwheel centers form spatial frequency dipoles in cat early visual cortex.

We have demonstrated mathematically in [27] that two natural principles, local exhaustivity of representation and parsimony, would constrain the orientation and spatial frequency maps to display co-located singularities around which the orientation is organized as a pinwheel and spatial frequency as a dipole. We have further focused on the theoretical implications of this structure. Using a computational model, we have shown that this architecture allows a trade-off in the local perception of orientation and spatial frequency, but this would occur for sharper selectivity than the tuning width reported in the literature. We therefore re-examined physiological data and have shown that indeed the spatial frequency selectivity substantially sharpens near maps singularities, bringing to the prediction that the system tends to optimize balanced detection between different attributes.

7.4.2. Modeling the timing of neurogenesis and control of the neuron pool : Enhanced abventricular proliferation compensates cell death in the embryonic cerebral cortex

Participants: Betty Freret-Hodara [IJM], Yi Cui, Amélie Griveau [IJM], Lisa Vigier [IJM], Yoko Arai [IJM], Jonathan Touboul, Alessandra Pierani [IJM].

Loss of neurons in the neocortex is generally thought to result in a final reduction of cerebral volume. Yet, little is known on how the developing cerebral cortex copes with death of early-born neurons. We have tackled this issue by taking advantage of a transgenic mouse model in which, from early embryonic stages to mid-corticogenesis, abundant apoptosis is induced in the postmitotic compartment. Unexpectedly, the thickness of the mutant cortical plate at E18.5 was normal, due to an overproduction of upper layer neurons at E14.5. We have developed and simulated a mathematical model to investigate theoretically the recovering capacity of the system and found that a minor increase in the probability of proliferative divisions of intermediate progenitors (IPs) is a powerful compensation lever. Combined with our experimental observations, these results illustrate the remarkable plasticity of neocortical progenitors to adapt to major embryonic insults via the modulation of abventricular divisions thereby ensuring the production of an appropriate number of neurons.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

Jonathan Touboul is member of the Kibord (KInetic models in Biology Or Related Domains) project obtained in 2014.

He is also PI of the projects “Mathematical modeling of synaptic plasticity” (with Laurent Venance, CIRB) funded as an interdisciplinary structuring project of INSB (Institut des Sciences Biologiques in CNRS) and “Altering Fear Memory” (with Sidney Wiener, CIRB and Karim Benchenane, ESPCI) funded by the PSL Labex MemoLife.

8.1.2. National Networks

- GdR REPRO (F. Clément is member of the direction board)
- MIA REM network: Réduction de modèles (PI Béatrice Laroche, INRA Jouy)

8.1.3. National Collaborations

- Center for Interdisciplinary Research in Biology (CIRB), Collège de France (Alain Prochiantz, Marie Manceau, Laurent Venance)
- UMR Physiologie de la Reproduction et des Comportements, INRA Centre- Val de Loire (Bios and Bingo teams)
- Université Pierre & Marie Curie (UPMC)
  - Jacques-Louis Lions Laboratory, Pierre & Marie Curie University (Jean-Pierre Françoise, Marie Postel)
  - Developmental Biology Laboratory, Institut de Biologie Paris Seine (IBPS), Pierre & Marie Curie University (Alice Karam, Sylvie Schneider Maunoury), in the framework of the NeuroMathMod, Sorbonne-Universités Émergence call
- Jacques Monod Institute (IJM)
  - Computational Biology and Biomathematics (Khashayar Pakdaman)
  - Génétique et développement du cortex cérébral (Alessandra Pierani)
- Centre de Recherche en Mathématiques de la Décision (CEREMADE), Paris Dauphine University (Stéphane Mischler)
- Unité de Neurosciences, Information & Complexité (UNIC), CNRS Gif-sur-Yvette (Alain Destexhe)

8.2. International Research Visitors

8.2.1. Visits to International Teams

Jonathan Touboul has visited Simon Levin in Princeton University (December 15-26)

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Colloquium on Dynamical Systems and Applications, May 19th 2016

Thematic session co-organized by Jonathan Touboul and Khashayar Pakdaman within the framework of the CIRB

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

Jonathan Touboul was member of the program committee of ICMNS 2016 (International Conference on Mathematical Neuroscience, Juan-les-Pins)

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Jonathan Touboul participates in the editorial boards of *Plos One* and *Frontiers in neuronal circuits*

9.1.3.2. Reviewer - Reviewing Activities

Annals of Applied Probability, Journal of Statistical Physics, eLife, PloS Computational Biology, SIAM Journal on Applied Dynamical Systems, SIAM Journal on Discrete Mathematics, Bulletin of Mathematical Biology

9.1.4. Invited Talks

Invited plenary conference of Frédérique Clément at ICAR2016 http://www.icar2016.org 18th International Congress on Animal Reproduction (over 900 attendees). Multiscale mathematical modeling of the hypothalamo-pituitary-gonadal axis. Tours (France) June 26-30th 2016.

9.1.5. Scientific Expertise

Frédérique Clément belongs to the expert board of the BCDE (Cell Biology, Development and Evolution) ITMO (Multi OrganizationThematic Institute) of the French National Alliance for Life and Health Sciences Aviesan.

Jonathan Touboul has been reviewer for the ANR and European Research Council in 2016.

9.2. Teaching - Supervision - Juries

9.2.1. Supervision

PhD in progress : Richard Bailleul. Modeling of the developmental mechanisms underlying the formation of color and appendage patterns in birds, since September 2015. Université Pierre & Marie Curie (ED515), supervisors: Benoît Perthame, Marie Manceau and Jonathan Touboul (funded by the ERC starting grant of Marie Manceau)

PhD in progress: Yi Cui. Role of Pax6 in neurodevelopment: experiments and models, since September 2014, Université Pierre & Marie Curie (ED158), supervisors: Jonathan Touboul, Alain Prochiantz and Alessandra Pierani

PhD in progress: Frédérique Robin. Multiscale modeling of the morphodynamics in ovarian follicles, since October 2016, Université Pierre & Marie Curie (ED386), supervisors: Frédérique Clément and Romain Yvinec (INRA)

PhD: Tanguy Cabana. Limits of randomly connected networks and their dynamics. Defended on December14th 2016, Université Pierre & Marie Curie (ED386), supervisors: Raphaël Krikorian, Jonathan Touboul

PhD: Elif Köksal Ersöz. A mathematical study on coupled multiple timescale systems, synchronization of populations of endocrine neurons. Defended on December 13th, Université Pierre & Marie Curie (ED386), supervisors: Frédérique Clément and Jean-Pierre Françoise, with the involvement of Mathieu Desroches

PhD: Lucile Megret, Explosion of limit cycles : qualitative analysis, numerical simulations and models. Defended on November 25th, Université Pierre & Marie Curie (ED386), supervisors: Jean-Pierre Françoise and Frédérique Clément, with the involvement of Mathieu Desroches

HDR: Alexandre Vidal. From qualitative analysis of complex dynamics to parameter estimation in neuronal models. Université d'Évry-Val-d'Essonne – Université Paris-Saclay, December 14th 2016

9.2.2. Juries

Jonathan Touboul participated in the PhD committee of Takafumi Arakaki (ED3C, supervisors: D. Hansel and A. Leblois), in the selection committee for the hiring of a professor at Technische Universitat Berlin, as well as in the Bernstein Award for Computational Neuroscience committee.

9.3. Popularization

Jonathan Touboul has given a presentation in the framework of the “demi-heure de science” : Mathematical exploration of the brain activity (January 7th).

Frédérique Clément has given a 3h lecture on “Multiscale modeling of folliculogenesis in mammals” in the M2 master “Predictive & integrative animal biology” (PRIAM) of Université Paris-Saclay.

10. Bibliography

Major publications by the team in recent years

[1] B. AYMARD, F. CLÉMENT, F. COQUEL, M. POSTEL.*Numerical simulation of the selection process of the ovarian follicles*, in "ESAIM Proc.", 2012, vol. 28, p. 99-117.

[2] B. AYMARD, F. CLÉMENT, F. COQUEL, M. POSTEL.*A numerical method for cell dynamics; kinetic equations with discontinuous coefficients*, in "SIAM J. Sci. Comput.", 2013, vol. 35, p. A2442-A2468.

[3] B. AYMARD, F. CLÉMENT, M. POSTEL.*Adaptive mesh refinement strategy for a non conservative transport problem*, in "ESAIM Math. Model. Numer. Anal.", 2014, vol. 48, n° 5, p. 1381-1412.

[4] J. BURKE, M. DESROCHES, A. BARRY, T. KAPER, M. KRAMER.*A showcase of torus canards in neuronal bursters*, in "J. Math. Neurosci.", 2012, vol. 2.

[5] F. CLÉMENT, J.-M. CORON, P. SHANG.*Optimal control of cell mass and maturity in a model of follicular ovulation*, in "SIAM J. Control Optim.", 2013, vol. 51, n° 2, p. 824-847.

[6] F. CLÉMENT, P. MICHEL, D. MONNIAUX, T. STIEHL.*Coupled somatic cell kinetics and germ cell growth: multiscale model-based insight on ovarian follicular development*, in "Multiscale Model. Simul.", 2013, vol. 11, n° 3, p. 719-746.

[7] F. CLÉMENT, D. MONNIAUX.*Multiscale modelling of follicular selection*, in "Prog. Biophys. Mol. Biol.", 2013, vol. 113, p. 398-408.

[8] M. DESROCHES, J. GUCKENHEIMER, B. KRAUSKOPF, C. KUEHN, H. OSINGA, M. WECHSELBERGER.*Mixed-mode oscillations with multiple time scales*, in "SIAM Rev.", 2012, vol. 54, p. 211–288.

[9] D. HEITZLER, G. DURAND, A. RIZK, S. AHN, J. KIM, J. VIOLIN, L. DUPUY, C. GAUTHIER, V. PIKETTY, P. CRÉPIEUX, A. POUPON, F. CLÉMENT, F. FAGES, R. LEFKOWITZ, E. REITER.*Competing G protein-coupled receptor kinases balance G protein and β-arrestin signaling*, in "Mol. Syst. Biol.", 2012, vol. 8, n° 590.

[10] M. KRUPA, A. VIDAL, F. CLÉMENT.*A network model of the periodic synchronization process in the dynamics of calcium concentration in GnRH neurons*, in "J. Math. Neurosci.", 2013, vol. 3, 4.

[11] P. SHANG.*Cauchy problem for multiscale conservation laws: Application to structured cell populations*, in "J. Math. Anal. Appl.", 2013, vol. 401, n° 2, p. 896-920.

[12] J. TOUBOUL.*Limits and dynamics of stochastic neuronal networks with random delays*, in "J. Stat. Phys.", 2012, n^{o} 149, p. 569–597.

[13] J. TOUBOUL.*Mean-Field equations for stochastic firing-rate neural fields with delays: derivation and noise-induced transitions*, in "Phys. D", 2012, vol. 241, p. 1223–1244.

[14] J. TOUBOUL.*Propagation Of Chaos In Neural Fields*, in "Ann. Appl. Probab.", 2014, vol. 24, n^{o} 3, p. 1298–1327.

[15] A. VIDAL, Q. ZHANG, C. MÉDIGUE, S. FABRE, F. CLÉMENT.*DynPeak: An algorithm for pulse detection and frequency analysis in hormonal time series*, in "PloS One", 2012, vol. 7, e39001.

Publications of the year

Articles in International Peer-Reviewed Journal

[16] B. AYMARD, F. CLÉMENT, D. MONNIAUX, M. POSTEL.*Cell-Kinetics Based Calibration of a Multiscale Model of Structured Cell Populations in Ovarian Follicles*, in "SIAM Journal on Applied Mathematics", 2016, vol. 76, n^{o} 4, p. 1471–1491 [*DOI :* 10.1137/15M1030327], https://hal.archives-ouvertes.fr/hal-01186381.

[17] F. CLÉMENT.*Multiscale mathematical modeling of the hypothalamo-pituitary-gonadal axis*, in "Theriogenology", July 2016, vol. 86, n^{o} 1, p. 11-21 [*DOI :* 10.1016/J.THERIOGENOLOGY.2016.04.063], https://hal.inria.fr/hal-01334304.

[18] S. FERNÁNDEZ-GARCÍA, M. KRUPA, F. CLÉMENT.*Mixed-Mode Oscillations in a piecewise linear system with multiple time scale coupling*, in "Physica D: Nonlinear Phenomena", July 2016, vol. 332, p. 9–22 [*DOI :* 10.1016/J.PHYSD.2016.06.002], https://hal.inria.fr/hal-01342978.

[19] B. FRERET-HODARA, Y. CUI, A. GRIVEAU, L. VIGIER, Y. ARAI, J. TOUBOUL, A. PIERANI.*Enhanced Abventricular Proliferation Compensates Cell Death in the Embryonic Cerebral Cortex*, in "Cerebral Cortex", September 2016, https://hal.archives-ouvertes.fr/hal-01412093.

[20] L. C. GARCIA DEL MOLINO, K. PAKDAMAN, J. TOUBOUL, G. WAINRIB.*The real Ginibre ensemble with k=O(n) real eigenvalues*, in "Journal of Statistical Physics", 2016, vol. 163, n^{o} 2, p. 303-323, https://hal.archives-ouvertes.fr/hal-01412352.

[21] M. KRUPA, J. TOUBOUL.*Canard explosion in delayed equations with multiple timescales*, in "Journal of Dynamics and Differential Equations", 2016, vol. 28, n^{o} 2, p. 471-491, https://hal.archives-ouvertes.fr/hal-01253412.

[22] E. KÖKSAL ERSÖZ, M. DESROCHES, M. KRUPA, F. CLÉMENT.*Canard-Mediated (De)Synchronization in Coupled Phantom Bursters*, in "SIAM Journal on Applied Dynamical Systems", March 2016, vol. 15, n^{o} 1, p. 580–608 [*DOI :* 10.1137/15M101840X], https://hal.inria.fr/hal-01256389.

[23] S. MISCHLER, C. QUIÑINAO, J. TOUBOUL.*On a Kinetic Fitzhugh–Nagumo Model of Neuronal Network*, in "Communications in Mathematical Physics", 2016, vol. 342, n^{o} 3, p. 1001–1042 [*DOI :* 10.1007/S00220-015-2556-9], https://hal.archives-ouvertes.fr/hal-01108872.

[24] D. MONNIAUX, P. MICHEL, M. POSTEL, F. CLÉMENT.*Multiscale modeling of ovarian follicular development: From follicular morphogenesis to selection for ovulation*, in "Biology of the Cell", June 2016, vol. 108, n^o 6, p. 1-12 [*DOI :* 10.1111/BOC.201500087], https://hal.inria.fr/hal-01294630.

[25] J. RIBOT, A. ROMAGNONI, C. MILLERET, D. BENNEQUIN, J. TOUBOUL.*Pinwheel-Dipole configuration in cat visual cortex*, in "NeuroImage", 2016, vol. 128, p. 63-73, https://hal.archives-ouvertes.fr/hal-01412346.

[26] P. ROBERT, J. TOUBOUL.*On the dynamics of random neuronal networks*, in "Journal of Statistical Physics", September 2016, vol. 165, n^o 3, p. 545-584 [*DOI :* 10.1007/S10955-016-1622-9], https://hal.inria.fr/hal-01075242.

[27] A. ROMAGNONI, J. RIBOT, D. BENNEQUIN, J. TOUBOUL.*Parsimony, exhaustivity and balanced detection in neocortex*, in "PLoS Computational Biology", 2016, vol. 11, n^o 11, e1004623 [*DOI :* 10.1371/JOURNAL.PCBI], https://hal.archives-ouvertes.fr/hal-01412347.

Scientific Books (or Scientific Book chapters)

[28] F. CLÉMENT, A. VIDAL.*Modeling the Dynamics of Gonadotropin-Releasing Hormone (GnRH) Secretion in the Course of an Ovarian Cycle* , in "Computational Neuroendocrinology", D. J. MACGREGOR, G. LENG (editors), Wiley-INF neuroendocrinology series, John Wiley & Sons, April 2016, https://hal.inria.fr/hal-01294646.

Other Publications

[29] F. CLÉMENT, F. COQUEL, M. POSTEL, K. L. TRAN.*Dimensional reduction of a multiscale model based on long time asymptotics*, June 2016, working paper or preprint, http://hal.upmc.fr/hal-01325275.

[30] J. RUBIN, J. SIGNERSKA-RYNKOWSKA, J. TOUBOUL, A. VIDAL.*Wild oscillations in a nonlinear neuron model with resets: (I) Bursting, spike adding and chaos*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01416002.

[31] J. SIGNERSKA-RYNKOWSKA, J. TOUBOUL, A. VIDAL.*A geometric mechanism for mixed-mode bursting oscillations in a hybrid neuron model*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01256368.

References in notes

[32] P. BRESSLOFF.*Spatiotemporal dynamics of continuum neural fields*, in "J. Phys. A: Math. Theor.", 2012, vol. 45, p. 033001–.

[33] F. CLÉMENT, A. VIDAL.*Foliation-based parameter tuning in a model of the GnRH pulse and surge generator*, in "SIAM J. Appl. Dyn. Syst.", 2009, vol. 8, n^o 4, p. 1591–1631.

[34] F. CLÉMENT, J.-P. FRANÇOISE.*Mathematical modeling of the GnRH-pulse and surge generator*, in "SIAM J. Appl. Dyn. Syst.", 2007, vol. 6, p. 441-456.

[35] N. ECHENIM, D. MONNIAUX, M. SORINE, F. CLÉMENT.*Multi-scale modeling of the follicle selection process in the ovary*, in "Math. Biosci.", 2005, vol. 198, p. 57-79.

[36] J. RIBOT, A. ROMAGNONI, C. MILLERET, D. BENNEQUIN, J. TOUBOUL.*Pinwheel-dipole configuration in cat early visual cortex*, 2014, p. 63–73.

[37] T. VO, R. BERTRAM, J. TABAK, M. WECHSELBERGER.*Mixed mode oscillations as a mechanism for pseudo-plateau bursting*, in "J. Comput. Neurosci.", 2010, vol. 28, n^{o} 3, p. 443–458.

[38] H. WILSON, J. COWAN.*Excitatory and inhibitory interactions in localized populations of model neurons*, in "Biophys. J.", 1972, vol. 12, p. 1–24.

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:

1.1.1. - Multicore
1.1.3. - Memory models
2.1.1. - Semantics of programming languages
2.1.3. - Functional programming
2.1.6. - Concurrent programming
2.1.8. - Synchronous languages
2.2.2. - Memory models
2.2.3. - Run-time systems
2.2.4. - Parallel architectures
2.2.5. - GPGPU, FPGA, etc.
2.2.6. - Adaptive compilation
2.3. - Embedded and cyber-physical systems
2.3.1. - Embedded systems
2.3.2. - Cyber-physical systems
2.3.3. - Real-time systems
2.4.3. - Proofs
3.1.3. - Distributed data
3.1.8. - Big data (production, storage, transfer)
6.2.1. - Numerical analysis of PDE and ODE
6.2.7. - High performance computing

Other Research Topics and Application Domains:

5.2.1. - Road vehicles
5.2.2. - Railway
5.2.3. - Aviation
6.4. - Internet of things
6.6. - Embedded systems
9.2.1. - Music, sound
9.4.1. - Computer science
9.4.2. - Mathematics

1. Members

Research Scientists

Timothy Bourke [Inria, Starting Research position]
Albert Cohen [Inria, Senior Researcher, HDR]
Francesco Zappa Nardelli [Inria, Senior Researcher, HDR]

Faculty Member

Marc Pouzet [Team leader, Univ. Paris VI, Professor]

Technical Staff

Jean-Baptiste Brejon [Inria, until Sep 2016]
Michael Kruse [Inria]
Mircea Namolaru [Inria, until Sep 2016, granted by FP7 EuroLab-4-HPC project]
Laurent Morin [Inria, External collaborator from Rennes]
Zhen Zhang [Inria, Postdoctoral Fellow, until Jan 2016]
Oleksandr Zinenko [Inria, Postdoctoral Fellow]

PhD Students

Guillaume Baudart [ENS Paris]
Julien Proy [ENS Paris and INVIA, CIFRE Fellowship]
Adilla Susungi [Mines Paristech, PSL Fellowship]
Ulysse Beaugnon [ENS Paris]
Lelio Brun [ENS Paris, from Apr 2016]
Prasanth Chatarasi [Rice University and Inria, from May 2016 until Aug 2016]
Nhat Minh Le [ENS Paris and Inria, until Nov 2016, granted by Min. du Redressement Productif]
Robin Morisset [ENS Paris and Inria, granted by Google Inc]
Chandan Reddy Gopal [ENS Paris and Inria]
Jie Zhao [Chinese scholarship from NDSC]

Post-Doctoral Fellows

Adrien Guatto [ENS Paris and Inria, until Jan 2016]
Guillaume Iooss [ENS Paris]

Administrative Assistant

Anna Bednarik [Inria]

Others

Tommaso Borghesi [Inria, Master Student, until Jan 2016]
Keyur Joshi [IIT Hyderabad and Inria, Master Student, May 2016 until Jul 2016]
Lilia Otmane Cherif [UPMC and Inria, Master Student, from Apr 2016 until Aug 2016]
Sven Verdoolaege [Independent contractor, Polly Labs grant of ARM]

2. Overall Objectives

2.1. Overall Objectives

The 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), 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 mutually beneficial and 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.

[0] http://www.esterel-technologies.com/products/scade-suite
[0] http://www.mathworks.com/products/simulink
[0] https://www.modelica.org

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.
- 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. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Awards

Marc Pouzet won the Inria/French Académie des Sciences/Dassault Systèmes Innovation award.

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: Pankaj Pawan, Francesco Zappa Nardelli, Robin Morisset, Anirudh Kumar, Pankaj Prateek Kewalramani and Pankaj More
- 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

- Participants: Albert Cohen, Tobias Grosser, Feng Li, Riyadh Baghdadi and Nhat Minh Le
- Contact: Albert Cohen
- URL: http://gcc.gnu.org/

5.3. Heptagon

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.

- Participants: Adrien Guatto, Marc Pouzet, Cédric Pasteur, Léonard Gerard, Brice Gelineau, Gwenael Delaval and Eric Rutten
- Contact: Marc Pouzet

5.4. 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: Scott Owens, Peter Sewell and Francesco Zappa Nardelli
- Contact: Francesco Zappa Nardelli
- URL: http://www.cl.cam.ac.uk/~pes20/lem/

5.5. 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.

- Contact: Marc Pouzet
- URL: http://www.di.ens.fr/~pouzet/lucid-synchrone/

5.6. 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: Albert Cohen, Adrien Guatto, Marc Pouzet and Louis Mandel
- Contact: Albert Cohen
- URL: https://www.lri.fr/~mandel/lucy-n/

5.7. 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/

5.8. 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: Sven Verdoolaege, Tobias Grosser, Riyadh Baghdadi and Albert Cohen
- Contact: Sven Verdoolaege
- URL: http://freshmeat.net/projects/ppcg

5.9. 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: Guillaume Baudart, Louis Mandel and Cédric Pasteur
- Contact: Guillaume Baudart
- URL: http://rml.lri.fr

5.10. SundialsML

Sundials/ML
KEYWORDS: Simulation - Mathematics - Numerical simulations
SCIENTIFIC DESCRIPTION

Sundials/ML is an OCaml interface to the Sundials suite of numerical solvers (CVODE, CVODES, IDA, IDAS, KINSOL, ARKODE). It supports all features except for the Hypre and PETSC nvectors (which require additional libraries). 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 an OCaml interface to the Sundials suite of numerical solvers (CVODE, CVODES, IDA, IDAS, KINSOL, ARKODE).
NEW PROGRESS

This year we updated our interface to work with versions 2.6.0 and 2.7.0 of the Sundials library. This included significant work to support the new ARKODE solver, sparse matrices and the KLU and SuperLU/MT linear solvers, OpenMP and Pthreads nvectors, and various new functions and linear solvers in existing solvers. The source files were completely reorganized. The OCaml types for nvectors were adapted to support multiple nvectors. Memory leaks were eliminated and the performance problems investigated. This work was presented at the ACM Workshop on ML [28].

- Participants: Marc Pouzet and Timothy Bourke
- Partner: UPMC, AIST (Jun Inoue)
- Contact: Timothy Bourke
- URL: http://inria-parkas.github.io/sundialsml/

5.11. Zélus

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

5.12. 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: Sven Verdoolaege, Tobias Grosser and Albert Cohen
- Contact: Sven Verdoolaege
- URL: http://freshmeat.net/projects/isl

6. New Results

6.1. Verified compilation of Lustre

Participants: Timothy Bourke, Lélio Brun, Marc Pouzet.

Synchronous dataflow languages and their compilers are increasingly used to develop safety-critical applications, like fly-by-wire controllers in aircraft and monitoring software for power plants. A striking example is the SCADE Suite tool of ANSYS/Esterel Technologies which is DO-178B/C qualified for the aerospace and defense industries. This tool allows engineers to develop and validate systems at the level of abstract block diagrams that are automatically compiled into executable code.

Formal modelling and verification in an interactive theorem prover can potentially complement the industrial certification of such tools to give very precise definitions of language features and increased confidence in their correct compilation; ideally, right down to the binary code that actually executes.

This year we integrated elements of the CompCert verified C compiler into our Lustre compiler. In particular, we modularized the syntax and semantics of our source Lustre language and intermediate Obc language to be independent of the underlying types and operators of the host language. All previous proofs are independent of the choice of host language. We integrated CompCert by instantiating the types and operators with those of the Clight language and by adding a function that compiles an Obc program into Clight. The key challenge in this compilation pass is to move from a model where program variables are stored in a tree structure where distinctness is manifest to a model where variables are stored in nested structures in a single memory block with concomitant problems of aliasing, alignment, and memory size. We addressed this challenge by extending a CompCert library for expressing separation assertions and applying it to express our recursive predicates.

A similar approach was taken to address the encoding of multiple return values (permitted in Obc but not in Clight). We made various practical improvements to our compiler and proofs including the addition of a verified parser, the addition of an elaboration pass with type and clock checking, and pretty-printers for intermediate languages. It is now possible to compile scheduled and normalized Lustre programs to assembly code with a proof correction that relates the generated transition function to the dataflow semantics of the source program.

The initial part of this work, reported last year, has been published [20].

In collaboration with Pierre-Évariste Dagand (CNRS), Lionel Reig (Collège de France), and Xavier Leroy (Inria, GALLIUM team).

6.2. Fence Optimisations for Multicore Architectures

Participants: Robin Morisset, Francesco Zappa Nardelli.

We have pursued our investigation of sound optimisations for modern multicore architectures. Last year we focused on optimisations that can be expressed inside the semantics of the C11/C++11 programming language; we thus moved to optimisations that can be expressed only at the harware level. In particular we have shown how partial redundancy elimination (PRE) can be instantiated to perform *provably correct* fence elimination for multi-threaded programs running on top of the x86, ARM and IBM Power relaxed memory models. We have implemented our algorithm in the x86, ARM and Power backends of the `LLVM` compiler infrastructure. The optimisation does not induce an observable overhead at compile-time and can result in up-to 10% speedup on some benchmarks.

This work has been published in CC 2017 [18]. The implementation of the optimisations will be submitted for inclusion in the LLVM compiler suite.

6.3. Compiling synchronous languages for multi-processor implementations

Participants: Timothy Bourke, Albert Cohen, Guillaume Iooss, Marc Pouzet.

Working together with industrial partners in the context of the ASSUME project, we have been working to treat a large-scale and complete case study of an industrial application. This has involved studying the original sources and adapting the Heptagon Lustre compiler. Three main extensions have been developed this year: a mechanism to calculate and exploit module interdependencies; an extension to the type system to allow operator overloading via ad hoc polymorphism; and modifications to the parser to accept the provided source code. We have also worked on a means to generate dependency graphs from the provided nonfunctional specifications.

Our current work centers on understanding how to formalize the peculiarities of this class of application and the target architecture in our framework, and on generating Lustre code from the non-functional specifications. The ultimate aim is to generate correct multi-processor task-parallel real-time code for an embedded target and to integrate with both the Heptagon and Vélus compilers.

In collaboration (this year) with Dumitru Potop-Butucaru (Inria, AOSTE team), Keryan Didier (Inria, AOSTE team), Jean Souyris (Airbus), and Adrien Gauffriau (Airbus).

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

Technology Transfer Project, partly funded by the TETRACOM grant and by Kalray.

7.2. Bilateral Grants with Industry

Polly Labs initiative. Funded by ARM for 4 years with complementary support from Xilinx, in cooperation with ETH Zürich and Qualcomm.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

ANR WMC project (program “jeunes chercheuses, jeunes chercheurs”), 2012–2016, 200 Keuros. F. Zappa Nardelli is the main investigator.

ANR Boole project (program “action blanche”), 2009-2014.

ANR CAFEIN, 2013-2015. Marc Pouzet.

8.1.2. Investissements d’avenir

Sys2Soft contract (Briques Génériques du Logiciel Embarqué). Partenaire principal: Dassault-Systèmes, etc. Inria contacts are Benoit Caillaud (HYCOMES, Rennes) and Marc Pouzet (PARKAS, Paris).

ManycoreLabs contract (Briques Génériques du Logiciel Embarqué). Partenaire principal: Kalray. Inria contacts are Albert Cohen (PARKAS, Paris), Alain Darte (COMPSYS, Lyon), Fabrice Rastello (CORSE, Grenoble).

8.1.3. Others

Marc Pouzet is scientific advisor for the Esterel-Technologies/ANSYS company.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. Eurolab-4-HPC

Title: EuroLab-4-HPC: Foundations of a European Research Center of Excellence in High Performance Computing Systems

Programm: H2020

Duration: September 2015 - September 2017

Coordinator: CHALMERS TEKNISKA HOEGSKOLA AB

Partners:

- Barcelona Supercomputing Center - Centro Nacional de Supercomputacion (Spain)
- Chalmers Tekniska Hoegskola (Sweden)
- Ecole Polytechnique Federale de Lausanne (Switzerland)
- Eidgenoessische Technische Hochschule Zuerich (Switzerland)
- Foundation for Research and Technology Hellas (Greece)
- Universitaet Stuttgart (Germany)
- Rheinisch-Westfaelische Technische Hochschule Aachen (Germany)
- Technion - Israel Institute of Technology (Israel)
- Universitaet Augsburg (Germany)
- The University of Edinburgh (United Kingdom)
- Universiteit Gent (Belgium)
- The University of Manchester (United Kingdom)

Inria contact: Albert Cohen

Europe has built momentum in becoming a leader in large parts of the HPC ecosystem. It has brought together technical and business stakeholders from application developers via system software to exascale systems. Despite such gains, excellence in high performance computing systems is often fragmented and opportunities for synergy missed. To compete internationally, Europe must bring together the best research groups to tackle the longterm challenges for HPC. These typically cut across layers, e.g., performance, energy efficiency and dependability, so excellence in research must target all the layers in the system stack. The EuroLab-4-HPC project's bold overall goal is to build connected and sustainable leadership in high-performance computing systems by bringing together the different and leading performance orientated communities in Europe, working across all layers of the system stack and, at the same time, fuelling new industries in HPC.

8.2.1.2. TETRACOM

Title: Technology Transfer in Computing Systems

Programm: FP7

Duration: September 2013 - August 2016

Coordinator: RHEINISCH-WESTFAELISCHE TECHNISCHE HOCHSCHULE AACHEN

Partners:

Imperial College of Science, Technology and Medicine (United Kingdom)

Rheinisch-Westfaelische Technische Hochschule Aachen (Germany)

Technische Universiteit Delft (Netherlands)

Tty-Saatio (Finland)

Universita di Pisa (Italy)

Inria contact: Albert Cohen

The mission of the TETRACOM Coordination Action is to boost European academia-to-industry technology transfer (TT) in all domains of Computing Systems. While many other European and national initiatives focus on training of entrepreneurs and support for start-up companies, the key differentiator of TETRACOM is a novel instrument called Technology Transfer Project (TTP). TTPs help to lower the barrier for researchers to make the first steps towards commercialisation of their research results. TTPs are designed to provide incentives for TT at small to medium scale via partial funding of dedicated, well-defined, and short term academia-industry collaborations that bring concrete R&D results into industrial use. This will be implemented via competitive Expressions-of-Interest (EoI) calls for TTPs, whose coordination, prioritization, evaluation, and management are the major actions of TETRACOM. It is expected to fund up to 50 TTPs. The TTP activities will be complemented by Technology Transfer Infrastructures (TTIs) that provide training, service, and dissemination actions. These are designed to encourage a larger fraction of the R&D community to engage in TTPs, possibly even for the first time. Altogether, TETRACOM is conceived as the major pilot project of its kind in the area of Computing Systems, acting as a TT catalyst for the mutual benefit of academia and industry. The projects primary success metrics are the number and value of coordinated TTPs as well as the amount of newly introduced European TT actors. It is expected to acquire around more than 20 new contractors over the project duration. TETRACOM complements and actually precedes the use of existing financial instruments such as venture capital or business angels based funding.

8.2.2. Collaborations in European Programs, Except FP7 & H2020

Program: ITEA 3

Project acronym: ASSUME

Project title: Affordable Safe & Secure Mobility Evolution

Duration: Sep 2015–Aug 2018

Coordinator: Udo Gleich

Other partners: AbsInt Angewandte Informatik GmbH, Airbus, Arcelik, Articus Systems AB, BTC Embedded Systems AG, Berner & Mattner Systemtechnik GmbH, Daimler AG, Eindhoven University of Technology, Ericsson, ANSYS, FindOut Technologies AB,

Ford Otosan, Forschungszentrum Informatik (FZI), Havelsan, KTH (Royal Institute of Technology), Kalray SA, Karlsruhe Institute of Technology (KIT), Kiel University, Koc University, KoçSistem, Model Engineering Solutions GmbH, Mälardalen University, NXP Semiconductors, OFFIS, Recore Systems BV, Robert Bosch GmbH, Safran Aircraft Engines SAS, Safran Electronics & Defense, Scania, TNO, Thales, UNIT Information Technologies R&D Ltd., University Pierre et Marie Curie, University of Technology in Munich, University of Twente, VDL Bus & Coach bv, Verum Software Tools BV, École normale supérieure.

Abstract: Future mobility solutions will increasingly rely on smart components that continuously monitor the environment and assume more and more responsibility for a convenient, safe and reliable operation. Currently the single most important roadblock for this market is the ability to come up with an affordable, safe multi-core development methodology that allows industry to deliver trustworthy new functions at competitive prices. ASSUME will provide a seamless engineering methodology, which addresses this roadblock on the constructive and analytic side.

8.2.3. Collaborations with Major European Organizations

Albert Cohen is an external member of the ARTEMIS-IA Working Group. Collaborating on the writing of the association's Strategic Research Agenda (SRA), and the ECSEL JU Multi-Annual Research and Innovation Agenda (MASRIA).

https://artemis-ia.eu

8.3. International Initiatives

8.3.1. POLYFLOW

Title: Polyhedral Compilation for Data-Flow Programming Languages

International Partner (Institution - Laboratory - Researcher):

IISc Bangalore (India) - Department of Computer Science and Automation (CSA) - Uday Kumar Reddy Bondhugula

Start year: 2016

See also: http://polyflow.gforge.inria.fr

The objective of the associate team is to foster collaborations on fundamental and applied research. It also supports training sessions, exchange of undergraduate and master students, and highlighting opportunities in the partners' research, education and economic environments.

Polyhedral techniques for program transformation are now used in several proprietary and open source compilers. However, most of the research on polyhedral compilation has focused on imperative languages, where computation is specified in terms of computational statements within nested loops and control structures. Graphical data-flow languages, where there is no notion of statements or a schedule specifying their relative execution order, have so far not been studied using a powerful transformation or optimization approach. These languages are extremely popular in the system analysis, modeling and design of embedded reactive control applications. They also underline the construction of domain-specific languages and compiler intermediate representations. The execution semantics of data-flow languages impose a different set of challenges for compilation and optimization. We are studying techniques enabling the extraction of a polyhedral representation from data-flow programs, to transform them with the goal of generating memory-efficient and high-performance code for modern architectures.

The research conducted in PolyFlow covers both fundamental and applied aspects. The partners also emphasize the development of solid research tools. The associate team will facilitate their dissemination as free software and their exploitation through industrial collaborations.

8.3.2. Inria International Partners

8.3.2.1. Informal International Partners

Pr. Peter Sewell, Computer Laboratory, University of Cambridge, UK. Regular visits and scientific collaboration.

Pr. Jan Vitek, College of Computer & Information Science Northeastern University, USA. Regular visits and scientific collaboration.

Prof. Uday Bondhugula, CSA department, Indian Institute of Science, India. See POLYFLOW associate team for details.

Prof. Ramakrishna Updadrasta, IIT Hyderabad, India, collaboration visits including internships.

Prof. P. Sadayappan, CS department, Ohio State University, USA. Joint publications, frequent visits, occasionally for several weeks.

Prof. M. Sheeran, Computer Science and Engineering Department, Chalmers University of Technology, Sweden. Regular visits. Continuing exchanges on languages and compilation for synchronous and hybrid systems.

Prof. C. Tinelli, CS department, University of IOWA, USA. Regular visits. Continuing exchanges on the verification of synchronous languages and programs.

Prof. R. von Hanxleden, Director at the Department of Computer Science, Head of the Real-Time and Embedded Systems Group, Kiel University, Germany. Regular visits and scientific collaboration.

Prof. M. Mendler, Head of the Informatics Theory Group, Bamberg University, Germany. Regular visits and scientific collaboration.

Dr. Sven Verdoolaege, CS department, K. U. Leuven, Belgium. Joint steering of the Polly Labs initiative and contractual cooperation in this context.

Dr. Tobias Grosser in the group of Prof. Torsten Hoeffler, ETH Zürich. Joint steering of the Polly Labs initiative. See Polly Labs for details.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

8.4.1.1. Internships

Prasanth Chatarasi, PhD student from Rice University.

Keyur Joshi, undergraduate student from IIT Hyderabad.

8.4.2. Visits to International Teams

8.4.2.1. Research Stays Abroad

Guillaume Baudart spent three months working at the IBM Thomas J. Watson Research Centre.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- Albert Cohen is the General Chair of PLDI 2017

9.1.1.2. Member of the Conference Program Committees

- Timothy Bourke was a member of the PC of EMSOFT 2016.
- Francesco Zappa Nardelli was a member of the PC of POPL 2017.
- Francesco Zappa Nardelli will be a member of the PC of ECOOP 2017.
- Albert Cohen was a PC member of ASPLOS, PACT, PPoPP, CGO, PLDI.

9.1.1.3. Reviewer

- Timothy Bourke was a reviewer for FM 2016 (Int. Symposium on Formal Methods).

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- Albert Cohen is an Associate Editor of ACM TACO.

9.1.2.2. Reviewer - Reviewing Activities

- Timothy Bourke was a reviewer for IEEE Embedded Systems Letters, ACM Transactions on Embedded Computing Systems, and IEEE Transactions on Software Engineering.

9.1.3. Invited Talks

- April, T. Bourke presented "Towards the verified compilation of Lustre" in the Gallium seminar series in Paris, France.
- December, T. Bourke presented "Verifying a Lustre Compiler (Part 1)" at the SYNCHRON workshop in Bamberg, Germany.
- November, F. Zappa Nardelli presented "Shared Memory Concurrency and Compiler Optimisations" at IMDEA, Madrid, Spain.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: F. Zappa Nardelli: "A Programmer's introduction to Computer Architectures and Operating Systems" (M1), 45h, École Polytechnique, France

Master: A. Cohen & F. Zappa Nardelli, "Semantics, languages and algorithms for multicore programming", Lecture, 9h+12h, M2, MPRI: Ecole normale supeérieure and Université Paris Diderot, France

Licence: F. Zappa Nardelli: "Conception et analyse d'algorithmes" (L3), PCs, 32h, E´cole Polytechnique, France

Master : M. Pouzet & T. Bourke: "Synchronous Systems" (M2), Lectures and TDs, MPRI, France

Master: T. Bourke participated in reviewing the M1 internships of students at the ENS, France.

Licence : M. Pouzet & T. Bourke: "Operating Systems" (L3), Lectures and TDs, ENS, France.

Licence : T. Bourke, "Digital Systems" (L3), Lectures and TDs, ENS, France

Marc Pouzet is Director of Studies for the CS department, at ENS.

9.2.2. Supervision

PhD in progress : Ulyssse Beaugnon, 2nd year, supervised by A. Cohen and M. Pouzet.

PhD in progress : Chandan Reddy, 2nd year, supervised by A. Cohen.

PhD in progress : Jie Zhao, 2nd year, supervised by A. Cohen.

PhD in progress : Guillaume Baudart, 3rd year, supervised by T. Bourke and M. Pouzet. This thesis will be defended in March.

PhD in progress : Lélio Brun, 1st year, supervised by T. Bourke and M. Pouzet.

PhD in progress : Robin Morisset, 3rd year, supervised by F. Zappa Nardelli. This thesis will be defended in April 2017.

9.2.3. Juries

Francesco Zappa Nardelli was an external reviewer of the PhD thesis of Carl Leonardsson, Uppsala University, Sweden.

Francesco Zappa Nardelli was a jury member of the PhD thesis of Nhat Minh Lê, ENS, Paris, France.

Timothy Bourke was an external reviewer of the masters thesis of Shruti Saini, The University of the South Pacific.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] A. ACHARYA, U. BONDHUGULA, A. COHEN.*Polyhedral Auto-Transformation almost without ILP*, in "ACM Transactions on Programming Languages and Systems (TOPLAS)", May 2016, https://hal.inria.fr/hal-01425546.

[2] G. BAUDART, A. BENVENISTE, T. BOURKE.*Loosely Time-Triggered Architectures*, in "ACM Transactions on Embedded Computing Systems (TECS)", August 2016, vol. 15, Article 71 [*DOI :* 10.1145/2932189], https://hal.inria.fr/hal-01408224.

[3] S. G. BHASKARACHARYA, U. BONDHUGULA, A. COHEN.*Automatic Storage Optimization for Arrays*, in "ACM Transactions on Programming Languages and Systems (TOPLAS)", 2016, Original submission, candidate for presentation at PLDI 2016, https://hal.archives-ouvertes.fr/hal-01257223.

[4] S. G. BHASKARACHARYA, U. BONDHUGULA, A. COHEN.*Automatic Storage Optimization for Arrays*, in "ACM Transactions on Programming Languages and Systems (TOPLAS)", 2016, vol. 38, p. 1 - 23 [*DOI :* 10.1145/2845078], https://hal.inria.fr/hal-01425564.

[5] U. BONDHUGULA, A. ACHARYA, A. COHEN.*The Pluto+ Algorithm: A Practical Approach for Parallelization and Locality Optimization of Affine Loop Nests*, in "ACM Transactions on Programming Languages and Systems (TOPLAS)", 2016, https://hal.archives-ouvertes.fr/hal-01257226.

[6] T. BOURKE, R. J. VAN GLABBEEK, P. HÖFNER.*Mechanizing a Process Algebra for Network Protocols*, in "Journal of Automated Reasoning", March 2016, vol. 56, p. 309-341 [*DOI :* 10.1007/s10817-015-9358-9], https://hal.inria.fr/hal-01408217.

[7] I. LLOPARD, C. FABRE, A. COHEN.*A From a Formalized Parallel Action Language to its Efficient Code Generation*, in "ACM Transactions on Embedded Computing Systems (TECS)", January 2017 [*DOI :* 10.1145/0000000.0000000], https://hal.inria.fr/hal-01425140.

International Conferences with Proceedings

[8] G. BAUDART, T. BOURKE, M. POUZET.*Soundness of the Quasi-Synchronous Abstraction*, in "Formal Methods in Computer-Aided Design (FMCAD)", Mountain View, CA, United States, Proceedings of the 16th International Conference on Formal Methods in Computer-Aided Design, October 2016, p. 9-16, https://hal.inria.fr/hal-01408208.

[9] S. G. BHASKARACHARYA, U. BONDHUGULA, A. COHEN.*SMO: An Integrated Approach To Intra-Array and Inter-Array Storage Optimization*, in "Symp. on Principles of Programming Languages (POPL)", St Petersburg, FL, United States, 2016, https://hal.archives-ouvertes.fr/hal-01257228.

[10] S. G. BHASKARACHARYA, U. BONDHUGULA, A. COHEN.*SMO: An Integrated Approach to Intra-array and Inter-array Storage Optimization*, in "POPL 2016 - ACM Symposium on Principles of Programming Languages", Saint Petersburg, United States, January 2016, p. 526-538 [*DOI :* 10.1145/2837614.2837636], https://hal.inria.fr/hal-01425888.

[11] A. COHEN, A. DARTE, P. FEAUTRIER.*Static Analysis of OpenStream Programs*, in "6th International Workshop on Polyhedral Compilation Techniques (IMPACT'16), held with HIPEAC'16", Prague, Czech Republic, Proceedings of the IMPACT series, Michelle Strout and Tomofumi Yuki, January 2016, https://hal.inria.fr/hal-01251845.

[12] A. COHEN, V. PERRELLE, D. POTOP-BUTUCARU, M. POUZET, E. SOUBIRAN, Z. ZHANG.*Hard Real Time and Mixed Time Criticality on Off-The-Shelf Embedded Multi-Cores*, in "International Conference on Embedded and Real-Time Software and Systems (ERTS2)", Toulouse, France, January 2016, https://hal.inria.fr/hal-01425887.

[13] X. K. DO, S. LOUISE, A. COHEN.*Transaction Parameterized Dataflow: A Model for Context-Dependent Streaming Applications*, in "Design, Automation & Test in Europe Conference & Exhibition (DATE)", Dresden, Germany, March 2016, https://hal.inria.fr/hal-01425902.

[14] A. DREBES, J.-B. BRÉJON, A. POP, K. HEYDEMANN, A. COHEN.*Language-Centric Performance Analysis of OpenMP Programs with Aftermath*, in "International Workshop on OpenMP (IWOMP)", Nara, Japan, October 2016, p. 237 - 250 [*DOI :* 10.1007/978-3-319-45550-1_17], https://hal.inria.fr/hal-01425903.

[15] A. DREBES, A. POP, K. HEYDEMANN, A. COHEN.*Interactive visualization of cross-layer performance anomalies in dynamic task-parallel applications and systems*, in "IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)", Uppsala, Sweden, April 2016, p. 274 - 283 [*DOI :* 10.1109/ISPASS.2016.7482102], https://hal.inria.fr/hal-01425892.

[16] A. DREBES, A. POP, K. HEYDEMANN, A. COHEN, N. DRACH.*Scalable Task Parallelism for NUMA: A Uniform Abstractionfor Coordinated Scheduling and Memory Management*, in "PACT : International Conference on Parallel Architectures and Compilation", Haifa, Israel, ACM New York, NY, USA, September 2016, p. 125-137 [*DOI :* 10.1145/2967938.2967946], http://hal.upmc.fr/hal-01365718.

[17] F. GINDRAUD, F. RASTELLO, A. COHEN, F. BROQUEDIS.*A bounded memory allocator for software-defined global address spaces*, in "ISMM 2016 - 2016 ACM SIGPLAN International Symposium on Memory Management", Santa Barbara, United States, June 2016, https://hal.inria.fr/hal-01412919.

[18] R. MORISSET, F. ZAPPA NARDELLI.*Partially Redundant Fence Elimination for x86, ARM and Power processors*, in "International Conference on Compiler Construction (CC)", Austin, United States, February 2017, https://hal.inria.fr/hal-01423612.

[19] S. VERDOOLAEGE, A. COHEN.*Live Range Reordering*, in "6þ Workshop on Polyhedral Compilation Techniques (IMPACT, associated with HiPEAC)", Prag, Czech Republic, 2016, https://hal.archives-ouvertes.fr/hal-01257224.

National Conferences with Proceeding

[20] T. BOURKE, P.-E. DAGAND, M. POUZET, L. RIEG.*Vérification de la génération modulaire du code impératif pour Lustre*, in "JFLA 2017 - Vingt-huitième Journées Francophones des Langages Applicatifs", Gourettes, France, January 2017, https://hal.inria.fr/hal-01403830.

Conferences without Proceedings

[21] A. COHEN, V. PERRELLE, D. POTOP-BUTUCARU, M. POUZET, E. SOUBIRAN, Z. ZHANG.*Hard Real Time and Mixed Time Criticality on Off-The-Shelf Embedded Multi-Cores*, in "8th European Congress on Embedded Real Time Software and Systems (ERTS 2016)", Toulouse, France, January 2016, https://hal.archives-ouvertes.fr/hal-01259157.

[22] A. DREBES, J.-B. BRÉJON, A. POP, K. HEYDEMANN, A. COHEN.*Language-Centric Performance Analysis of OpenMP Programs with Aftermath*, in "International Workshop on OpenMP", Nara, Japan, October 2016, http://hal.upmc.fr/hal-01343686.

[23] C. HONG, W. BAO, A. COHEN, S. KRISHNAMOORTHY, L.-N. POUCHET, F. RASTELLO, J. RAMANUJAM, S. PONNUSWANY.*Effective padding of multidimensional arrays to avoid cache conflict misses*, in "PLDI 2016: Proceedings of the 37th ACM SIGPLAN Conference on Programming Language Design and Implementation", Santa Barbara, United States, June 2016, https://hal.inria.fr/hal-01335346.

Scientific Books (or Scientific Book chapters)

[24] S. POP, A. COHEN.*SSA-based Compiler Design*, in "SSA-based Compiler Design", F. RASTELLO (editor), springer, August 2016, vol. Loop tree and induction variables, ISBN 978-1-4419-6201-0, https://hal.archives-ouvertes.fr/hal-01257229.

Research Reports

[25] D. BARTHOU, G. GROSDIDIER, K. PETROV, M. KRUSE, C. EISENBEIS, O. PÈNE, O. BRAND-FOISSAC, C. TADONKI, R. DOLBEAU.*Automated Code Generation for Lattice QCD Simulation*, University of Bordeaux, University of Paris Sud, Inria, University of Paris Sud, Mines ParisTech, CAPS Entreprise, June 2016, https://hal-mines-paristech.archives-ouvertes.fr/hal-01433302.

[26] A. BENVENISTE, B. CAILLAUD, M. POUZET, H. ELMQVIST, M. OTTER.*Structural Analysis of Multi-Mode DAE Systems*, Inria, July 2016, n^{o} RR-8933, 32, https://hal.inria.fr/hal-01343967.

[27] A. COHEN, A. DARTE, P. FEAUTRIER.*Static Analysis of OpenStream Programs*, CNRS ; Inria ; ENS Lyon, January 2016, n^{o} RR-8764, 26, Corresponding publication at IMPACT'16 (http://impact.gforge.inria.fr/impact2016), https://hal.inria.fr/hal-01184408.

Other Publications

[28] T. BOURKE, J. INOUE, M. POUZET.*Sundials/ML: interfacing with numerical solvers*, September 2016, ACM Workshop on ML, https://hal.inria.fr/hal-01408230.

[29] A. DREBES, A. POP, K. HEYDEMANN, N. DRACH, A. COHEN.*NUMA-aware scheduling and memory allocation for data-flow task-parallel applications*, ACM New York, NY, USA, March 2016, p. 44:1-44:2, ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, Poster [*DOI :* 10.1145/2851141.2851193], http://hal.upmc.fr/hal-01365746.

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:

2.1.1. - Semantics of programming languages
2.1.3. - Functional programming
2.1.11. - Proof languages
2.4.3. - Proofs
7.2. - Discrete mathematics, combinatorics
7.4. - Logic in Computer Science
7.6. - Computer Algebra

Other Research Topics and Application Domains:

6.1. - Software industry
6.6. - Embedded systems

1. Members

Research Scientists

Pierre-Louis Curien [Team leader, CNRS, Senior Researcher, HDR]
Yves Guiraud [Inria, Researcher]
Hugo Herbelin [Inria, Senior Researcher, HDR]
Jean-Jacques Lévy [Inria, Emeritus Senior Researcher, HDR]
Alexis Saurin [CNRS, Researcher]
Matthieu Sozeau [Inria, Researcher]

Faculty Members

Thierry Coquand [Inria International Chair, Professor, Göteborg University]
Pierre Letouzey [Univ. Paris 7, Assistant Professor]
Yann Régis-Gianas [Univ. Paris 7, Assistant Professor]

Technical Staff

Daniel de Rauglaudre [Inria]
Matej Kosik [Inria, until September 2016]

PhD Students

Amina Doumane [Univ. Paris 7]
Cyrille Chenavier [Univ. Paris 7until December 2016, ATER 2016-2017]
Maxime Lucas [Univ. Paris 7]
Thibaut Girka [Univ. Paris 7, CIFRE grant (with Mitsubishi Rennes)]
Étienne Miquey [Univ. Paris 7]
Jovana Obradović [Univ. Paris 7]
Cyprien Mangin [Univ. Paris 7]
Guillaume Claret [Univ. Paris 7]
Gabriel Lewertowski [Univ. Paris 7, until September 2016]
Théo Zimmermann [Univ. Paris 7]

Visiting Scientist

Paul Aaron Steckler [MIT & Inria, from Apr 2016 until May 2016]

Administrative Assistant

Lindsay Polienor [Inria]

Others

Bruno Barras [External collaborator, Inria Saclay, Research Scientist]
Philippe Malbos [External collaborator, Univ. Lyon 1, Faculty Member]
Samuel Mimram [External collaborator, Ecole Polytechnique, Faculty Member]
Arnaud Spiwack [External collaborator, Tweag I/O]

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 Gentzen [57] 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 [52], then by Howard and de Bruijn at the end of the 60's [70], [89], 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 [49], [50] stemmed out – a formalism that is both a logic and a programming language and that is at the source of the Coq system [87].

3.1.2. Towards the calculus of constructions

The λ-calculus, defined by Church [48], 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 [44].

To explain the Curry-Howard correspondence, it is important to distinguish between intuitionistic and classical logic: following Brouwer at the beginning of the 20$^{\text{th}}$ 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 [80].

In 1985, Coquand and Huet [49], [50] in the Formel team of Inria-Rocquencourt explored an alternative approach based on Girard-Reynolds' system F [58], [83]. 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 [51] 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

Since 1984, 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 is now. 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 nowadays 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 by employees of Inria, the CNAM and Paris 7.

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 software.

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). 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 [59] 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 [76] and Reynolds [82] and started to be studied in an abstract way in the 80's by Felleisen *et al* [55], leading to Parigot's $\lambda\mu$-calculus [81], 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 [75] and Krivine's abstract machine for call-by-name evaluation [72], [71]. 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 [56]: 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 [43]) 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 [54]).

Since a few years, the categorification process has reached logic, with the introduction of homotopy type theory. After a preliminary result that had identified categorical structures in type theory [69], 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 ressemblance 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 [86], and independently by Burroni [47], 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 [73], [79], [60], [61], and beyond: Petri nets [63] and formal proofs of classical and linear logic have been expressed in this framework [62]. 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 [46].

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 work of Anick [41] and Squier [85], [84] in the 1980s: 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 [9], [10], [66], [67], [68].

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 [9] and algebras [65]. 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, each of which corresponding 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 [74], [42]). 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 [7], and in homological algebra [64], [65].

4. New Software and Platforms

4.1. Coq

KEYWORDS: Proof - Certification - Formalisation
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, ...

- Closest participants: Benjamin Grégoire, Enrico Tassi, Bruno Barras, Yves Bertot, Pierre Courtieu, Maxime Dénès, Hugo Herbelin, Matej Košík, Pierre Letouzey, Assia Mahboubi, Cyprien Mangin, Guillaume Melquiond, Jean-Marc Notin, Pierre-Marie Pédrot, Yann Régis-Gianas, Matthieu Sozeau, Arnaud Spiwack, Théo Zimmermann.
- Partners: CNRS - ENS Lyon - Université Paris-Diderot - Université Paris-Sud
- Contact: Matthieu Sozeau
- URL: http://coq.inria.fr/

4.1.1. *Coq 8.6*

The 8.6 version of Coq was released in December 2016. It initiates a time-based release cycle and concentrates on a smaller set of features than Coq 8.5 for which compatibility and testing were done more intensively. In the πr^2 team, Hugo Herbelin, Cyprien Mangin, Théo Zimmermann and Matthieu Sozeau contributed to the coordination of the development, to the discussion of the roadmap, to the implementation of the features, and to many bugfixes.

Matthieu Sozeau followed up his work on universe polymorphism making the explicit annotation system more accessible and resolving issues in the minimization algorithm used during refinement, resulting in a more predictable system. These improvements were used in the Coq/HoTT library for Homotopy Type Theory, which is described in an upcoming article [26].

Matthieu Sozeau implemented a new variant of the proof-search tactic for typeclasses that is set to replace the existing auto and eauto tactics in the following version. The new variant fully benefits from the features of the underlying proof engine, and allows much more control on proof-search (patterns used consistently, modes for triggering hints, ...). It is at the basis of the work of Théo Zimmermann described below.

4.1.2. The Equations plugin

Cyprien Mangin and Matthieu Sozeau continued work on the Equations plugin, modularizing it so that the use of axioms can be minimized, and making it compatible with developments in Homotopy Type Theory. To achieve this, it has moved to a simplification engine in ML based on telescopes and is able to produce axiom-free proofs of the examples that were previously implicitly using them. This work will be presented at the POPL workshop Type-Theoretic Tools (TTT), next January 2017.

4.1.3. Maintenance

Among other contributions, Hugo Herbelin, Pierre Letouzey, Matej Košík and Matthieu Sozeau worked at the maintenance of the system.

In particular, Pierre Letouzey vastly reworked the build mechanism of Coq, taking advantage of code evolutions driven by Pierre-Marie Pédrot. Pierre Letouzey also administrated (and improved) several machines or systems that are critical for the Coq community (web server, build test server, git repositories ...), in coordination with Inria's SIC support team.

Matej Košík developed a new benchmarking infrastructure based on Jenkins and continuous integration (http://ci.inria.fr). It allows easily testing any developer branch on the benchmark suite prior to integration to the main archive.

4.1.4. Coordination and animation

After 10 years coordinating the Coq development team, Hugo Herbelin handed over the coordination to Matthieu Sozeau.

A Coq working group is organised every two months (5 times a year). Discussions about the development happen, in particular, on `coq-dev@inria.fr`, Coq's GitHub http://github.com/coq and http://coq.inria.fr/bugs. This year, a week-long working group organized in Sophia-Antipolis was devoted to the 8.6 roadmap discussion.

4.1.5. Documentation and stabilization of Coq's programming interface

Matej Košík worked on the programming interfaces of Coq, starting to isolate a subset of key functions to be used by Coq plugin developers.

4.2. Other software developments

In collaboration with François Pottier (Inria Gallium), Yann Régis-Gianas maintained Menhir, an LR parser generator for OCaml. Yann Régis-Gianas develops the "Hacking Dojo", a web platform to automatically grade programming exercises. The platform is now used in several courses of the University Paris Diderot. Yann Régis-Gianas develops a reference implementation of a syntactic analyzer for the POSIX shell programming language. This analyzer is used by the Colis project to analyze the scripts embedded in the packages of the Debian GNU/Linux distribution. In collaboration with Beta Ziliani (LIIS, Cordoba, Argentine), Yann Régis-Gianas, Béatrice Carré and Jacques-Pascal Deplaix develop MetaCoq, an extension of Coq to use Coq as a metalanguage for itself.

Yves Guiraud has updated the Catex tool for Latex, whose purpose is to automatise the production of string diagrams from algebraic expressions http://www.irif.fr/~guiraud/catex/catex.zip. Yves Guiraud collaborates with Samuel Mimram (LIX) to develop the prototype Rewr that implements several algorithms developed in the "Effective higher-dimensional algebra" research direction, including the homotopical completion-reduction procedure of [10]. An online version is available at http://www.lix.polytechnique.fr/Labo/Samuel.Mimram/rewr.

5. New Results

5.1. Effects in proof theory and programming

Participants: Hugo Herbelin, Gabriel Lewertowski, Étienne Miquey, Alexis Saurin, Matthieu Sozeau.

5.1.1. A classical sequent calculus with dependent types

Dependent types are a key feature of type systems, typically used in the context of both richly-typed programming languages and proof assistants. Control operators, which are connected with classical logic along the proof-as-program correspondence, are known to misbehave in the presence of dependent types [11], unless dependencies are restricted to values. As a step in his work to develop a sequent-calculus version of Hugo Herbelin's dPA_ω system [13], Étienne Miquey proposed a sequent calculus with classical logic and dependent types. His calculus—named dL—is an extension of the $\mu\widetilde{\mu}$-calculus with a syntactical restriction of dependent types to the fragment of *negative-elimination free* proofs. The corresponding type system includes a list of explicit dependencies, which maintains type safety. He showed that a continuation-passing style translation can be derived by adding delimited continuations, and how a chain of dependencies can be related to a manipulation of the return type of this continuations. This work has been accepted for publication at ESOP 2017 [39].

5.1.2. Logical foundations of call-by-need evaluation

Alexis Saurin, in collaboration with Pierre-Marie Pédrot, extended their reconstruction of call-by-need based on linear head reduction with control. They showed how linear head reduction could be adapted to the $\lambda\mu$-calculus. This classical linear head reduction lifts the usual properties of the intuitionistic one (with respect to σ-equivalence) to the $\lambda\mu$-calculus (and its σ-equivalence already formulated by Olivier Laurent in his PhD thesis). Moreover, they showed that substitution sequences of the $\lambda\mu$-calculus' linear head reduction are in correspondence with the classical Krivine abstract machine substitution sequences, validating the known fact that the KAM implements linear head reduction. This work has been published at ESOP'16 [29]. They plan to lift to the $\lambda\mu$-calculus their three-step transformation from linear head reduction to call-by-need, and to study the correspondence with Ariola, Herbelin and Saurin's classical call-by-need.

5.1.3. Call-by-name forcing for Dependent Type Theory

Guilhem Jaber, Gabriel Lewertowski, Pierre-Marie Pédrot, Matthieu Sozeau, and Nicolas Tabareau studied a variant of the forcing translation for dependent type theory, moving from the call-by-value variant to a call-by-name version which naturally preserves definitional equalities, avoiding the coherence pitfalls of the former one. This new version was inspired by Pierre-Marie Pédrot's former decomposition of forcing in call-by-push-value. It allows to show various metatheoretical results in a succint fashion, notably for the independence of axioms. Work is ongoing to produce more positive results including abstracting reasoning on step-indexing using this technique. This work was presented at LICS 2016 [28].

5.1.4. Classical realizability and implicative algebras

Étienne Miquey has been working with Alexandre Miquel in Montevideo on the topic of implicative algebras. Implicative algebras are an algebraization of the structure needed to develop a realizability model. In particular, they give rise to the usual ordered combinatory algebras and thus to the triposes used to model classical realizability. An implicative algebra is given by an implicative structure (which consists of a complete semi-lattice with a binary operation $\rightarrow$) together with a separator containing the element interpreted as true in the structure. Étienne Miquey has been working on a formalization of implicative algebras theory in Coq. Following the work of Guillaume Munch-Maccagnoni on focalization and classical realizability, he also worked on alternative presentations within structures based on other connectives, (negation, "par", tensor), rather than $\rightarrow$. Such connectives correspond to the decomposition of the arrow according to the strategy of evaluation (call-by-name/call-by-value). The aim of this work is to obtain a classification of the possible algebraic structures to interpret classical realizability, in order to prove that different strategies of evaluation actually provide us with equivalent models.

5.2. Reasoning and programming with infinite data

Participants: Amina Doumane, Yann Régis-Gianas, Alexis Saurin.

This theme is part of the ANR project Rapido (see the National Initiatives section).

5.2.1. Proof theory of infinitary and circular proofs

In collaboration with David Baelde, Amina Doumane and Alexis Saurin developed further the theory of infinite proofs. In their study of the proof theory of circular and infinitary proofs in $\mu MALL$, they established two fundamental proof-theoretical and computational results, namely cut-elimination and focalisation. This result appeared in CSL 2016 (long version in [33]).

The usual result of focalisation for linear logic can actually be extended to circular proofs, but, contrarily to finitary $\mu MALL$ proofs where fixed-points operators can be given an arbitrary polarity, the least fixed-points must be set to be a positive construction and the greatest fixed-points to be negative, which is consistent with intuition from programming with inductive and co-inductive datatypes. An interesting phenomenon arising with focalisation is that some infinite but regular proofs may not have any regular focused proofs. This is similar to what happens for cut-elimination of regular proofs.

The proof of cut-elimination is quite involved and proceeds in two steps relying on semantic arguments, even though the paper actually proves a cut-elimination result and not only a cut-admissibility result as usual semantic arguments provide. A first part of the proof shows that some cut-reduction strategy is actually productive while a second part of the proof shows that the proof-object produced is actually a correct proof in the sense that it satisfies the validity condition of $\mu MALL$ infinite proofs. Previous cut-elimination results were only known for the restricted additive fragment of linear logic with fixed points, a result due to Santocanale and Fortier.

Baelde, Doumane and Saurin are currently working with Jaber to extend the cut-elimination result to a more expressive validity condition for $\mu MALL$ infinite proofs.

5.2.2. Automata theory meets proof theory: proof certificates for Büchi inclusion

In a joint work with David Baelde and Lucca Hirschi, Amina Doumane and Alexis Saurin carried out a proof-theoretical investigation of the linear-time μ-calculus, proposing well-structured proof systems and showing constructively that they are complete for inclusions of Büchi automata suitably encoded as formulas.

They do so in a way that combines the advantages of two lines of previous work: Kaivola gave a proof of completeness for an axiomatisation that amounts to a finitary proof system, but his proof is non-constructive and yields no reasonable procedure. On the other hand, Dax, Hofmann and Lange recently gave a deductive system that is appropriate for algorithmic proof search, but their proofs require a global validity condition and do not have a well understood proof theory.

They work with well-structured proof systems, effectively constructing proofs in a finitary sequent calculus that enjoys local correctness and cut elimination. This involves an intermediate circular proof system in which one can obtain proofs for all inclusions of parity automata, by adapting Safra's construction. In order to finally obtain finite proofs of Büchi inclusions, a translation result from circular to finite proofs is designed.

These results appeared in LICS 2016 (long version in [37]). Since then, Doumane extended the result and obtained a constructive proof of completeness for the full linear-time μ-calculus.

5.2.3. Co-patterns

In collaboration with Paul Laforgue (Master 1, University Paris Diderot), Yann Régis-Gianas studied the mechanisms of co-patterns introduced by Abel and Pientka from a programming language perspective. More precisely, they defined an untyped version of this calculus as well as an abstract machine to efficiently evaluate cofunctions. In addition, they designed several (type preserving) encodings of co-patterns using generalized algebraic datatypes and purely functional objects. Finally, they started to revisit an optimisation called "stream fusion" in a purely equational way by application of copattern-based program definitions.

5.2.4. Functional reactive programming

In collaboration with Sylvain Ribstein (Master 1, University Paris Diderot), Yann Régis-Gianas defined an OCaml library for differential functional reactive programming (DFRP). This framework extends standard functional reactive programming with the possibility to modify past events and to compute the consequences of this modification in all the events that depend on it. A paper is in preparation.

Saurin and Tasson co-advised in the spring/summer of 2016 the master internship of Rémi Nollet who started his PhD thesis under their supervision in September 2016. The topic of his thesis is the extension of Curry-Howard correspondence between FRP and LTL as recently noticed by Jeffrey and Jeltsch. During his internship, Nollet studied various proof systems for LTL and compared them to type systems for FRP. He notably studied various translations between natural deduction and sequent calculus, which led him to study precisely the role played by structural rules in those translations and preparing the work for future extensions to classical constructive LTL, and to work out the foundations for an extension of Curien-Herbelin's system L, closer to abstract machines, for LTL.

5.3. Effective higher-dimensional algebra

Participants: Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud, Maxime Lucas, Philippe Malbos, Samuel Mimram, Jovana Obradović.

5.3.1. Rewriting and Garside theory

Yves Guiraud has collaborated with Patrick Dehornoy (LNO, Univ. Caen) to develop an axiomatic setting for monoids with a special notion of quadratic normalisation map with good computational properties. This theory generalises the normalisation procedure known for monoids that admit a special family of generators called a Garside family [53] to a much wider class that also includes the plactic monoids. It is proved that good quadratic normalisation maps correspond to quadratic convergent presentations, together with a sufficient condition for this to happen, based on the shape of the normalisation paths on length-three words. This work has been published in the International Journal of Algebra and Computation [21].

Building on this last article, Yves Guiraud currently collaborates with Matthieu Picantin (IRIF, Univ. Paris 7) to generalise the main results of Gaussent, Guiraud and Malbos on coherent presentations of Artin monoids [7], to monoids with a Garside family. This will allow 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 in collaboration with Matthieu Picantin, Yves Guiraud develops an improvement of the classical Knuth-Bendix completion procedure, called the KGB 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 [21]. The KGB completion procedure is partially implemented in the prototype Rewr, developed by Yves Guiraud and Samuel Mimram.

5.3.2. Higher-dimensional linear rewriting

With Eric Hoffbeck (LAGA, Univ. Paris 13), Yves Guiraud and Philippe Malbos have introduced in [65] the setting of linear polygraphs to formalise a theory of linear rewriting, generalising Gröbner bases. They have adapted the method of Guiraud and Malbos [9] to compute polygraphic resolutions of associative algebras, with applications to the decision of the Koszul homological property. They are currently finishing the major overhaul of this work, started in 2015, whose main goal is to ease the adaptation of the results to other algebraic varieties, like commutative algebras or Lie algebras.

Cyrille Chenavier, supervised by Yves Guiraud and Philippe Malbos, explored the use of Berger's theory of reduction operators [45] to improve the theory of Gröbner bases for associative algebras. This work has permitted to unveil two interesting algebraic structures that are hidden in rewriting theory. First, the operations that associate a normal form to an arbitrary word admit a structure of lattice, that gives a new algebraic characterisation of confluence and a new algorithm for completion, based on an iterated use of the meet-operation of the lattice. Second, under mild technical conditions, the different normalisation strategies are related through braid-like relations, as in Artin monoids, that have been used to propose a new method for a particular problem in homological algebra (namely, the construction of a contracting homotopy for the Koszul complex). The second result is published in Algebra and Representation Theory [20], the first one is submitted for publication [35], and both are contained in Cyrille Chenavier's PhD thesis [19].

5.3.3. Rewriting methods for coherence

Yves Guiraud and Philippe Malbos have written a survey on the use of rewriting methods in algebra, centered on a formulation of Squier's homotopical and homological theorems in the modern language of higher-dimensional categories. This article is intended as an introduction to the domain, mainly for graduate students, and will appear in Mathematical Structures in Computer Science [23].

Maxime Lucas, supervised by Yves Guiraud and Pierre-Louis Curien, has applied the rewriting techniques of Guiraud and Malbos [68] to prove coherence theorems for bicategories and pseudofunctors. He obtained a coherence theorem for pseudonatural transformations thanks to a new theoretical result, improving on the former techniques, that relates the properties of rewriting in 1- and 2-categories. This result is published in the Journal of Pure and Applied Algebra [25]. Maxime is currently engaged into a major rework of the results of [9], that will produce improved methods to build Squier's polygraphic resolution from a convergent presentation, based on the use of cubical higher categories instead of globular ones. He has already achieved a first result in this direction [77], and conducted a major foundational work towars the full result [78], which have just been submitted for publication.

Pierre-Louis Curien and Jovana Obradović pursued their work on cyclic operads (started in [36], now acepted in the Journal Applied Categorical Structures). They established the notion of categorified cyclic operad. Categorification involves weakening the axioms of cyclic operads (from equalities to natural isomorphisms) and formulating conditions concerning these isomorphisms which ensure coherence. For entries-only cyclic operads, this coherence is of the same kind as the coherence of symmetric monoidal categories: all diagrams made of associator and commutator isomorphisms are required to commute. However, in the setting of cyclic operads, where the existence of objects and morphisms depends on the shape of a fixed unrooted tree, these arrows do not always exist. In other words, the coherences that Mac Lane established for symmetric monoidal categories do not solve the coherence problem of categorified cyclic operads. They exhibited the appropriate conditions of this setting and proved the coherence theorem, relying on a result of Došen and

Petrić, coming from the coherence of categorified operads. Additionally, by the equivalence between the two possible characterisations of cyclic operads, for cyclic operads introduced as operads with extra structure (that exchanges the output of an operation with one of its inputs), i.e. for exchangeable-output cyclic operads, they examined which of the axioms of the extra structure needs to be weakened (in order to lift that equivalence to weakened structures), and they exhibited the appropriate coherence conditions in this setting as well.

5.4. Incrementality

Participants: Thibaut Girka, Yann Régis-Gianas.

5.4.1. Incrementality in proof languages

In collaboration with Paolo Giarrusso and Yufei Cai (Univ Marburg, Allemagne), Yann Régis-Gianas developed a new method to incrementalise higher-order programs using formal derivatives and static caching. Yann Régis-Gianas has developed a mechanized proof for this transformation. A paper will be submitted to ICFP 2017.

5.4.2. Difference languages

In collaboration with David Mentré (Mitsubishi), Thibaut Girka and Yann Régis-Gianas have developed a theoretical framework to define a notion of differential operational semantics: a general mathematical object to characterise the difference of behavior of two close programs. A paper is under submission. A technical report is available [8].

Thibaut Girka and Yann Régis-Gianas presented this work in several working groups: Gallium (Paris), “Journée annuelle du groupe LTP” of the GDR GPL (Saclay), LIMA (Nantes), IRIF (Paris).

5.5. Metatheory and development of Coq

Participants: Hugo Herbelin, Pierre Letouzey, Yann Régis-Gianas, Matthieu Sozeau.

5.5.1. Dependent pattern-matching

Hugo Herbelin supervised the internship of Meven Bertrand on compiling dependent pattern-matching using a combination of techniques known as small inversion and generalization, as a following of Pierre Boutillier’s PhD.

5.5.2. Transferring theorems along isomorphisms

Théo Zimmermann has developed a tool for transferring theorems along isomorphic structures. The long-term objective is to provide a language of proof methods matching the level of abstraction common in mathematics. Théo Zimmermann is applying his tool to introduce higher "mathematical" levels of abstraction to the basic Coq method for applying theorems. The proof of concept of this idea will be presented at the TTT POPL workshop in January.

5.5.3. Unification

Matthieu Sozeau worked in collaboration with Beta Ziliani (assistant professor at Córdoba, Argentina) on a journal version of the formalisation of the unification algorithm used in Coq, which is central for working with advanced type inference features like Canonical Structures. The presentation of this journal version is incremental (it is presented feature by feature), with an aim of easing the understanding of how the algorithm actually works for users who want to take advantage of it. It has been accepted for publication in the Journal of Functional Programming.

5.5.4. Explicit Cumulativity

Pierre Letouzey started exploring with the help of Matthieu Sozeau a version of Coq's logic (CIC) where the cumulativity rule would be explicit. This cumulativity rule is a form of coercion between Coq universes, and is done silently in Coq up to now. Having a version of CIC where the use of the cumulativity bewteen Prop and Type is traceable woud be of great interest. In particular this would lead to a solid ground for the Coq extraction tool and solve some of its current limitations. Moreover, an explicit cumulativity would also help significantly the studies of Coq theoretical models. Preliminary results are encouraging, but this work has not been finalized yet. This work is related to the studies of Ali Assaf (Google Zurich, formerly PhD student in the team Deducteam), but uses different technical choices for different goals. This work is now pursued by Gaëtan Gilbert (PhD student of Nicolas Tabareau and Matthieu Sozeau at the École des Mines in Nantes), with the goal of providing a version of the calculus of constructions with definitional proof-irrelevance. The absence of explicit cumulativity between Prop and Type was identified in earlier work by Benjamin Werner and Giesik Lee as an important obstacle to building models of the theory, we hence expect this work to simplify the (relative) consistency proof of the theory.

5.6. Formalisation work

Participants: Jean-Jacques Lévy, Daniel de Rauglaudre.

5.6.1. Proofs of algorithms on graphs

Jean-Jacques Lévy and Chen Ran (a PhD student of the Institute of Software, Beijing, visiting the Toccata team) pursue their work about formal proofs of algorithms. Their goal is to provide proofs of algorithms which ought to be both checked by computer and easily human readable. If these kinds of proofs exist for algorithms on inductive structures or recursive algorithms on arrays, they seem less easy to design for combinatorial structures such as graphs. In 2016, they completed proofs for algorithms computing the strongly connected components in graphs. There are mainly two algorithms: one by Kosaraju (1978) working in two phases (some formal proofs of it have already been achieved by Pottier with Coq-classic and by Théry and Gonthier with Coq-ssreflect), one by Tarjan (1972) working in a single pass.

Their proofs use a first-order logic with definitions of inductive predicates. This logic is the one defined in Why3 (research-team Toccata, Saclay). They widely use automatic provers interfaced by Why3. A very minor part of these proofs is also achieved in Coq. The difficulty of this approach is to combine automatic provers and intuitive design.

Part of this work (Tarjan 1972) is presented at JFLA 2017 in Gourette [30] A more comprehensive version is under submission to another conference [34]. Scripts of proofs can be found at http://jeanjacqueslevy.net/why3.

5.6.2. Formalization of theorems in Coq

This section reports on formalisation work by Daniel de Rauglaudre.

5.6.2.1. Puiseux' Theorem

Puiseux' theorem states that the set of Puiseux series (series with rational powers) is an algebraically closed field, i.e. every non-constant polynomial with Puiseux series coefficients admits a zero. This theorem was formalized in Coq a couple of years ago, but it depended on five ad hoc axioms. This year, all these axioms have been grouped together into the only axiom LPO (Limited Principle of Omniscience), stating that for each sequence of booleans, we can decide whether it is always false or if there is at least one true element. This formalized theorem now depends only on this axiom.

5.6.2.2. Banach-Tarski Paradox

Banach-Tarski Paradox states that, if we admit the axiom of choice, a sphere is equidecomposable into two spheres identical to the initial one. The equidecomposability is a property of geometric objects: two objects (sets) are equidecomposable if we can partition them into a same finite number of sets, and each set of the first object is mapped to a set of the second object by only rotations and translations. In other words, we break the first object into a finite number of pieces, and with them, we reconstitute the second object. Its pen and paper proof was done in 1924 by Banach and Tarski.

Its formal proof in Coq has been started this year. About 80% of the proof has been done. The already proved part includes a lemma which says that the sphere without some specific countable number of points is equidecomposable into twice itself. It also includes a formal proof that equidecomposability is an equivalence relation. This makes about 7000 lines of Coq. The remaining part is to formalize the proof that the sphere is equidecomposable into the sphere without this countable set of points.

The version of axiom of choice used for this proof is named TTCA (Type Theoretical Axiom of Choice, introduced by Benjamin Werner [88]), stating that for each equivalence relation, there exists a function mapping each relation class to one of its elements.

6. Partnerships and Cooperations

6.1. National Initiatives

Alexis Saurin (coordinator) and Yann Régis-Gianas are members of the four-year RAPIDO ANR project, started in January 2015. 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 fixed points 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.

Pierre-Louis Curien (coordinator), Yves Guiraud (local coordinator), Philippe Malbos and Samuel Mimram have been members of the three-year Focal project of the IDEX Sorbonne Paris Cité (July 2013 to June 2016). This project, giving the support for the PhD grant of Cyrille Chenavier, concerns the interactions between higher-dimensional rewriting and combinatorial algebra. This project is joint with mathematicians form LAGA (Univ. Paris 13).

Pierre-Louis Curien (coordinator), Yves Guiraud (local coordinator), Philippe Malbos and Samuel Mimram are members of the four-year Cathre ANR project, started in January 2014. This project, giving the support for the PhD grant of Maxime Lucas, investigates the general theory of higher-dimensional rewriting, the development of a general-purpose library for higher-dimensional rewriting, and applications in the fields of combinatorial linear algebra, combinatorial group theory and theoretical computer science. This project is joint with mathematicians and computer scientists from LAGA (Univ. Paris 13), LIX (École Polytechnique), ICJ (Univ. Lyon 1 and Univ. Saint-Étienne), I2M (Univ. Aix-Marseille) and IMT (Univ. Toulouse 3).

Pierre-Louis Curien, Yves Guiraud, Hugo Herbelin, Philippe Malbos, Samuel Mimram and Alexis Saurin are members of the GDR Informatique Mathématique, in the Géocal (Geometry of computation) and LAC (Logic, algebra and computation) working groups.

Pierre-Louis Curien, Yves Guiraud (local coordinator), Philippe Malbos, Samuel Mimram 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.

Hugo Herbelin was the coordinator of the PPS site for the ANR Récré (January 2012 to mid 2016). Récré is about realisability and rewriting, with applications to proving with side-effects and concurrency.

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 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).

Matthieu Sozeau is a member of the CoqHoTT project led by Nicolas Tabareau (Ascola team, École des Mines de Nantes), funded by an ERC Starting Grant. The PhD grant of Gabriel Lewertowski was funded by the CoqHoTT ERC.

6.2. European Initiatives

6.2.1. FP7 & H2020 Projects

Hugo Herbelin is a deputy representative of France in the COST action EUTYPES.

6.3. International Initiatives

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

Pierre-Louis Curien participates to the Associated Team CRECOGI (Concurrent, Resourceful and Effectful Computation, by Geometry of Interaction) between the project-team Focus (Bologna) and the University of Tokyo (principal investigators Ugo dal Lago and Ichiro Hasuo, started in 2015).

6.3.2. Inria International Partners

6.3.2.1. Informal International Partners

The project-team has collaborations with University of Aarhus (Denmark), University of Oregon, University of Tokyo, University of Sovi Sad and the Institute of Mathematics of the Serbian Academy of Sciences, University of Nottingham, Institute of Advanced Study, MIT, University of Cambridge, and Universidad Nacional de Córdoba.

6.3.3. Participation in Other International Programs

Pierre-Louis Curien participates to the ANR International French-Chinese project LOCALI (Logical Approach to Novel Computational Paradigms), coordinated by Gilles Dowek (Deducteam).

6.4. International Research Visitors

6.4.1. Visits of International Scientists

Paolo Giarrusso (Univ. of Marburg) visited Yann Régis-Gianas in February 2016.

Lourdes del Carmen Gonzalez Huesca (Univ. of Mexico) visited Yann Régis-Gianas in December 2016.

6.4.2. Visits to International Teams

6.4.2.1. Research Stays Abroad

Pierre-Louis Curien visited the Category Theory group at Macquarie University in June-July 2016 (collaborative work on the combinatorial structure of type dependency).

As a part of his joint PhD, Étienne Miquey worked most of the year in Montevideo within the Logic group of the Universidad de la República of Uruguay.

7. Dissemination

7.1. Promoting Scientific Activities

7.1.1. Scientific Events Organisation

7.1.1.1. Member of the Organizing Committees

Yann Régis-Gianas is multimedia chair of the organizing committee of POPL 2017 that will be held in Paris in January 2017.

Yves Guiraud, Philippe Malbos and Samuel Mimram have organised the second edition of the Higher-Dimensional Rewriting and Applications (HDRA) workshop of the Formal Structures for Computation and Deduction conference (FSCD), held in Porto in June 2016. They plan to organise the third edition of HDRA, still with FSCD, in September 2017 in Oxford.

Yves Guiraud and Alexis Saurin, with Christine Tasson (IRIF), have organised the annual meeting of the Géocal and LAC working groups of the GDR Informatique Mathématique in Paris, in November 2016.

Yves Guiraud and Samuel Mimram, with Dimitri Ara (Univ. Aix-Marseille) are currently organising the Categories in Homotopy and Rewriting one-week conference, that will be held at the CIRM, in Marseille, in September 2017.

7.1.2. Scientific Events Selection

7.1.2.1. Member of the Conference Program Committees

Matthieu Sozeau was member of the program committees of FSCD'16, ITP'16 and CoqPL'16.

7.1.2.2. Member of the Conference Steering Committees

Hugo Herbelin is a member of the steering committee of the conference *Formal Structures for Computation and Deduction* (FSCD).

Pierre-Louis Curien is member of the steering committee of the international workshop Games for Logic and Programming Languages (GaLop).

Matthieu Sozeau is member of the steering committee of the Dependently Typed Programming international workshop (DTP).

7.1.3. Journal

7.1.3.1. 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).

7.1.3.2. Reviewer - Reviewing Activities

The members of the team reviewed papers for numerous journals and international conferences.

7.1.4. Invited Talks

Pierre-Louis Curien and Samuel Mimram gave invited talks at the annual meeting of the Géocal and LAC working groups of the GDR Informatique Mathématique (Paris, November).

Pierre-Louis Curien gave an invited talk at the annual meeting of the international ANR project Pace (between Univ. of Bologna, ENS Lyon and Shanghai Jiaotong University) on "Categorified cyclic operads" (Shanghai, November).

Hugo Herbelin gave an invited talk on “Proving with side-effects” at the Days in Logic meeting in Lisbon, Portugal.

Jean-Jacques Lévy gave an invited talk about “Strongly connected components in graphs, Formal proof of Tarjan 1972 algorithm” at the LTP (Langages, Types et Preuves) day, Saclay [38].

Matthieu Sozeau gave invited talks at the DeepSpec kickoff meeting in Princeton, NJ, USA, June 8th 2016, on “Coq 8.6” (together with Maxime Dénès), at the International Conference on Mathematical Software in Berlin, Germany, July 14th 2016, on “Coq for HoTT”, at the Categorical Logic and Univalent Foundations workshop, Leeds, UK, July 28th 2016, on “Forcing Translations in Type Theory”, and at the Coq Workshop in Nancy, France, August 26th 2016, on “Coq 8.6”.

7.1.5. Scientific Expertise

Pierre-Louis Curien has been member of the “Comité de Sélection” for a professor position in discrete mathematics at the University Paul Sabatier in Toulouse.

Yann Régis-Gianas and Hugo Herbelin have been members of the “Comité de Sélection” for an assistant professor position at CNAM in Paris.

Yann Régis-Gianas has been member of the “Comité de Sélection” for an assistant professor position at IRIF in Paris.

Hugo Herbelin has been member of the “Comité de Sélection” for a starting researcher position at Inria Saclay.

7.1.6. Scientific expertise

Pierre-Louis Curien is a member of the Scientific Committee of the CIRM (since June 2013).

7.1.7. Research Administration

Pierre-Louis Curien, Hugo Herbelin and Yves Guiraud are members of the scientific council of the Computer Science deparment of University Paris 7.

Yves Guiraud is the head of the Preuves, Programmes and Systèmes (PPS) team of the IRIF laboratory (since April 2016), and a member of the IRIF council (since January 2016).

7.1.8. Presentation of papers

Étienne Miquey gave a talk on a computational reduction of dependent choice in classical logic to system F at TYPES’16 (Novi Sad, Serbia, May 2016).

Étienne Miquey gave a talk on realizability games for the specification problem during the workshop Realizability in Uruguay 2016 (Piriápolis, Uruguay, July 2016).

Cyrille Chenavier gave a talk at the workshop IWC, Obergurgl, Austria (September 2016).

Cyrille Chenavier, Maxime Lucas and Jovana Obradović gave talks at the workshop Categories, Homotopy and Rewriting (Toulouse, January) and at the workshop HDRA (Porto, June).

Jovana Obradović presented her works on cyclic operads at the Types Conference 2016 (Novi Sad, Serbia, May 2016) and at the Conference Logic and Applications 2016 (Dubrovnik, Croatia, September).

Hugo Herbelin gave a talk on proving Gödel’s completeness theorem with side-effects at the Mathematics for Computation workshop in Niederalteich, Germany, May 2016.

7.1.9. Talks in seminars

Pierre-Louis Curien gave a talk at the Séminaire de Topologie of the University of Angers on the semantics of dependent types (January).

Yves Guiraud gave a talk in the Séminaire de Combinatoire of the University Paris 7 on an introduction to Squier’s theory (November).

Hugo Herbelin gave a talk on a proof-as-program interpretation of the classical axiom of dependent choice at the Séminaire “Logique et Interactions” of the “Logique de la Programmation” team of the “Institut de Mathématiques de Marseille” (University Aix-Marseille, February).

Yann Régis-Gianas gave a talk about control operators in the history of programming at the Séminaire “Code Sources” organized by Baptiste Mélès.

Yann Régis-Gianas gave a talk about the writing style in programming at the conference “Current issues in the philosophy of practice of mathematics and informatics” (University of Toulouse, April).

Thibaut Girka gave a talk about difference languages at the Gallium seminar (Paris, September 2016) and at the TLP group of the GDR GPL (Saclay, November 2016).

Yann Régis-Gianas gave a talk about difference languages at the LIMA laboratory (Nantes, October 2016) and at the Semantic Working Group of IRIF (Paris, December 2016).

Matthieu Sozeau gave a talk about Equations: a function definition toolbox for Coq at Dagstuhl in March 2016.

Cyrille Chenavier gave a talk about confluence algebras at the Algebra working group of the LMPA, Calais, in February 2016.

Jovana Obradović gave a talk about categorified cyclic operads at the Proof Theory Seminar of the Mathematical Institute of the Serbian Academy of Sciences and Arts (Belgrade, December 2016).

7.1.10. Attendance to conferences, workshops, schools,...

Pierre-Louis Curien attended the conferences Types 2016 in Novi Sad (Serbia, May) and Logic and Applications in Dubrovnik (Croatia, September).

Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud, Maxime Lucas, Philippe Malbos, Samuel Mimram and Jovana Obradović attended the Category, Homotopy and Rewriting workshop in Toulouse (January 2016).

Cyrille Chenavier, Maxime Lucas, Philippe Malbos and Samuel Mimram attended the HDRA workshop in Lisbon (June 2016).

Hugo Herbelin attended the Days in Logic meeting in Lisbon (Portugal, January), the Mathematics for Computation workshop in Niederalteich (Germany, May), the conferences Types 2016 in Novi Sad (Serbia, May), the Coq coding sprint in Sophia-Antipolis (May-June), the DeepSpec kick-off meeting in Princeton (USA, June), the FSCD conference in Porto (Portugal, June), the Coq workshop and ITP 2016 (Nancy, August), as well as the Dagstuhl seminar on universality of proofs (October).

Jean-Jacques Lévy participated to CPP and POPL 2016 conferences, Saint Petersburgh, USA, January 18-22, and the Robin Milner Award reception, the Royal Society, London, November 24 (X. Leroy (research team Gallium) was awarded).

Matthieu Sozeau attended POPL 2016, ICMS 2016, ITP 2016, the Coq coding sprint, the DeepSpec kick-off meeting in Princeton as well as the Dagstuhl seminar on proofs of functional programs (March).

Théo Zimmermann attended the conference CICM 2016 in Białystok (Poland, July). He gave a talk there to present his PhD subject. He also attended the Coq coding sprint.

7.1.11. Groupe de travail Théorie des types et réalisabilité

This is one of the working groups of PPS, jointly organised by Hugo Herbelin and Matthieu Sozeau.

7.1.12. Groupe de travail Catégories supérieures, polygraphes et homotopie

Several members of the team participate actively in this weekly working group of PPS, organised by François Métayer (IRIF) since 2009.

7.2. Teaching - Supervision - Juries

7.2.1. Teaching

Master: Pierre-Louis Curien teaches in the course Models of programming languages: domains, categories, games of the MPRI (together with Thomas Ehrhard and Paul-André Melliès).

Master: Hugo Herbelin teaches the course on the proof-as-program correspondence for classical logic and beyond at the LMFI.

Master: Pierre Letouzey teaches two short courses to the LMFI Master 2 students : “Models of programming” and ”Introduction to computed-aided formal proofs”. These two courses come in addition to Pierre Letouzey’s regular duty as teacher in the Computer Science department of Paris 7 (including a course on Compilation to M2-Pro students).

Master: Yann Régis-Gianas took part in the MPRI course entitled “Type systems”: he gave a 12-hour course about generalised algebraic data types, higher-order Hoare logic and dependently typed programming.

Master: Matthieu Sozeau taught the MPRI course on Advanced uses of proof assistants (12 hours + a project), together with Assia Mahboubi (Inria SpecFun).

MOOC: In collaboration with Roberto Di Cosmo and Ralf Treinen, Yann Régis-Gianas has created a MOOC about the OCaml programming language. The first edition took place in 2015, the second edition in 2016.

7.2.2. Supervision

Internship: Yves Guiraud has supervised the M2 internship of Amina Bendjaafar.

Internship: Hugo Herbelin has supervised the L3 internship of Meven Bertrand.

Internship: Hugo Herbelin has supervised the pre-doctoral internship of Théo Zimmermann.

Internship: Yann Régis-Gianas has supervised the M1 internship of Paul Laforgue.

Internship: Yann Régis-Gianas has supervised the M1 internship of Sylvain Ribstein.

PhD (completed): Cyrille Chenavier, supervised by Yves Guiraud and Philippe Malbos, successfully defended in December 2016

PhD in progress: Guillaume Claret, Programmation avec effets en Coq, (started in September 2012), supervised by Hugo Herbelin and Yann Régis-Gianas, defense planned in February 2017.

PhD in progress: Amina Doumane, supervised by Alexis Saurin, David Baelde and Pierre-Louis Curien.

PhD in progress: Thibaut Girka, Differential semantics (started in January 2014), supervised by Roberto Di Cosmo and Yann Régis-Gianas.

PhD in progress: Maxime Lucas, supervised by Yves Guiraud and Pierre-Louis Curien.

Phd in progress: Cyprien Mangin, Dependent Pattern-Matching, induction-induction and higher inductive types, September 2015, supervised by Matthieu Sozeau and Bruno Barras.

PhD in progress: Étienne Miquey, Réalisabilité classique et effets de bords, September 2014, supervised by Hugo Herbelin and Alexandre Miquel.

PhD in progress: Jovana Obradović, Cyclic operads: syntactic, algebraic and categorified aspects, supervised by Pierre-Louis Curien.

PhD stopped: Gabriel Lewertowski, On forcing in type theory, supervised by Matthieu Sozeau and Nicolas Tabareau. Gabriel stopped his PhD in september 2016 and is now working at la Pitié Salpêtrière as an engineer.

PhD starting: Gae¨tan Gilbert, Definitional Proof Irrelevance, supervised by Nicolas Tabareau and Matthieu Sozeau.

PhD starting: Théo Zimmermann, supervised by Hugo Herbelin.

7.2.3. Juries

Pierre-Louis Curien was referee for the habilitations of Emmanuel Haucourt (Paris 7, September) and Samuel Mimram (Paris 7, September). He was president of the jury of the thesis of Matteo Acclavio (Univ. de la Méditerranée, December).

Pierre-Louis Curien (president), Yves Guiraud and Philippe Malbos were members of the jury of the thesis of Cyrille Chenavier (Univ. Paris 7, December).

Hugo Herbelin was referee for the habilitation of Nicolas Tabareau (Nantes, November). He was a referee of the jury of the thesis of Jirka Maršík (LORIA, December).

Matthieu Sozeau was a member of the jury of the thesis of Kevin Quirin (EMN Nantes, December).

Yann Régis-Gianas is a member of the jury of the competitive examination for the entrance to the Écoles Normales Supérieures and the École Polytechnique.

7.3. Popularization

Yann Régis-Gianas co-organised the "Journée Francilienne de Programmation", a programming contest between undergraduate students of three universities of Paris (UPD, UPMC, UPS). Yann Régis-Gianas organised, and Étienne Miquey took part in the animation of the (computer science part of the) "Fête de la Science" event at the University Paris 7. Yann Régis-Gianas gave several presentations about "What is programming?" in primary and high schools of Paris and its region.

8. 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] P.-L. CURIEN.*Operads, clones, and distributive laws*, in "Operads and Universal Algebra : Proceedings of China-France Summer Conference", Tianjin, China, L. G. CHENGMING BAI, 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.

[4] 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.

[5] 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/.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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/.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] M. SOZEAU, N. OURY.*First-Class Type Classes*, in "Theorem Proving in Higher Order Logics, 21st International Conference, TPHOLs 2008, Montreal, Canada, August 18-21, 2008. Proceedings", O. A. MOHAMED, C. MUÑOZ, S. TAHAR (editors), Lecture Notes in Computer Science, Springer, 2008, vol. 5170, p. 278-293.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[19] 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.

Articles in International Peer-Reviewed Journal

[20] C. CHENAVIER.*Confluence Algebras and Acyclicity of the Koszul Complex*, in "Algebras and Representation Theory", 2016 [*DOI :* 10.1007/s10468-016-9595-6], https://hal.archives-ouvertes.fr/hal-01141738.

[21] P. DEHORNOY, Y. GUIRAUD.*Quadratic normalisation in monoids*, in "International Journal of Algebra and Computation", August 2016, vol. 26, n^{o} 5, p. 935-972 [*DOI :* 10.1142/S0218196716500399], https://hal.archives-ouvertes.fr/hal-01141226.

[22] M. GUILLERMO, É. MIQUEY.*Classical realizability and arithmetical formulæ*, in "Mathematical Structures in Computer Science", 2016 [*DOI :* 10.1017/S0960129515000559], https://hal.inria.fr/hal-01247989.

[23] Y. GUIRAUD, P. MALBOS.*Polygraphs of finite derivation type*, in "Mathematical Structures in Computer Science", September 2016, in press, 46 pages [*DOI :* 10.1017/S0960129516000220], https://hal.archives-ouvertes.fr/hal-00932845.

[24] F. LOULERGUE, W. BOUSDIRA, J. TESSON.*Calculating Parallel Programs in Coq using List Homomorphisms*, in "International Journal of Parallel Programming", 2016, 20, https://hal.inria.fr/hal-01159182.

[25] M. LUCAS.*A coherence theorem for pseudonatural transformations*, in "Journal of Pure and Applied Algebra", 2017, vol. 221, n^{o} 5, p. 1146-1217 [*DOI :* 10.1016/J.JPAA.2016.09.005], https://hal.archives-ouvertes.fr/hal-01191867.

International Conferences with Proceedings

[26] A. BAUER, G. JASON, P. LUMSDAINE, M. SHULMAN, M. SOZEAU, B. SPITTERS.*The HoTT Library: A Formalization of Homotopy Type Theory in Coq*, in "CPP'17", Paris, France, CPP'17, ACM, January 2017, 9 [*DOI :* 10.1145/3018610.3018615], https://hal.inria.fr/hal-01421212.

[27] P.-L. CURIEN, M. FIORE, G. MUNCH-MACCAGNONI.*A theory of effects and resources: adjunction models and polarised calculi*, in "Principles of Programmming Languages", Saint-Petersbourg, Florida, United States, Proceedings POPL 2016, January 2016 [*DOI :* 10.1145/2837614.2837652], https://hal.archives-ouvertes.fr/hal-01256092.

[28] G. JABER, G. LEWERTOWSKI, P.-M. PÉDROT, M. SOZEAU, N. TABAREAU.*The Definitional Side of the Forcing*, in "Logics in Computer Science", New York, United States, May 2016 [*DOI :* 10.1145/2933575.2935320], https://hal.archives-ouvertes.fr/hal-01319066.

[29] P.-M. PÉDROT, A. SAURIN.*Classical by-need*, in "European Symposium on Programming", Eindhoven, Netherlands, European Symposium on Programming, April 2016, https://hal.archives-ouvertes.fr/hal-01257348.

National Conferences with Proceeding

[30] R. CHEN, J.-J. LÉVY.*Une preuve formelle de l'algorithme de Tarjan-1972 pour trouver les composantes fortement connexes dans un graphe*, in "JFLA 2017 - Vingt-huitièmes Journées Francophones des Langages Applicatifs", Gourette, France, Vingt-huitièmes Journées Francophones des Langages Applicatifs, January 2017, https://hal.inria.fr/hal-01422215.

Conferences without Proceedings

[31] H. HERBELIN, É. MIQUEY.*A continuation-passing-style interpretation of simply-typed call-by-need λ-calculus with control within System F*, in "CL&C'16. Sixth International Workshop on. Classical Logic and Computation", Porto, Portugal, June 2016, https://hal.inria.fr/hal-01302696.

Research Reports

[32] 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.

Other Publications

[33] D. BAELDE, A. DOUMANE, A. SAURIN.*Infinitary proof theory : the multiplicative additive case* , June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01339037.

[34] R. CHEN, J.-J. LEVY.*Formal proofs of two algorithms for strongly connected components in graphs*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01422216.

[35] C. CHENAVIER.*Reduction Operators and Completion of Rewriting Systems*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01325907.

[36] P.-L. CURIEN, J. OBRADOVIC.*On the various definitions of cyclic operads*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01254649.

[37] A. DOUMANE, D. BAELDE, L. HIRSCHI, A. SAURIN.*Towards Completeness via Proof Search in the Linear Time mu-Calculus: The case of Büchi inclusions*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01275289.

[38] J.-J. LEVY, R. CHEN.*Strongly Connected Components in graphs, formal proof of Tarjan1972 algorithm*, November 2016, Groupe de travail LTP du GDR GPL , Travail présenté à JFLA 2017, https://hal.inria.fr/hal-01422227.

[39] É. MIQUEY.*A Classical Sequent Calculus with Dependent Types* , December 2016, working paper or preprint, https://hal.inria.fr/hal-01375977.

[40] J. OBRADOVIC.*The Bénabou-Roubaud monadic descent theorem via string diagrams*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01254637.

References in notes

[41] D. J. ANICK.*On the Homology of Associative Algebras*, in "Trans. Amer. Math. Soc.", 1986, vol. 296, n^{o} 2, p. 641–659.

[42] D. ARA, F. MÉTAYER.*The Brown-Golasiński Model Structure on strict ∞-groupoids revisited*, in "Homology, Homotopy and Applications", 2011, vol. 13, n^{o} 1, p. 121–142.

[43] J. BAEZ, A. CRANS.*Higher-dimensional algebra. VI. Lie 2-algebras*, in "Theory Appl. Categ.", 2004, vol. 12, p. 492–538.

[44] H. P. BARENDREGT.*The Lambda Calculus: Its Syntax and Semantics*, North Holland, Amsterdam, 1984.

[45] R. BERGER.*Confluence and Koszulity*, in "J. Algebra", 1998, vol. 201, n^{o} 1, p. 243–283.

[46] 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.

[47] 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.

[48] A. CHURCH.*A set of Postulates for the foundation of Logic*, in "Annals of Mathematics", 1932, vol. 2, p. 33, 346-366.

[49] T. COQUAND.*Une théorie des Constructions*, University Paris 7, January 1985.

[50] 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.

[51] 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.

[52] H. B. CURRY, R. FEYS, W. CRAIG.*Combinatory Logic*, North-Holland, 1958, vol. 1, §9E.

[53] P. DEHORNOY, F. DIGNE, E. GODELLE, D. KRAMMER, J. MICHEL.*Foundations of Garside theory*, EMS Tracts Math., European Mathematical Society, 2015, vol. 22, xiv + 689 pages.

[54] P. DELIGNE.*Action du groupe des tresses sur une catégorie*, in "Invent. Math.", 1997, vol. 128, n^{o} 1, p. 159–175.

[55] 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.

[56] 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.

[57] G. GENTZEN.*Untersuchungen über das logische Schließen*, in "Mathematische Zeitschrift", 1935, vol. 39, p. 176–210,405–431.

[58] 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.

[59] 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.

[60] Y. GUIRAUD.*Présentations d'opérades et systèmes de réécriture*, Univ. Montpellier 2, 2004.

[61] Y. GUIRAUD.*Termination Orders for 3-Dimensional Rewriting*, in "Journal of Pure and Applied Algebra", 2006, vol. 207, n^o 2, p. 341–371.

[62] Y. GUIRAUD.*The Three Dimensions of Proofs*, in "Annals of Pure and Applied Logic", 2006, vol. 141, n^o 1–2, p. 266–295.

[63] Y. GUIRAUD.*Two Polygraphic Presentations of Petri Nets*, in "Theoretical Computer Science", 2006, vol. 360, n^o 1–3, p. 124–146.

[64] 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.

[65] Y. GUIRAUD, E. HOFFBECK, P. MALBOS.*Linear polygraphs and Koszulity of algebras*, June 2014, 42 pages, https://hal.archives-ouvertes.fr/hal-01006220.

[66] Y. GUIRAUD, P. MALBOS.*Higher-dimensional categories with finite derivation type*, in "Theory Appl. Categ.", 2009, vol. 22, n^o 18, p. 420-478.

[67] 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.

[68] Y. GUIRAUD, P. MALBOS.*Coherence in monoidal track categories*, in "Math. Structures Comput. Sci.", 2012, vol. 22, n^o 6, p. 931–969.

[69] 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.

[70] 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.

[71] J.-L. KRIVINE.*A call-by-name lambda-calculus machine*, in "Higher Order and Symbolic Computation", 2005.

[72] J.-L. KRIVINE.*Un interpréteur du lambda-calcul*, 1986, Unpublished.

[73] Y. LAFONT.*Towards an Algebraic Theory of Boolean Circuits*, in "Journal of Pure and Applied Algebra", 2003, vol. 184, p. 257-310.

[74] 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.

[75] P. LANDIN.*The mechanical evaluation of expressions*, in "The Computer Journal", January 1964, vol. 6, n^{o} 4, p. 308–320.

[76] 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.

[77] M. LUCAS.*A cubical Squier's theorem*, 2016, 6 pages, arXiv:1612.06541.

[78] M. LUCAS.*Cubical (omega,p)-categories*, 2016, 38 pages, arXiv:1612.07050.

[79] P. MALBOS.*Critères de finitude homologique pour la non convergence des systèmes de réécriture de termes*, Univ. Montpellier 2, 2004.

[80] P. MARTIN-LÖF.*A theory of types*, University of Stockholm, 1971, n^{o} 71-3.

[81] 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.

[82] 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.

[83] 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.

[84] 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.

[85] 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.

[86] R. STREET.*Limits Indexed by Category-Valued 2-Functors*, in "Journal of Pure and Applied Algebra", 1976, vol. 8, p. 149–181.

[87] THE COQ DEVELOPMENT TEAM.*The Coq Reference Manual, version 8.2*, September 2008, http://coq.inria.fr/doc.

[88] B. WERNER.*Sets in types, types in sets*, M. ABADI, T. ITO (editors), Springer Berlin Heidelberg, Berlin, Heidelberg, 1997, p. 530–546, http://dx.doi.org/10.1007/BFb0014566.

[89] 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

Université Pierre et Marie Curie (Paris 6)

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

Keywords:

Computer Science and Digital Science:

2.4. - Verification, reliability, certification
4.3. - Cryptography
4.3.1. - Public key cryptography
6.2.6. - Optimization
6.2.7. - High performance computing
7.2. - Discrete mathematics, combinatorics
7.3. - Optimization
7.5. - Geometry, Topology
7.6. - Computer Algebra

Other Research Topics and Application Domains:

5. - Industry of the future
5.2. - Design and manufacturing
6. - IT and telecom
6.3. - Network functions
6.5. - Information systems
9.4.1. - Computer science
9.4.2. - Mathematics
9.8. - Privacy

1. Members

Research Scientists

Jean-Charles Faugère [Team leader, Inria, Senior Researcher, HDR]
Alain Jacquemard [Délégation Inria, Senior Researcher, Univ. Bourgogne, Professor, until Aug. 2016, HDR]
Elias Tsigaridas [Inria, Researcher]

Faculty Members

Jérémy Berthomieu [UPMC, Associate Professor]
Ludovic Perret [UPMC, Associate Professor, HDR]
Guénaël Renault [UPMC, Associate Professor, HDR]
Mohab Safey El Din [UPMC, Professor, HDR]

Technical Staff

Jérôme Govinden [SATT-LUTECH, until Mar. 2016]

PhD Students

Ivan Bannwarth [UPMC]
Matías Bender [Inria]
Ulrick Severin [Dassault Aviation, until Aug. 2016]
Thibaut Verron [UPMC, until Sep. 2016]
Alexandre Wallet [Inria, until Dec. 2016]

Visiting Scientist

Christian Eder [Technische Universität Kaiserslautern, Germany, regularly]

Administrative Assistants

Georgette Bonpapa [UPMC]
Laurence Bourcier [Inria]
Virginie Collette [Inria]
Irphane Khan [UPMC]
Nelly Maloisel [Inria]

Others

Daniel Lazard [UPMC, Professor, Émérite, HDR]
Emmanuel Prouff [ANSSI, Safran Identity and Security, Associate Member, HDR]
Dongming Wang [CNRS, Senior Researcher, Associate Member, HDR]
Sènan Dossa [ENS Lyon, Internship, from May 2016 until Sep. 2016]
Vincent Guisse [Min. de l'Éducation Nationale, Internship, from Apr. 2016 until Jul. 2016]
Ramon Ronzon [École polytechnique, Internship, from Mar. 2016 until Sep. 2016]

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 a building blocks for higher level algorithms who 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: Jean-Charles Faugère, Mohab Safey El Din, Elias Tsigaridas, Guénaël Renault, Dongming Wang, Jérémy Berthomieu, Thibaut Verron.

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, Daniel Lazard, Elias Tsigaridas, Ivan Bannwarth.

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.

[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.

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.

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, Christian Eder, Elias Tsigaridas.

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.FGB is an efficient library 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 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.

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: Jean-Charles Faugère, Ludovic Perret, Guénaël Renault, Jérémy Berthomieu.

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

The goal of the RISQ project is to prepare the security industry to the upcoming shift of classical cryptography to quantum-safe cryptography. The RISQ project is a massive effort at the French level to embrace the quantum-safe revolution. The project gather 15 partners : ANSSI, C&S, CEA, Crypto Experts, EADS, ENS Lyon, ENS Paris, Gemalto, Orange, PCQC, POLSYS (Inria de Paris), Université de Rennes, Secure IC, Thales CS, and Université de Versailles.

The RISQ project is certainly the biggest (in term of number of partners, as well as funding) industrial project ever organized in quantum-safe cryptography. RISQ is one of few projects accepted in the “Grands Défis du Numérique” which is managed by BPI France, and will be funded thanks to the PIA.

POLSYS actively participated to gather the partners of RISQ, and in defining the proposal. POLSYS will lead the academic effort in RISQ.

Jointly with LAAS (D. Henrion, S. Naldi), we have released a new MAPLE library SPECTRA for finding a real point $x = (x_1, ..., x_n)$ such that the symmetric matrix $A(x) = A_0 + A_1\,x_1 + \cdots + A_n\,x_n$ is positive semidefinite using exact arithmetic (see http://homepages.laas.fr/henrion/software/spectra/).

Our open source C library SLV has been officially released this year with a presentation at ISSAC. It aims at solating and approximating the real roots of univariate polynomials with integer coefficients (see http://www-polsys.lip6.fr/~elias/soft.html)

4.1.1. Awards

Matías Bender received the Distinguished Student Author Award of ISSAC2016 for his paper [22] written with J.-Ch. FAUGÈRE, L. PERRET and E. TSIGARIDAS.
BEST PAPERS AWARDS :
[22] **ISSAC '16 - 41st International Symposium on Symbolic and Algebraic Computation**. M. R. BENDER, J.-C. FAUGÈRE, L. PERRET, E. TSIGARIDAS.

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

FUNCTIONAL DESCRIPTION

FGb is a powerful software for computing Groebner 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 Faugère
- Contact: Jean-Charles Faugère
- URL: http://polsys.lip6.fr/~jcf/Software/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/Software/FGb/

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/Software/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. SLV

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 hundreds of Megabytes. Currently the code consists of $\sim 5\,000$ lines.

- Contact: Elias Tsigaridas
- URL: http://www-polsys.lip6.fr/~elias/soft

5.8. SPECTRA

Semidefinite Programming solved Exactly with Computational Tools of Real Algebra
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 Grobner bases and provides either certified numerical approximations of the solutions or exact representations of them.

- Contact: Mohab Safey El Din
- URL: http://homepages.laas.fr/henrion/software/spectra/

6. New Results

6.1. Fundamental algorithms and structured polynomial systems

6.1.1. Linear Algebra for Computing Gröbner Bases of Linear Recursive Multidimensional Sequences

The so-called Berlekamp – Massey – Sakata algorithm computes a Gröbner basis of a 0-dimensional ideal of relations satisfied by an input table. It extends the Berlekamp – Massey algorithm to n-dimensional tables, for $n > 1$.

In the extended version [6], we investigate this problem and design several algorithms for computing such a Gröbner basis of an ideal of relations using linear algebra techniques. The first one performs a lot of table queries and is analogous to a change of variables on the ideal of relations.

As each query to the table can be expensive, we design a second algorithm requiring fewer queries, in general. This FGLM-like algorithm allows us to compute the relations of the table by extracting a full rank submatrix of a *multi-Hankel* matrix (a multivariate generalization of Hankel matrices).

Under some additional assumptions, we make a third, adaptive, algorithm and reduce further the number of table queries. Then, we relate the number of queries of this third algorithm to the *geometry* of the final staircase and we show that it is essentially linear in the size of the output when the staircase is convex. As a direct application to this, we decode n-cyclic codes, a generalization in dimension n of Reed Solomon codes.

We show that the multi-Hankel matrices are heavily structured when using the LEX ordering and that we can speed up the computations using fast algorithms for quasi-Hankel matrices. Finally, we design algorithms for computing the generating series of a linear recursive table.

6.1.2. Guessing Linear Recurrence Relations of Sequence Tuples and P-recursive Sequences with Linear Algebra

Given several n-dimensional sequences, we first present in [23] an algorithm for computing the Gröbner basis of their module of linear recurrence relations.

A P-recursive sequence $(u_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^n}$ satisfies linear recurrence relations with polynomial coefficients in $\mathbf{i}$, as defined by Stanley in 1980. Calling directly the aforementioned algorithm on the tuple of sequences $\left((\mathbf{i}^{\mathbf{j}}\, u_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^n}\right)_{\mathbf{j}}$ for retrieving the relations yields redundant relations. Since the module of relations of a P-recursive sequence also has an extra structure of a 0-dimensional right ideal of an Ore algebra, we design a more efficient algorithm that takes advantage of this extra structure for computing the relations.

Finally, we show how to incorporate Gröbner bases computations in an Ore algebra $\mathbb{K}\langle t_1,...,t_n,x_1,...,x_n\rangle$, with commutators $x_k\,x_\ell - x_\ell\,x_k = t_k\,t_\ell - t_\ell\,t_k = t_k\,x_\ell - x_\ell\,t_k = 0$ for $k\neq\ell$ and $t_k\,x_k - x_k\,t_k = x_k$, into the algorithm designed for P-recursive sequences. This allows us to compute faster the Gröbner basis of the ideal spanned by the first relations, such as in 2D/3D-space walks examples.

6.1.3. On the Connection Between Ritt Characteristic Sets and Buchberger-Gröbner Bases

For any polynomial ideal I, let the minimal triangular set contained in the reduced Buchberger–Gröbner basis of I with respect to the purely lexicographical term order be called the W-characteristic set of I. In [18], we establish a strong connection between Ritt's characteristic sets and Buchberger's Gröbner bases of polynomial ideals by showing that the W-characteristic set C of I is a Ritt characteristic set of I whenever C is an ascending set, and a Ritt characteristic set of I can always be computed from C with simple pseudo-division when C is regular. We also prove that under certain variable ordering, either the W-characteristic set of I is normal, or irregularity occurs for the jth, but not the $(j+1)$th, elimination ideal of I for some j. In the latter case, we provide explicit pseudo-divisibility relations, which lead to nontrivial factorizations of certain polynomials in the Buchberger–Gröbner basis and thus reveal the structure of such polynomials. The pseudo-divisibility relations may be used to devise an algorithm to decompose arbitrary polynomial sets into normal triangular sets based on Buchberger–Gröbner bases computation.

6.1.4. On the complexity of computing Gröbner bases for weighted homogeneous systems

Solving polynomial systems arising from applications is frequently made easier by the structure of the systems. Weighted homogeneity (or quasi-homogeneity) is one example of such a structure: given a system of weights $W=(w_1,\cdots,w_n)$, W-homogeneous polynomials are polynomials which are homogeneous w.r.t the weighted degree $\deg_W(X_1^{\alpha_1},\cdots,X_n^{\alpha_n})=\sum w_i\alpha_i$.

Gröbner bases for weighted homogeneous systems can be computed by adapting existing algorithms for homogeneous systems to the weighted homogeneous case. In [12], we show that in this case, the complexity estimate for Algorithm F5 $\left(\binom{n+d_{\max}-1}{d_{\max}}^{\omega}\right)$ can be divided by a factor $\left(\prod w_i\right)^{\omega}$. For zero-dimensional systems, the complexity of Algorithm FGLMnD^{ω} (where D is the number of solutions of the system) can be divided by the same factor $\left(\prod w_i\right)^{\omega}$. Under genericity assumptions, for zero-dimensional weighted homogeneous systems of W-degree $(d_1,\cdots,d_n)$, these complexity estimates are polynomial in the weighted Bézout bound $\prod_{i=1}^n d_i/\prod_{i=1}^n w_i$.

Furthermore, the maximum degree reached in a run of Algorithm F5 is bounded by the weighted Macaulay bound $\sum(d_i-w_i)+w_n$, and this bound is sharp if we can order the weights so that $w_n=1$. For overdetermined semi-regular systems, estimates from the homogeneous case can be adapted to the weighted case.

We provide some experimental results based on systems arising from a cryptography problem and from polynomial inversion problems. They show that taking advantage of the weighted homogeneous structure yields substantial speed-ups, and allows us to solve systems which were otherwise out of reach.

6.1.5. A Superfast Randomized Algorithm to Decompose Binary Forms

Symmetric Tensor Decomposition is a major problem that arises in areas such as signal processing, statistics, data analysis and computational neuroscience. It is equivalent to a homogeneous polynomial in n variables of degree D as a sum of Dth powers of linear forms, using the minimal number of summands. This minimal number is called the rank of the polynomial/tensor. We consider the decomposition of binary forms, that corresponds to the decomposition of symmetric tensors of dimension 2 and order D. This problem has its roots in Invariant Theory, where the decompositions are known as canonical forms. As part of that theory, different algorithms were proposed for the binary forms. In recent years, those algorithms were extended for the general symmetric tensor decomposition problem. We present in [22] a new randomized algorithm that enhances the previous approaches with results from structured linear algebra and techniques from linear recurrent sequences. It achieves a softly linear arithmetic complexity bound. To the best of our knowledge, the previously known algorithms have quadratic complexity bounds.

6.1.6. On the Bit Complexity of Solving Bilinear Polynomial Systems

In [29] we bound the Boolean complexity of computing isolating hyperboxes for all complex roots of systems of bilinear polynomials. The resultant of such systems admits a family of determinantal Sylvester-type formulas, which we make explicit by means of homological complexes. The computation of the determinant of the resultant matrix is a bottleneck for the overall complexity. We exploit the quasi-Toeplitz structure to reduce the problem to efficient matrix-vector products, corresponding to multivariate polynomial multiplication. For zero-dimensional systems, we arrive at a primitive element and a rational univariate representation of the roots. The overall bit complexity of our probabilistic algorithm is $\widetilde{O}_B(n^4D^4 + n^2D^4\tau)$, where n is the number of variables, D equals the bilinear Bézout bound, and τ is the maximum coefficient bitsize. In addition, a careful infinitesimal symbolic perturbation of the system allows us to treat degenerate and positive dimensional systems, thus making our algorithms and complexity analysis applicable to the general case.

6.2. Solving Systems over the Reals and Applications

6.2.1. Exact algorithms for linear matrix inequalities

Let $A(x) = A_0 + x_1A_1 + ... + x_nA_n$ be a linear matrix, or pencil, generated by given symmetric matrices $A_0, A_1, ..., A_n$ of size m with rational entries. The set of real vectors x such that the pencil is positive semidefinite is a convex semi-algebraic set called spectrahedron, described by a linear matrix inequality (LMI). In [13], we design an exact algorithm that, up to genericity assumptions on the input matrices, computes an exact algebraic representation of at least one point in the spectrahedron, or decides that it is empty. The algorithm does not assume the existence of an interior point, and the computed point minimizes the rank of the pencil on the spectrahedron. The degree d of the algebraic representation of the point coincides experimentally with the algebraic degree of a generic semidefinite program associated to the pencil. We provide explicit bounds for the complexity of our algorithm, proving that the maximum number of arithmetic operations that are performed is essentially quadratic in a multilinear Bézout bound of d. When m (resp. n) is fixed, such a bound, and hence the complexity, is polynomial in n (resp. m). We conclude by providing results of experiments showing practical improvements with respect to state-of-the-art computer algebra algorithms.

6.2.2. Real root finding for determinants of linear matrices

Let $A_0, A_1, ..., A_n$ be given square matrices of size m with rational coefficients. In [14], we focus on the exact computation of one point in each connected component of the real determinantal variety $\{x \in \mathbb{R}^n : \det(A_0 + x_1A_1 + \cdots + x_nA_n) = 0\}$. Such a problem finds applications in many areas such as control theory, computational geometry, optimization, etc. Using standard complexity results this problem can be solved using $m^{O(n)}$ arithmetic operations. Under some genericity assumptions on the coefficients of the matrices, we provide an algorithm solving this problem whose runtime is essentially quadratic in $\binom{n+m}{n}^3$. We also report on experiments with a computer implementation of this algorithm. Its practical performance illustrates the complexity estimates. In particular, we emphasize that for subfamilies of this problem where m is fixed, the complexity is polynomial in n.

6.2.3. A nearly optimal algorithm for deciding connectivity queries in smooth and bounded real algebraic sets

A roadmap for a semi-algebraic set S is a curve which has a non-empty and connected intersection with all connected components of S. Hence, this kind of object, introduced by Canny, can be used to answer connectivity queries (with applications, for instance, to motion planning) but has also become of central importance in effective real algebraic geometry, since it is used in higher-level algorithms. In [15], we provide a probabilistic algorithm which computes roadmaps for smooth and bounded real algebraic sets. Its output size and running time are polynomial in $(nD)^{n\log(d)}$, where D is the maximum of the degrees of the input polynomials, d is the dimension of the set under consideration and n is the number of variables. More precisely, the running time of the algorithm is essentially subquadratic in the output size. Even under our assumptions, it is the first roadmap algorithm with output size and running time polynomial in $(nD)^{n\log(d)}$.

6.2.4. Determinantal sets, singularities and application to optimal control in medical imagery

Control theory has recently been involved in the field of nuclear magnetic resonance imagery. The goal is to control the magnetic field optimally in order to improve the contrast between two biological matters on the pictures. Geometric optimal control leads us here to analyze mero-morphic vector fields depending upon physical parameters, and having their singularities defined by a determinantal variety. The involved matrix has polynomial entries with respect to both the state variables and the parameters. Taking into account the physical constraints of the problem, one needs to classify, with respect to the parameters, the number of real singularities lying in some prescribed semi-algebraic set. In [24], we develop a dedicated algorithm for real root classification of the singularities of the rank defects of a polynomial matrix, cut with a given semi-algebraic set. The algorithm works under some genericity assumptions which are easy to check. These assumptions are not so restrictive and are satisfied in the aforementioned application. As more general strategies for real root classification do, our algorithm needs to compute the critical loci of some maps, intersections with the boundary of the semi-algebraic domain, etc. In order to compute these objects, the determinantal structure is exploited through a stratification by the rank of the polynomial matrix. This speeds up the computations by a factor 100. Furthermore, our implementation is able to solve the application in medical imagery, which was out of reach of more general algorithms for real root classification. For instance, computational results show that the contrast problem where one of the matters is water is partitioned into three distinct classes.

6.2.5. Optimal Control of an Ensemble of Bloch Equations with Applications in MRI

The optimal control of an ensemble of Bloch equations describing the evolution of an ensemble of spins is the mathematical model used in Nuclear Resonance Imaging and the associated costs lead to consider Mayer optimal control problems. The Maximum Principle allows to parameterize the optimal control and the dynamics is analyzed in the framework of geometric optimal control. This leads to numerical implementations or suboptimal controls using averaging principle as presented in [25].

6.2.6. Critical Point Computations on Smooth Varieties: Degree and Complexity bounds

Let $V \subset \mathbb{C}^n$ be an equidimensional algebraic set and g be an n-variate polynomial with rational coefficients. Computing the critical points of the map that evaluates g at the points of V is a cornerstone of several algorithms in real algebraic geometry and optimization. Under the assumption that the critical locus is finite and that the projective closure of V is smooth, we provide in [31] sharp upper bounds on the degree of the critical locus which depend only on $\deg(g)$ and the degrees of the generic polar varieties associated to V. Hence, in some special cases where the degrees of the generic polar varieties do not reach the worst-case bounds, this implies that the number of critical points of the evaluation map of g is less than the currently known degree bounds. We show that, given a lifting fiber of V, a slight variant of an algorithm due to Bank, Giusti, Heintz, Lecerf, Matera and Solernó computes these critical points in time which is quadratic in this bound up to logarithmic factors, linear in the complexity of evaluating the input system and polynomial in the number of variables and the maximum degree of the input polynomials.

6.3. Solving Systems in Finite Fields, Applications in Cryptology and Algebraic Number Theory.

6.3.1. Structural Cryptanalysis of McEliece Schemes with Compact Key.

A very popular trend in code-based cryptography is to decrease the public-key size by focusing on subclasses of alternant/Goppa codes which admit a very compact public matrix, typically quasi-cyclic (QC), quasi-dyadic (QD), or quasi-monoidic (QM) matrices. We show in [11] that the very same reason which allows to construct a compact public-key makes the key-recovery problem intrinsically much easier. The gain on the public-key size induces an important security drop, which is as large as the compression factor p on the public-key. The fundamental remark is that from the $k \times n$ public generator matrix of a compact McEliece, one can construct a $k/p \times n/p$ generator matrix which is – from an attacker point of view – as good as the initial public-key. We call this new smaller code the *folded code*. Any key-recovery attack can be deployed equivalently on this smaller generator matrix. To mount the key-recovery in practice, we also improve the algebraic technique of Faugère, Otmani, Perret and Tillich (FOPT). In particular, we introduce new algebraic equations allowing to include codes defined over any prime field in the scope of our attack. We describe a so-called "structural elimination" which is a new algebraic manipulation which simplifies the key-recovery system. As a proof of concept, we report successful attacks on many cryptographic parameters available in the literature. All the parameters of CFS-signatures based on QD/QM codes that have been proposed can be broken by this approach. In most cases, our attack takes few seconds (the hardest case requires less than 2 hours). In the encryption case, the algebraic systems are harder to solve in practice. Still, our attack succeeds against several cryptographic challenges proposed for QD and QM encryption schemes. We mention that some parameters that have been proposed in the literature remain out of reach of the methods given here. weakness arising from Goppa codes with QM or QD symmetries. Indeed, the security of such schemes is not relying on the bigger compact public matrix but on the small folded code which can be efficiently broken in practice with an algebraic attack for a large set of parameters

6.3.2. Folding Alternant and Goppa Codes with Non-Trivial Automorphism Groups

The main practical limitation of the McEliece public-key encryption scheme is probably the size of its key. A famous trend to overcome this issue is to focus on subclasses of alternant/Goppa codes with a non trivial automorphism group. Such codes display then *symmetries* allowing compact parity-check or generator matrices. For instance, a key-reduction is obtained by taking *quasi-cyclic* (QC) or *quasi-dyadic* (QD) alternant/Goppa codes. We show in [10], that the use of such *symmetric* alternant/Goppa codes in cryptography introduces a fundamental weakness. It is indeed possible to reduce the key-recovery on the original symmetric public-code to the key-recovery on a (much) smaller code that has no symmetry anymore. This result is obtained thanks to an operation on codes called *folding* that exploits the knowledge of the automorphism group. This operation consists in adding the coordinates of codewords which belong to the same orbit under the action of the automorphism group. The advantage is twofold: the reduction factor can be as large as the size of the orbits, and it preserves a fundamental property: folding the dual of an alternant (*resp*. Goppa) code provides the dual of an alternant (*resp*. Goppa) code. A key point is to show that all the existing constructions of alternant/Goppa codes with symmetries follow a common principal of taking codes whose support is globally invariant under the action of affine transformations (by building upon prior works of T. Berger and A. Dür). This enables not only to present a unified view but also to generalize the construction of QC,QD and even *quasi-monoidic* (QM) Goppa codes. Lastly, our results can be harnessed to boost up any key-recovery attack on McEliece systems based on symmetric alternant or Goppa codes, and in particular algebraic attacks.

6.3.3. Factoring $N = p^r q^s$ for Large r and s

D. Boneh, G. Durfee, and N. Howgrave-Graham showed at Crypto 99 that moduli of the form $N = p^r q$ can be factored in polynomial time when $r \simeq \log p$. Their algorithm is based on Coppersmith's technique for finding small roots of polynomial equations. In [27], we show that $N = p^r q^s$ can also be factored in polynomial time when r or s is at least $(\log p)^3$; therefore we identify a new class of integers that can be efficiently factored.

We also generalize our algorithm to moduli equal to a product of k factors of prime powers $p_i^{r_i}$; we show that a non-trivial factor of N can be extracted in polynomial-time if one of the exponents r_i is large enough.

6.3.4. On the p-adic stability of the FGLM algorithm

Nowadays, many strategies to solve polynomial systems use the computation of a Gröbner basis for the graded reverse lexicographical ordering, followed by a change of ordering algorithm to obtain a Gröbner basis for the lexicographical ordering. The change of ordering algorithm is crucial for these strategies. In [33], we study the p-adic stability of the main change of ordering algorithm, FGLM. We show that FGLM is stable and give explicit upper bound on the loss of precision occuring in its execution. The variant of FGLM designed to pass from the grevlex ordering to a Gröbner basis in shape position is also stable. Our study relies on the application of Smith Normal Form computations for linear algebra.

6.3.5. Binary Permutation Polynomial Inversion and Application to Obfuscation Techniques

Whether it is for constant obfusation, opaque predicate or equation obfuscation, Mixed Boolean-Arithmetic (MBA) expressions are a powerful tool providing concrete ways to achieve obfuscation. Recent results introduced ways to mix such a tool with permutation polynomials modulo 2^n in order to make the obfuscation technique more resilient to SMT solvers. However, because of limitations regarding the inversion of such permutations, the set of permutation polynomials presented suffers some restrictions. Those restrictions allow several methods of arithmetic simplification, decreasing the effectiveness of the technique at hiding information. In [19], we present general methods for permutation polynomials inversion. These methods allow us to remove some of the restrictions presented in the literature, making simplification attacks less effective. We discuss complexity and limits of these methods, and conclude that not only current simplification attacks may not be as effective as we thought, but they are still many uses of polynomial permutations in obfuscation that are yet to be explored.

6.3.6. Horizontal Side-Channel Attacks and Countermeasures on the ISW Masking Scheme

A common countermeasure against side-channel attacks consists in using the masking scheme originally introduced by Ishai, Sahai and Wagner (ISW) at Crypto 2003, and further generalized by Rivain and Prouff at CHES 2010. The countermeasure is provably secure in the probing model, and it was showed by Duc, Dziembowski and Faust at Eurocrypt 2014 that the proof can be extended to the more realistic noisy leakage model. However the extension only applies if the leakage noise increases at least linearly with the masking order n, which is not necessarily possible in practice. In [20], we investigate the security of an implementation when the previous condition is not satisfied, for example when the masking order n increases for a constant noise. We exhibit two (template) horizontal side-channel attacks against the Rivain-Prouff's secure multiplication scheme and we analyze their efficiency thanks to several simulations and experiments. Eventually, we describe a variant of Rivain-Prouff's multiplication that is still provably secure in the original ISW model, and also heuristically secure against our new attacks.

6.3.7. Faster Evaluation of SBoxes via Common Shares

In [28], we describe a new technique for improving the efficiency of the masking countermeasure against side-channel attacks. Our technique is based on using common shares between secret variables, in order to reduce the number of finite field multiplications. Our algorithms are proven secure in the ISW probing model with $n > t + 1$ shares against t probes. For AES, we get an equivalent of 2.8 non-linear multiplications for every SBox evaluation, instead of 4 in the Rivain-Prouff countermeasure. We obtain similar improvements for other block-ciphers. Our technique is easy to implement and performs relatively well in practice, with roughly a 20% speed-up compared to existing algorithms.

6.3.8. Information Extraction in the Presence of Masking with Kernel Discriminant Analysis

To reduce the memory and timing complexity of the Side-Channel Attacks (SCA), dimensionality reduction techniques are usually applied to the measurements. They aim to detect the so-called Points of Interest (PoIs), which are time samples which (jointly) depend on some sensitive information (e.g. secret key sub-parts), and exploit them to extract information. The extraction is done through the use of functions which combine the measurement time samples. Examples of combining functions are the linear combinations provided by the Principal Component Analysis or the Linear Discriminant Analysis. When a masking countermeasure is properly implemented to thwart SCAs, the selection of PoIs is known to be a hard task: almost all existing methods have a combinatorial complexity explosion, since they require an exhaustive search among all possible d-tuples of points. In this paper we propose an efficient method for informative feature extraction in presence of masking countermeasure. This method, called Kernel Discriminant Analysis, consists in completing the Linear Discriminant Analysis with a so-called kernel trick, in order to efficiently perform it over the set of all possible d-tuples of points without growing in complexity with d. We identify and analyse the issues related to the application of such a method. Afterwards, its performances are compared to those of the Projection Pursuit (PP) tool for PoI selection up to a 4th-order context. Experiments show that the Kernel Discriminant Analysis remains effective and efficient for high-order attacks, leading to a valuable alternative to the PP in constrained contexts where the increase of the order d does not imply a growth of the profiling datasets.

6.3.9. Polynomial Evaluation and Side Channel Analysis

Side Channel Analysis (SCA) is a class of attacks that exploits leakage of information from a cryptographic implementation during execution. To thwart it, masking is a common countermeasure. The principle is to randomly split every sensitive intermediate variable occurring in the computation into several shares and the number of shares, called the masking order, plays the role of a security parameter. The main issue while applying masking to protect a block cipher implementation is to specify an efficient scheme to secure the S-box computations. Several masking schemes, applicable for arbitrary orders, have been recently introduced. Most of them follow a similar approach originally introduced in the paper of Carlet et al published at FSE 2012; the S-box to protect is viewed as a polynomial and strategies are investigated which minimize the number of field multiplications which are not squarings. The paper [32] aims at presenting all these works in a comprehensive way. The methods are discussed, their differences and similarities are identified and the remaining open problems are listed.

6.3.10. Redefining the Transparency Order

In [7], we consider the multi-bit Differential Power Analysis (DPA) in the Hamming weight model. In this regard, we revisit the definition of Transparency Order (TO) from the work of Prouff (FSE 2005) and find that the definition has certain limitations. Although this work has been quite well referred in the literature, surprisingly, these limitations remained unexplored for almost a decade. We analyse the definition from scratch, modify it and finally provide a definition with better insight that can theoretically capture DPA in Hamming weight model for hardware implementation with precharge logic. At the end, we confront the notion of (revised) transparency order with attack simulations in order to study to what extent the low transparency order of an s-box impacts the efficiency of a side channel attack against its processing.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants with Industry

Until the mid 2000's, multivariate cryptography was developing very rapidly, producing many interesting and versatile public-key schemes. However, many of them were soon successfully cryptanalysed (a lot have been done in this group). As a consequence, the confidence in multivariate cryptography cryptosystems declined. It seems that there have emerged new important reasons for renewal of the interest in a new generation of multivariate schemes. In the past two years, the algorithms for solving the Discrete Logarithm Problem over small characteristic fields underwent an extraordinary development. This clearly illustrates the risk to not consider alternatives to classical assumptions based on number theory. In parallel, two of the most important standardization bodies in the world, NIST and ETSI have recently started initiatives for developing cryptographic standards not based on number theory, with a particular focus on primitives resistant to quantum algorithms. An objective here is then to focus on the design of multivariate schemes.

The team is now involved in the industrial transfer of post-quantum cryptography. The project is supervised by SATT-LUTECH. SATT-LUTECH specializes 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, Université Panthéon-Assas, Paris Sorbonne University and National School of Industrial Creation).

The team has recently developed, in partnership with a mobile application development company (WASSA), an Android app for smartphones (Samsung G5 type) that uses multivariate cryptography. The application has been tested mid-November in a series of experiments supervised by DGA and French Ministry of Defense. The experiment gathered a total of hundred participants from various operational units. This is a first milestone in the maturation project whose goal is to create a start-up.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

- **ANR Grant HPAC: High Performance Algebraic Computing (2012-2016).** The pervasive ubiquity of parallel architectures and memory hierarchy has led to a new quest for parallel mathematical algorithms and software capable of exploiting the various levels of parallelism: from hardware acceleration technologies (multi-core and multi-processor system on chip, GPGPU, FPGA) to cluster and global computing platforms. For giving a greater scope to symbolic and algebraic computing, beyond the optimization of the application itself, the effective use of a large number of resources (memory and specialized computing units) is expected to enhance the performance multi-criteria objectives: time, resource usage, reliability, even energy consumption. The design and the implementation of mathematical algorithms with provable, adaptive and sustainable performance is a major challenge. In this context, this project is devoted to fundamental and practical research specifically in exact linear algebra and system solving that are two essential "dwarfs" (or "killer kernels") in scientific and algebraic computing. The project should lead to progress in matrix algorithms and challenge solving in cryptology, and should provide new insights into high performance programming and library design problems (J.-C. Faugère [contact], L. Perret, G. Renault, M. Safey El Din).
- **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).

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. A3

Type: PEOPLE

Instrument: Career Integration Grant

Duration: May 2013 - April 2017

Coordinator: Jean-Charles Faugère

Partner: Institut National de Recherche en Informatique et en Automatique (Inria), France

Inria contact: Elias Tsigaridas

Abstract: The project Algebraic Algorithms and Applications (A3) is an interdisciplinary and multidisciplinary project, with strong international synergy. It consists of four work packages The first (Algebraic Algorithms) focuses on fundamental problems of computational (real) algebraic geometry: effective zero bounds, that is estimations for the minimum distance of the roots of a polynomial system from zero, algorithms for solving polynomials and polynomial systems, derivation of non-asymptotic bounds for basic algorithms of real algebraic geometry and application of polynomial system solving techniques in optimization. We propose a novel approach that exploits structure and symmetry, combinatorial properties of high dimensional polytopes and tools from mathematical physics. Despite the great potential of the modern tools from algebraic algorithms, their use requires a combined effort to transfer this technology to specific problems. In the second package (Stochastic Games) we aim to derive optimal algorithms for computing the values of stochastic games, using techniques from real algebraic geometry, and to introduce a whole new arsenal of algebraic tools to computational game theory. The third work package (Non-linear Computational Geometry), we focus on exact computations with implicitly defined plane and space curves. These are challenging problems that commonly arise in geometric modeling and computer aided design, but they also have applications in polynomial optimization. The final work package (Efficient Implementations) describes our plans for complete, robust and efficient implementations of algebraic algorithms.

8.2.2. Collaborations in European Programs, Except FP7 & H2020

Program: COST

Project acronym: CryptoAction

Project title: Cryptography for Secure Digital Interaction

Duration: 04 2014 - 04 2018

Coordinator: Claudio ORLANDI

Abstract: As increasing amounts of sensitive data are exchanged and processed every day on the Internet, the need for security is paramount. Cryptography is the fundamental tool for securing digital interactions, and allows much more than secure communication: recent breakthroughs in cryptography enable the protection - at least from a theoretical point of view - of any interactive data processing task. This includes electronic voting, outsourcing of storage and computation, e-payments, electronic auctions, etc. However, as cryptography advances and becomes more complex, single research groups become specialized and lose contact with "the big picture". Fragmentation in this field can be dangerous, as a chain is only as strong as its weakest link. To ensure that the ideas produced in Europe’s many excellent research groups will have a practical impact, coordination among national efforts and different skills is needed. The aim of this COST Action is to stimulate interaction between the different national efforts in order to develop new cryptographic solutions and to evaluate the security of deployed algorithms with applications to the secure digital interactions between citizens, companies and governments. The Action will foster a network of European research centers thus promoting movement of ideas and people between partners.

Program: COST

Project acronym: CRYPTACUS

Project title: Cryptanalysis of ubiquitous computing systems

Duration: 12 2014 - 12 2018

Coordinator: Gildas AVOINE

Abstract: Recent technological advances in hardware and software have irrevocably affected the classical picture of computing systems. Today, these no longer consist only of connected servers, but involve a wide range of pervasive and embedded devices, leading to the concept of “ubiquitous computing systems”. The objective of the Action is to improve and adapt the existent cryptanalysis methodologies and tools to the ubiquitous computing framework. Cryptanalysis, which is the assessment of theoretical and practical cryptographic mechanisms designed to ensure security and privacy, will be implemented along four axes: cryptographic models, cryptanalysis of building blocks, hardware and software security engineering, and security assessment of real-world systems. Researchers have only recently started to focus on the security of ubiquitous computing systems. Despite the critical flaws found, the required highly-specialized skills and the isolation of the involved disciplines are a true barrier for identifying additional issues. The Action will establish a network of complementary skills, so that expertise in cryptography, information security, privacy, and embedded systems can be put to work together. The outcome will directly help industry stakeholders and regulatory bodies to increase security and privacy in ubiquitous computing systems, in order to eventually make citizens better protected in their everyday life.

8.3. International Initiatives

8.3.1. Inria International Labs

8.3.1.1. GOAL

Title: Geometry and Optimization with ALgebraic methods.

International Partner (Institution - Laboratory - Researcher):

University of California Berkeley (United States) - Dept. of Mathematics - Bernd Sturmfels

Start year: 2015

See also: http://www-polsys.lip6.fr/GOAL/index.html

Polynomial optimization problems form a subclass of general global optimization problems, which have received a lot of attention from the research community recently; various solution techniques have been designed. One reason for the spectacular success of these methods is the potential impact in many fields: data mining, big data, energy savings, etc. More generally, many areas in mathematics, as well as applications in engineering, biology, statistics, robotics etc. require a deeper understanding of the algebraic structure of their underlying objects.

A new trend in the polynomial optimization community is the combination of algebraic and numerical methods. Understanding and characterizing the algebraic properties of the objects occurring in numerical algorithms can play an important role in improving the efficiency of exact methods. Moreover, this knowledge can be used to estimate the quality (for example the number of significant digits) of numerical algorithms. In many situations each coordinate of the optimum is an algebraic number. The degree of the minimal polynomials of these algebraic numbers is the Algebraic Degree of the problem. From a methodological point of view, this notion of Algebraic Degree emerges as an important complexity parameter for both numerical and the exact algorithms. However, algebraic systems occurring in applications often have special algebraic structures that deeply influence the geometry of the solution set. Therefore, the (true) algebraic degree could be much less than what is predicted by general worst case bounds (using Bézout bounds, mixed volume, etc.), and would be very worthwhile to understand it more precisely.

The goal of this proposal is to develop algorithms and mathematical tools to solve geometric and optimization problems through algebraic techniques. As a long-term goal, we plan to develop new software to solve these problems more efficiently. These objectives encompass the challenge of identifying instances of these problems that can be solved in polynomial time with respect to the number of solutions and modeling these problems with polynomial equations.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

Carlos Améndola Cerón

Date: May 2016

Institution: Technische Universität Berlin, Germany

Christoph Koutschan

Date: Nov. 2016

Institution: Österreichische Akademie der Wissenschaften, Linz

Didier Henrion

Date: Nov. 2016

Institution: LAAS, CNRS

Simone Naldi

Date: Nov. 2016

Institution: TU Univ. Dortmund, Germany.

Ioannis Psarros

Date: May. 2016

Institution: University of Athens, Greece.

8.4.1.1. Internships

Vincent Guisse

Date: Apr. 2016 - Jul. 2016

Institution: Université Paris – Diderot

Supervisor: Jean-Charles Faugère, Jérémy Berthomieu

Ramon Ronzon

Date: Mar. 2016 - Sep. 2016

Institution: École polytechnique

Supervisor: Jean-Charles Faugère, Ludovic Perret

Sènan Dossa

Date: Mar. 2016 - Sep. 2016

Institution: ENS Lyon

Supervisor: Jean-Charles Faugère, Ludovic Perret

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific events organisation

9.1.1.1. Member of the organizing committees

Dongming Wang was involved in the organization of the following conferences

- Special Session on Software of Polynomial Systems at the 5th International Congress on Mathematical Software (ICMS 2016) (Berlin, Germany, July 11-14, 2016).

9.1.2. Scientific events selection

9.1.2.1. Member of the conference program committees

Emmanuel Prouff was member of the program committees of the following conferences

- Conference on Cryptographic Hardware and Embedded Systems 2016 (CHES 2016) (Santa Barbara, CA, USA, Aug. 17-19, 2016);
- Smart Card Research and Advanced Application Conference (CARDIS 2016) (Cannes, France, Nov. 7-9, 2016);
- International Workshop on Constructive Side-Channel Analysis and Secure Design (COSADE 2016) (Graz, Austria, Apr. 14-15);
- 23rd ACM Conference on Computer and Communications Security (ACM CCS 2016) (Vienna, Austria, Oct. 24-28).

Dongming Wang was member of the program committees of the following conferences

- 11th International Workshop on Automated Deduction in Geometry (ADG 2016) (Strasbourg, France, June 27-29, 2016);
- 7th International Symposium on Symbolic Computation in Software Science (SCSS 2016) (Tokyo, Japan, March 28-31, 2016).

Elias Tsigaridas was member of the program committees of the following conferences

- Computer Algebra in Scientific Computing (CASC 2016), Sept 2016 Bucharest, Romania.

9.1.3. Journal

9.1.3.1. Member of the editorial boards

Ludovic Perret is Member of the Editorial Board of Designs, Codes and Cryptography.

Emmanuel Prouff is member of the editorial board of Journal of Cryptographic Engineering.

Mohab Safey El Din is member of the editorial board of Journal of Symbolic Computation.

Dongming Wang has the following editorial activities:

- Editor-in-Chief and Managing Editor for the journal
  Mathematics in Computer Science (published by Birkhäuser/Springer, Basel).
- Executive Associate Editor-in-Chief for the journal
  SCIENCE CHINA Information Sciences (published by Science China Press, Beijing and Springer, Berlin).
- Member of the Editorial Boards for the
  - Journal of Symbolic Computation (published by Academic Press/Elsevier, London),
  - Frontiers of Computer Science (published by Higher Education Press, Beijing and Springer, Berlin),
  - Texts and Monographs in Symbolic Computation (published by Springer, Wien New York),
- Member of the International Advisory Board for the Communications of JSSAC (Japan Society for Symbolic and Algebraic Computation) (published by JSSAC).

9.1.4. Invited talks

Emmanuel Prouff was invited speaker at

- EUROCRYPT 2016 (invited tutorial), Vienna, Austria, on Securing Cryptography Implementations in Embedded Systems.
- SPACE 2016 (invited speaker), Hyderabad, India on Breaking Cryptographic Implementations Using Deep Learning Techniques.

Mohab Safey El Din was invited speaker at

- the SMAI-MODE session on semi-algebraic optimization, Toulouse, March 2016, France.
- the AIM Workshop on Algebraic Vision which was held at the American Institute of Mathematics, San Jose, May 2016, USA.
- the NCSU seminar on Symbolic Computation, Raleigh, May 2016, USA.
- the PGMO session on Semi-Definite Programming, Palaiseau, October 2016, France.

Ludovic Perret was invited speaker at 17th World Conference on Information Security Applications (WISA 2016, August, Korea).

Elias Tsigaridas was invited speaker at

- the Department Seminar Series, of the Computer Science Department of the University of Liverpool, Apr 2016, UK.
- the Seminar of RICAM, University of Linz, Austria (Dec. 2016)

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Jérémy Berthomieu had the following teaching activities:

Master : Modeling and problems numerical and symbolic solving through MAPLE and MATLAB software, 34 hours, M1, Université Pierre-et-Marie-Curie, France

Master : In charge of Basics of Algebraic Algorithms, 70 hours, M1, Université Pierre-et-Marie-Curie, France

Master : Introduction to Security, 20 hours, M1, Université Pierre-et-Marie-Curie, France

Master : Projects supervision, 8 hours, L2, Université Pierre-et-Marie-Curie, France

Licence : Introduction to Algorithmics, 49 hours, L3, Université Pierre-et-Marie-Curie, France

Licence : Representations and Numerical Methods, 41 hours, L2, Université Pierre-et-Marie-Curie, France

Licence : Projects supervision, 10 hours, L2, Université Pierre-et-Marie-Curie, France

Jean-Charles Faugère had the following teaching activities:

Master: Fundamental Algorithms in Real Algebraic Geometry, 13,5 hours, M2, ENS de Lyon, France

Master : Polynomial Systems solving, 12 hours, M2, MPRI

Ludovic Perret had the following teaching activities amounting to around 220 hours:

Master : Polynomial Systems solving, M2, MPRI

Master : In charge of Introduction to Security, M1, Université Pierre-et-Marie-Curie, France

Master : In charge of Complexity, M1, Université Pierre-et-Marie-Curie, France

Licence : Introduction to Algorithmic, L2, Université Pierre-et-Marie-Curie, France

Licence : In charge of the Computer Science – Applied Mathematics Program (PIMA) in Licence, L2, Université Pierre-et-Marie-Curie, France

Licence : Project supervision, L2, Université Pierre-et-Marie-Curie, France

Guénaël Renault had the following teaching activities:

Master : In charge of the Security, Reliability and Numerical Efficiency Program in Master, 45 hours, M1 and M2, Université Pierre-et-Marie-Curie, France

Master : In charge of Advanced and Applied Cryptology, 70 hours, M2, Université Pierre-et-Marie-Curie, France

Master : In charge of Security and Side-channels, 10 hours, M2, Université Pierre-et-Marie-Curie, France

Master : In charge of Threats and Attacks Modeling, 40 hours, M1, Université Pierre-et-Marie-Curie, France

Master : Pro/Research internships supervision, 40 hours, M2, Université Pierre-et-Marie-Curie, France

Master : Projects supervision, 20 hours, M1, Université Pierre-et-Marie-Curie, France

Licence : In charge of Introduction to Cryptology, 30 hours, L3, Université Pierre-et-Marie-Curie, France

Licence : Project supervision, 10 hours, L2, Université Pierre-et-Marie-Curie, France

Mohab Safey El Din had the following teaching activities:

Master : In charge of Modeling and problems numerical and symbolic solving through MAPLE and MATLAB software, 36 hours, M1, Université Pierre-et-Marie-Curie, France

Master : In charge of Introduction to polynomial system solving, 48 hours, M2, Université Pierre-et-Marie-Curie, France

Master: In charge of Fundamental Algorithms in Real Algebraic Geometry, 22,5 hours, M2, ENS de Lyon, France

Master : In charge of the Security, Reliability and Numerical Efficiency Program in Master, 12 hours, M1 and M2, Université Pierre-et-Marie-Curie, France

Master : Introduction to Security, 10 hours, M1, Université Pierre-et-Marie-Curie, France

Licence : Introduction to Cryptology, 20 hours, L3, Université Pierre-et-Marie-Curie, France

Licence : In charge of the Computer Science – Applied Mathematics Program (PIMA) in Licence, L2 and L3, Université Pierre-et-Marie-Curie, France

9.2.2. Supervision

PhD in progress : Ivan Bannwarth, Fast algorithms for studying real algebraic sets, started in Sept. 2014, Mohab Safey El Din

PhD in progress : Matías Bender, Algorithms for Sparse Gröbner basis and applications, started in Dec. 2015, Jean-Charles Faugère and Elias Tsigaridas

PhD in progress : Eleonora Cagli, Analysis and interest points research in the attacks by observation context, Emmanuel Prouff and Cécile Dumas

PhD in progress : Clayton Eduardo Lente da Silva, Planar discontinuous dynamical system, Universidade Estadual Paulista (São José do Rio Preto), started in Sep. 2013, Paulo Ricardo da Silva and Alain Jacquemard

HdR : Ludovic Perret, Université Pierre-et-Marie-Curie, defended in Dec. 2016

HdR : Guénaël Renault, Université Pierre-et-Marie-Curie, defended in Dec. 2016

PhD : Thársis Souza Silva, Relay Systems, Universidade Federal de Goiás, Goiânia, defended in May 2016, Ronaldo Alves Garcia and Alain Jacquemard

PhD : Adrian Thillard, Countermeasures to Side-Channel Attacks and Secure- Multi-Party Computation, ENS Paris, defended in Dec. 2016 Damien Vergnaud and Emmanuel Prouff

PhD : Thibaut Verron, Gröbner bases and structured polynomial systems, Université Pierre-et-Marie-Curie, defended in Sept. 2016, Jean-Charles Faugère and Mohab Safey El Din

PhD : Alexandre Wallet, The point decomposition problem in Jacobian varieties, Université Pierre-et-Marie-Curie, defended in Dec. 2016, Jean-Charles Faugère

9.2.3. Juries

Jean-Charles Faugère was examiner in the PhD committees of C. Chenavier, V. Neiger, T. Verron and A. Wallet and in the HDR committees of L. Perret and G. Renault.

Alain Jacquemard was examiner in the PhD committee of T.S. Silva.

Emmanuel Prouff was reviewer of the PhD theses of A. Battistello and D. Martin. He was examiner in the PhD committee of A. Battistello, D. Martin and A. Thillard and in the HDR committees of G. Renault.

Mohab Safey El Din was examiner in the PhD committees of T. Verron and A. Wallet and in the HDR committees of L. Perret and G. Renault.

9.3. Popularization

J.-C. Faugère and L. Perret wrote a paper “Le grand défi du post-quantique” for MISC (HS 13, April 2016).

10. Bibliography

Major publications by the team in recent years

[1] M. R. BENDER, J.-C. FAUGÈRE, L. PERRET, E. TSIGARIDAS.*A Superfast Randomized Algorithm to Decompose Binary Forms*, in "ISSAC '16 - 41st International Symposium on Symbolic and Algebraic Computation", Waterloo, Canada, ACM, July 2016, p. 79-86 [*DOI :* 10.1145/2930889.2930896], https://hal.inria.fr/hal-01363545.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[2] L. PERRET.*Gröbner bases techniques in Quantum-Safe Cryptography*, UPMC - Paris 6 Sorbonne Universités, December 2016, Habilitation à diriger des recherches, https://tel.archives-ouvertes.fr/tel-01417808.

[3] G. RENAULT.*Contribution à la Résolution Algébrique et Applications en Cryptologie*, UPMC - Paris 6 Sorbonne Universités, December 2016, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01416242.

[4] T. VERRON.*Regularisation of Gröbner basis computations for weighted and determinantal systems, and an application to medical imagery*, Université Pierre et Marie Curie, September 2016, https://tel.archives-ouvertes.fr/tel-01404406.

[5] A. WALLET.*The point decomposition problem in Jacobian varieties* , Université Pierre & Marie Curie - Paris 6, November 2016, https://tel.archives-ouvertes.fr/tel-01407675.

Articles in International Peer-Reviewed Journal

[6] J. BERTHOMIEU, B. BOYER, J.-C. FAUGÈRE.*Linear Algebra for Computing Gröbner Bases of Linear Recursive Multidimensional Sequences*, in "Journal of Symbolic Computation", 2016, 48 [*DOI :* 10.1016/J.JSC.2016.11.005], https://hal.inria.fr/hal-01253934.

[7] K. CHAKRABORTY, S. SARKAR, S. MAITRA, B. MAZUMDAR, D. MUKHOPADHYAY, E. PROUFF.*Redefining the transparency order*, in "Designs, Codes and Cryptography", 2016 [*DOI :* 10.1007/S10623-016-0250-3], https://hal.archives-ouvertes.fr/hal-01399584.

[8] C. EDER, J.-C. FAUGÈRE.*A survey on signature-based algorithms for computing Gröbner basis computations*, in "Journal of Symbolic Computation", 2016, p. 1-75 [*DOI :* 10.1016/J.JSC.2016.07.031], https://hal.inria.fr/hal-00974810.

[9] J.-C. FAUGÈRE, C. MOU.*Sparse FGLM algorithms*, in "Journal of Symbolic Computation", May 2017, vol. 80, n^{o} 3, p. 538 - 569 [*DOI :* 10.1016/J.JSC.2016.07.025], https://hal.inria.fr/hal-00807540.

[10] J.-C. FAUGÈRE, A. OTMANI, L. PERRET, F. DE PORTZAMPARC, J.-P. TILLICH.*Folding Alternant and Goppa Codes with Non-Trivial Automorphism Groups*, in "IEEE Transactions on Information Theory", 2016, vol. 62, n^{o} 1, p. 184 - 198 [*DOI :* 10.1109/TIT.2015.2493539], https://hal.inria.fr/hal-01244609.

[11] J.-C. FAUGÈRE, A. OTMANI, L. PERRET, F. DE PORTZAMPARC, J.-P. TILLICH.*Structural Cryptanalysis of McEliece Schemes with Compact Keys*, in "Designs, Codes and Cryptography", April 2016, vol. 79, n^{o} 1, p. 87-112 [*DOI :* 10.1007/S10623-015-0036-Z], https://hal.inria.fr/hal-00964265.

[12] J.-C. FAUGÈRE, M. SAFEY EL DIN, T. VERRON.*On the complexity of computing Gröbner bases for weighted homogeneous systems*, in "Journal of Symbolic Computation", 2016 [*DOI :* 10.1016/J.JSC.2015.12.001], https://hal.inria.fr/hal-01097316.

[13] D. HENRION, S. NALDI, M. SAFEY EL DIN.*Exact algorithms for linear matrix inequalities*, in "SIAM Journal on Optimization", September 2016, vol. 26, n^{o} 4, p. 2512–2539 [*DOI :* 10.1137/15M1036543], https://hal.archives-ouvertes.fr/hal-01184320.

[14] D. HENRION, S. NALDI, M. SAFEY EL DIN.*Real root finding for determinants of linear matrices*, in "Journal of Symbolic Computation", May 2016, vol. 74, p. 205-238 [*DOI :* 10.1016/J.JSC.2015.06.010], https://hal.archives-ouvertes.fr/hal-01077888.

[15] M. SAFEY EL DIN, E. SCHOST.*A nearly optimal algorithm for deciding connectivity queries in smooth and bounded real algebraic sets*, in "Journal of the ACM", 2016, Major revision, accepted for publication to Journal of the ACM, https://hal.inria.fr/hal-00849057.

[16] A. STRZEBONSKI, E. TSIGARIDAS.*Univariate real root isolation over a single logarithmic extension of real algebraic numbers*, in "Springer Proceedings in Mathematics & Statistics", 2017, Ilias S. Kotsireas and Edgar Marti´nez-Moro, https://hal.inria.fr/hal-01001820.

[17] E. TSIGARIDAS.*SLV: a software for real root isolation*, in "ACM Communications in Computer Algebra", November 2016, vol. 50, n^{o} 3, p. 117 - 120 [*DOI :* 10.1145/3015306.3015317], https://hal.inria.fr/hal-01422209.

[18] D. WANG.*On the Connection Between Ritt Characteristic Sets and Buchberger–Gröbner Bases*, in "Mathematics in Computer Science", 2016, vol. 10, n^{o} 4, p. 479–492 [*DOI :* 10.1007/S11786-016-0279-8], https://hal.inria.fr/hal-01399579.

International Conferences with Proceedings

[19] L. BARTHELEMY, N. EYROLLES, G. RENAULT, R. ROBLIN.*Binary Permutation Polynomial Inversion and Application to Obfuscation Techniques*, in "2nd International Workshop on Software PROtection", Vienna, Austria, ACM, October 2016 [*DOI :* 10.1145/2995306.2995310], https://hal.inria.fr/hal-01388108.

[20] A. BATTISTELLO, J.-S. CORON, E. PROUFF, R. ZEITOUN.*Horizontal Side-Channel Attacks and Countermeasures on the ISW Masking Scheme*, in "18th Conference on Cryptographic Hardware and Embedded Systems (CHES 2016)", Santa Barbara, CA, United States, Cryptographic Hardware and Embedded Systems – CHES 2016, Springer, August 2016, vol. 9813, p. 23 - 39 [*DOI :* 10.1007/978-3-662-53140-2_2], https://hal.archives-ouvertes.fr/hal-01399577.

[21] S. BELAID, F. BENHAMOUDA, A. PASSELÈGUE, E. PROUFF, A. THILLARD, D. VERGNAUD.*Randomness Complexity of Private Circuits for Multiplication*, in "EUROCRYPT 2016", Vienna, Austria, May 2016, p. 616-648 [*DOI :* 10.1007/978-3-662-49896-5_22], https://hal.archives-ouvertes.fr/hal-01324823.

[22] *Best Paper*
M. R. BENDER, J.-C. FAUGÈRE, L. PERRET, E. TSIGARIDAS.*A Superfast Randomized Algorithm to Decompose Binary Forms*, in "ISSAC '16 - 41st International Symposium on Symbolic and Algebraic Computation", Waterloo, Canada, ACM, July 2016, p. 79-86 [*DOI :* 10.1145/2930889.2930896], https://hal.inria.fr/hal-01363545.

[23] J. BERTHOMIEU, J.-C. FAUGÈRE.*Guessing Linear Recurrence Relations of Sequence Tuples and P-recursive Sequences with Linear Algebra*, in "41st International Symposium on Symbolic and Algebraic Computation", Waterloo, ON, Canada, July 2016, p. 95-102 [*DOI :* 10.1145/2930889.2930926], https://hal.inria.fr/hal-01314266.

[24] B. BONNARD, J.-C. FAUGÈRE, A. JACQUEMARD, M. SAFEY EL DIN, T. VERRON.*Determinantal sets, singularities and application to optimal control in medical imagery*, in "International symposium on symbolic and algebraic computations", Waterloo, Canada, ACM, July 2016, p. 103-110 [*DOI :* 10.1145/2930889.2930916], https://hal.inria.fr/hal-01307073.

[25] B. BONNARD, A. JACQUEMARD, J. ROUOT.*Optimal Control of an Ensemble of Bloch Equations with Applications in MRI*, in "55th IEEE Conference on Decision and Control - CDC", Las Vegas, United States, December 2016, https://hal.inria.fr/hal-01287290.

[26] B. BOYER, C. EDER, J.-C. FAUGÈRE, S. LACHARTRE, F. MARTANI.*GBLA – Gröbner Basis Linear Algebra Package*, in "41st International Symposium on Symbolic and Algebraic Computation", Waterloo, ON, Canada, July 2016, p. 135-142 [*DOI :* 10.1145/2930889.2930914], https://hal.inria.fr/hal-01276346.

[27] J.-S. CORON, J.-C. FAUGÈRE, G. RENAULT, R. ZEITOUN.*Factoring $N = p^r q^s$ for Large r and s*, in "RSA Conference Cryptographers' Track", San Francisco, United States, Topics in Cryptology – CT-RSA 2016, February 2016 [*DOI :* 10.1007/978-3-319-29485-8_26], https://hal.inria.fr/hal-01250302.

[28] J.-S. CORON, A. GREUET, E. PROUFF, R. ZEITOUN.*Faster Evaluation of SBoxes via Common Shares*, in "18th International Conference on Cryptographic Hardware and Embedded Systems (CHES 2016)", Santa Barbara, CA, United States, Cryptographic Hardware and Embedded Systems – CHES 2016, Springer, August 2016, vol. 9813, p. 498 - 514 [*DOI :* 10.1007/978-3-662-53140-2_24], https://hal.archives-ouvertes.fr/hal-01399578.

[29] I. Z. EMIRIS, A. MANTZAFLARIS, E. TSIGARIDAS.*On the Bit Complexity of Solving Bilinear Polynomial Systems*, in "ISSAC '16 - 41st International Symposium on Symbolic and Algebraic Computation", Waterloo, Canada, ACM, July 2016, p. 215-222 [*DOI :* 10.1145/2930889.2930919], https://hal.inria.fr/hal-01401134.

[30] J.-C. FAUGÈRE, P.-J. SPAENLEHAUER, J. SVARTZ.*Computing Small Certificates of Inconsistency of Quadratic Fewnomial Systems*, in "International Symposium on Symbolic and Algebraic Computation (ISSAC 2016)", Waterloo, Canada, ACM, July 2016, p. 223-230 [*DOI :* 10.1145/2930889.2930927], https://hal.inria.fr/hal-01314651.

[31] M. SAFEY EL DIN, P.-J. SPAENLEHAUER.*Critical Point Computations on Smooth Varieties: Degree and Complexity bounds*, in "International Symposium on Symbolic and Algebraic Computation (ISSAC)", Waterloo, Canada, July 2016, p. 183–190 [*DOI :* 10.1145/2930889.2930929], https://hal.inria.fr/hal-01312750.

Scientific Books (or Scientific Book chapters)

[32] C. CARLET, E. PROUFF.*Polynomial Evaluation and Side Channel Analysis*, in "The New Codebreakers", Lecture Notes in Computer Science, Springer, 2016, vol. 9100, p. 315 - 341 [*DOI :* 10.1007/978-3-662-49301-4_20], https://hal.archives-ouvertes.fr/hal-01399573.

Other Publications

[33] G. RENAULT, T. VACCON.*On the p-adic stability of the FGLM algorithm*, 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01266071.

[34] M. SAFEY EL DIN, E. SCHOST.*Bit complexity for multi-homogeneous polynomial system solving Application to polynomial minimization*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01319729.

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:

1.1. - Architectures
1.1.8. - Security of architectures
1.2. - Networks
1.2.8. - Network security
1.3. - Distributed Systems
2. - Software
2.1. - Programming Languages
2.1.1. - Semantics of programming languages
2.1.3. - Functional programming
2.1.7. - Distributed programming
2.1.11. - Proof languages
2.2. - Compilation
2.2.1. - Static analysis
2.2.3. - Run-time systems
2.4. - Verification, reliability, certification
2.4.2. - Model-checking
2.4.3. - Proofs
2.5. - Software engineering
4. - Security and privacy
4.3. - Cryptography
4.3.3. - Cryptographic protocols
4.5. - Formal methods for security
4.6. - Authentication
4.8. - Privacy-enhancing technologies

Other Research Topics and Application Domains:

6. - IT and telecom
6.1. - Software industry
6.1.1. - Software engineering
6.3. - Network functions
6.3.1. - Web
6.3.2. - Network protocols
6.4. - Internet of things
9. - Society and Knowledge
9.8. - Privacy

1. Members

Research Scientists

Karthikeyan Bhargavan [Team leader, Inria, Senior Researcher, HDR]
Bruno Blanchet [Inria, Senior Researcher, HDR]
Harry Halpin [Inria, Starting Research position]
Catalin Hritcu [Inria, Researcher]
David Baelde [ENS Cachan, Délégation à Inria, Researcher]

Technical Staff

Gergely Bana [Inria, granted by FP7 ERC CIRCUS project]
Natalia Kulatova [Inria, from May 2016]
Marc Sylvestre [Inria, from Oct 2016]

PhD Students

Benjamin Beurdouche [Inria, granted by FP7 ERC CIRCUS project]
Yannis Juglaret [Inria, until Sep 2016]
Nadim Kobeissi [Inria]
Jean Karim Zinzindohoué [Min. Ecologie]

Visiting Scientist

Jonathan Protzenko [Microsoft Research, Oct 2016]

Administrative Assistant

Anna Bednarik [Inria]

Others

Alejandro Aguirre [Inria, from Apr 2016 until Aug 2016]
Abhishek Bichhawat [Inria, from Sep 2016 until Dec 2016]
Diane Gallois-Wong [ENS Paris, until Aug 2016]
Ritobroto Maitra [Inria, from May 2016 until Aug 2016]
Guido Martinez [Inria, until Jun 2016]
Jianyang Pan [Inria, from May 2016 until Aug 2016]
Marina Polubelova [Inria, from Sep 2016 until Nov 2016]
Vinay Yogendra [Inria, from May 2016 until Jul 2016]

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 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 manual proofs. Programs written in F* can be translated to OCaml, F#, or C for execution.

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 security type systems F7 and F* for the programming language F#. 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 computationally sound type system F* for applications written in F#. 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 compilers for realistic programming languages, both low-level (the C language) and high-level (ML and F*, a dependently-typed variant). These compilers 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 compilers will 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.

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 2014) 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, the 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
- F7: a security typechecker for cryptographic applications written in F#

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 [42]; it has motivated a serious research effort on the formal analysis of cryptographic protocols, starting with [40] 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* [36]. ProVerif also distinguishes itself from other tools by the variety of cryptographic primitives it can handle, defined by rewrite rules or by some equations, and the variety of security properties it can prove: secrecy [34], [25], correspondences (including authentication) [35], and observational equivalences [33]. 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 is the only tool that proves equivalences for an unbounded number of sessions.

Using ProVerif, it is now possible to verify large parts of industrial-strength protocols, such as TLS [30], JFK [26], and Web Services Security [32], 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 30 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 [37], results which were extended by formal analysis work using a Dolev-Yao style model of the standard [38]. 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 F7 and 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, that models typically ignore. This leads to a situation that a protocol may have been proved secure in theory, but 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# [30]. 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 such as typechecking. F7 [28] is a refinement typechecker for F#, developed jointly at Microsoft Research Cambridge and Inria. It implements a dependent type-system that allows us to specify security assumptions and goals as first-order logic annotations directly inside the program. It has been used for the modular verification of large web services security protocol implementations [31]. F* (see below) is an extension of F7 with higher-order kinds and a certifying typechecker. Both F7 and F* have a growing user community. The cryptographic protocol implementations verified using F7 and F* already represent the largest verified cryptographic applications to our knowledge.

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 already designed the automatic tool CryptoVerif, which generates proofs by sequences of games. Much work is still needed in order to develop this approach, so that it is applicable to more protocols. We also plan to design and implement techniques for proving implementations of protocols secure in the computational model, by generating them from CryptoVerif specifications that have been proved secure, or by automatically extracting CryptoVerif models from implementations.

A different approach is to directly verify cryptographic applications in the computational model by typing. A recent work [41] 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 Designed for Program Verification

F* [43] is a verification system for ML programs developed collaboratively by Inria and Microsoft Research. ML types are extended with logical predicates that can conveniently express precise specifications for programs (pre- and post- conditions of functions as well as stateful invariants), including functional correctness and security properties. The F* typechecker implements a weakest-precondition calculus to produce first-order logic formulas that are automatically discharged using the Z3 SMT solver. The original F* implementation has been successfully used to verify nearly 50,000 lines of code, including cryptographic protocol implementations, web browser extensions, cloudhosted web applications, and key parts of the F* typechecker and compiler (itself written in F*). F* has also been used for formalizing the semantics of other languages, including JavaScript and a compiler from a subset of F* to JavaScript, and TS*, a secure subset of TypeScript. Programs verified with F* can be extracted to F#, OCaml, C, and JavaScript and then efficiently executed and integrated into larger code bases.

The latest version of F* is written entirely in F*, and bootstraps in OCaml and F#. It is open source and under active development on GitHub. A detailed description of this new F* version is available in a POPL 2016 paper [20] and a POPL 2017 one [6]. We continue to evolve and develop F* and we use it to develop large case studies of verified cryptographic applications, such as miTLS.

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 compilers for realistic programming languages, both low-level (the C language) and high-level (ML and F*, a dependently-typed variant). These compilers 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 compilers will 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 experimentally evaluate and carefully optimize the efficiency of our secure compilers on realistic workloads and standard benchmark suites. We will use property-based testing and formal verification to provide high confidence that our compilers are indeed secure. Formally, we will construct machine-checked proofs of full abstraction with respect to a secure high-level semantics. This strong property complements compiler correctness and ensures that no machine-code attacker can do more harm to securely compiled components than a component in the secure source language already could.

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 [29] [29] and TS* [45], 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 the 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.

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

This year, we published 18 articles in international peer-reviewed journals and conferences, including papers in prestigious conferences such as POPL, IEEE S&P Oakland, ACM CCS, NDSS, CSF, and WPES. Notably, Bruno Blanchet published a book surveying the use of ProVerif, his state-of-the-art protocol verification tool. We also won several research awards for our work, detailed below. Our work also exposed two new attacks, SLOTH and SWEET32, on Transport Layer Security, resulting in security updates and CVEs in popular web browsers and VPN software.

5.1.1. Awards

- Catalin Hritcu was awarded an ERC Starting Grant
- Catalin Hritcu was awarded an ANR Jeune Chercheur/Jeune Chercheuse Grant
- Karthikeyan Bhargavan was awarded an ERC Consolidator Grant
- Karthikeyan Bhargavan and Gaëtan Leurent won a Best Paper award at NDSS 2016
- Karthikeyan Bhargavan, Cedric Fournet, Markulf Kohlweiss, and Alfredo Pironti were awarded the Levchin prize for contributions to Real-World Cryptography
- Karthikeyan Bhargavan was awarded a Microsoft Outstanding Collaborator Award
- Karthikeyan Bhargavan was awarded the Prix Inria – Académie des sciences du Jeune chercheur

6. New Software and Platforms

6.1. ProVerif

Participants: Bruno Blanchet [correspondant], Xavier Allamigeon [April–July 2004], Vincent Cheval [Sept. 2011–Dec. 2014], Benjamin Smyth [Sept. 2009–Feb. 2010], Marc Sylvestre [Oct. 2016–].

PROVERIF (http://proverif.inria.fr) 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 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.

The PROVERIF verifier can prove the following properties:

- secrecy (the adversary cannot obtain the secret);
- authentication and more generally correspondence properties, of the form "if an event has been executed, then other events have been executed as well";
- strong secrecy (the adversary does not see the difference when the value of the secret changes);
- equivalences between processes that differ only by terms.

PROVERIF is widely used by the research community on the verification of security protocols (see http://proverif.inria.fr/proverif-users.html for references).

PROVERIF is freely available on the web, at http://proverif.inria.fr, under the GPL license.

6.2. CryptoVerif

Participants: Bruno Blanchet [correspondant], David Cadé [Sept. 2009–Dec. 2013].

CRYPTOVERIF(http://cryptoverif.inria.fr) 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.

The generated proofs are proofs by sequences of games, as used by cryptographers. These proofs are valid for a number of sessions polynomial in the security parameter, in the presence of an active adversary. CRYPTOVERIF can also evaluate the probability of success of an attack against the protocol as a function of the probability of breaking each cryptographic primitive and of the number of sessions (exact security).

CryptoVerif has been used in particular for a study of Kerberos in the computational model, and as a back-end for verifying implementations of protocols in F# and C.

CryptoVerif is freely available on the web, at http://cryptoverif.inria.fr, under the CeCILL license.

6.3. miTLS

Participants: Karthikeyan Bhargavan [correspondant], Cedric Fournet [Microsoft Research], Markulf Kohlweiss [Microsoft Research], Antoine Delignat-Lavaud [Microsoft Research], Nikhil Swamy [Microsoft Research], Santiago Zanella-Béguelin [Microsoft Research], Jean Karim Zinzindohoué, Benjamin Beurdouche, Alfredo Pironti.

miTLS is a verified reference implementation of the TLS security protocol in F#, a dialect of OCaml for the .NET platform. It supports SSL version 3.0 and TLS versions 1.0-1.2 and interoperates with mainstream web browsers and servers. miTLS has been verified for functional correctness and cryptographic security using the refinement typechecker F7.

Papers describing the miTLS library was published at IEEE S&P 2013, CRYPTO 2014, and IEEE S&P Journal 2016. miTLS is now being developed on GitHub with dozens of contributors and regular updates. The miTLS team was awarded the Levchin prize for contributions to Real-World Cryptography in 2016. The software and associated research materials are available from http://mitls.org.

6.4. F*

Participants: Alejandro Aguirre, Danel Ahman [University of Edinburgh], Benjamin Beurdouche, Karthikeyan Bhargavan, Antoine Delignat-Lavaud [Microsoft Research], Cédric Fournet [Microsoft Research], Catalin Hritcu, Chantal Keller [Université Paris-Sud], Kenji Maillard, Guido Martínez, Gordon Plotkin, Samin Ishtiaq [Microsoft Research], Markulf Kohlweiss [Microsoft Research], Jonathan Protzenko [Microsoft Research], Tahina Ramananandro [Microsoft Research], Aseem Rastogi [Microsoft Research], Nikhil Swamy [Microsoft Research], Peng Wang [MIT], Santiago Zanella-Béguelin [Microsoft Research], Jean Karim Zinzindohoué.

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 Fω (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.

A detailed description of F* (circa 2011) appeared in the Journal of Functional Programming [44]. F* has evolved substantially since then. The latest version of F* is written entirely in F*, and bootstraps in OCaml and F#. It is under active development at GitHub: https://github.com/FStarLang and the official webpage is at http://fstar-lang.org.

6.5. HACL*

Participants: Karthikeyan Bhargavan, Jean Karim Zinzindohoué, Marina Polubelova, Benjamin Beurdouche, Jonathan Protzenko [Microsoft Research].

HACL* is a verified cryptographic library written in F*. It implements modern primitives, including elliptic curves like Curve25519, symmetric ciphers like Chacha20, and MAC algorithms like Poly1305. These primitives are then composed into higher-level constructions like Authenticated Encryption with Additional Data (AEAD) and the NaCl API. All the code in HACL* is verified for memory safety, side channel resistance, and where applicable, also for functional correctness and absence of integer overflow. HACL* code is used as the basis for cryptographic proofs for security in the miTLS project.

HACL* code can be compiled to OCaml using the standard F* compiler, or to C, using the Kremlin backend of F*. The generated C code is as fast as state-of-the-art cryptographic libraries written in C. HACL* is being actively developed on Github; see https://github.com/mitls/hacl-star

6.6. ProScript

Participants: Nadim Kobeissi [correspondant], Karthikeyan Bhargavan, Bruno Blanchet.

Defensive JavaScript (DJS) is a subset of the JavaScript language that guarantees the behaviour of trusted scripts when loaded in an untrusted web page. Code in this subset runs independently of the rest of the JavaScript environment. When propertly wrapped, DJS code can run safely on untrusted pages and keep secrets such as decryption keys. ProScript is a typed subset of JavaScript, inspired by DJS, that is focused on writing verifiable cryptographic protocol implementations. In addition to DJS typing, ProScript imposes a functional style that results in more readable and easily verifiable ProVerif models. ProScript has been used to write and verify a full implementation of the Signal and TLS 1.3 protocols in JavaScript.

The ProScript compiler and various libraries written in ProScript are being developed on Github and will be publicly released in 2017.

6.7. QuickChick

Participants: Maxime Dénès [Inria Sophia-Antipolis], Catalin Hritcu, John Hughes [Chalmers University], Leonidas Lampropoulos [University of Pennsylvania], Zoe Paraskevopoulou [Princeton University], Benjamin Pierce [University of Pennsylvania].

QuickChick is a verified plugin that integrates property-based testing and proving in the Coq proof assistant. This integration is aimed at reducing the cost of formal verification and at providing stronger, formal foundations to property-based testing. https://github.com/QuickChick/QuickChick

6.8. Luck

Participants: Leonidas Lampropoulos [University of Pennsylvania], Diane Gallois-Wong [ENS and Inria Paris], Catalin Hritcu, John Hughes [Chalmers University], Benjamin Pierce [University of Pennsylvania], Li-Yao Xia [ENS and Inria Paris].

Property-based random testing a la QuickCheck requires building efficient generators for well-distributed random data satisfying complex logical predicates, but writing these generators can be difficult and error prone. We propose a domain-specific language in which generators are conveniently expressed by decorating predicates with lightweight annotations to control both the distribution of generated values and the amount of constraint solving that happens before each variable is instantiated. This language, called Luck, makes generators easier to write, read, and maintain. https://github.com/QuickChick/Luck

6.9. Privacy-preserving federated identity

Participants: Harry Halpin, George Danezis [University College London].

security protocols, secure messaging, decentralization, blockchain

Working with the partners in the NEXTLEAP project, we helped design a protocol for decentralized and privacy-preserving identity and key management, creating both a fix privacy vulnerabilities in OAuth (UnlimitID) and on generic versions of blockchain for usages such as key management (ClaimChain). While this software has been prototyped in 2016 using Python working with NEXTLEAP project partner University College London (George Danezis), we expected either formal analysis using ProVerif or verified implementations using Fstar in 2017, as well and integrate this work with formal verification work done in Prosecco on end-to-end secure messaging.

7. New Results

7.1. Verification of Security Protocols in the Symbolic Model

Participants: Bruno Blanchet, Marc Sylvestre.

security protocols, symbolic model, automatic verification The applied pi calculus is a widely used language for modeling security protocols, including as a theoretical basis of PROVERIF. However, the seminal paper that describes this language [27] does not come with proofs, and detailed proofs for the results in this paper were never published. Martín Abadi, Bruno Blanchet, and Cédric Fournet wrote detailed proofs of all results of this paper. This work appears as a research report [21] and is submitted to a journal.

Stéphanie Delaune, Mark Ryan, and Ben Smyth [39] introduced the idea of swapping data in order to prove observational equivalence. For instance, ballot secrecy in electronic voting is formalized by saying that A voting a and B voting b is observationally equivalent to (indistinguishable from) A voting b and B voting a. Proving such an equivalence typically requires swapping the votes. However, Delaune et al's approach was never proved correct. Bruno Blanchet and Ben Smyth filled this gap by formalizing the approach and providing a detailed soundness proof [12], [23]. This extension is implemented in ProVerif. Moreover, Marc Sylvestre implemented a graphical display of attacks in ProVerif. The extended tool is available at http://proverif.inria.fr.

Bruno Blanchet wrote a survey on ProVerif, available both as a book and as a journal paper [3].

7.2. Verification of Security Protocols in the Computational model

Participant: Bruno Blanchet.

security protocols, computational model, verification Bruno Blanchet implemented extensions of his computational protocol verifier CryptoVerif. In particular, the tool collects more precise information at each program point, in order to improve the simplification of cryptographic games and the proof of correspondence assertions (authentication). For instance, this extension allows one to prove injective correspondences for protocols with a replay cache. Another extension provides a query to show that several variables are independent secrets. The extended tool is available at http://cryptoverif.inria.fr.

7.3. Verification of Avionic Security Protocols

Participant: Bruno Blanchet.

security protocols, symbolic model, computational model, verification Within the ANR project AnaStaSec, Bruno Blanchet studied an air-ground avionic security protocol, the ARINC823 public key protocol [24]. He verified this protocol both in the symbolic model of cryptography, using ProVerif, and in the computational model, using CryptoVerif. While this study confirmed the main security properties of the protocol (entity and message authentication, secrecy), he found several weaknesses and imprecisions in the standard. He proposed fixes for these problems. He delivered this work to the ANR and he plans to submit it for publication next year.

7.4. The F* programming language

Participants: Alejandro Aguirre, Danel Ahman [University of Edinburgh], Benjamin Beurdouche, Karthikeyan Bhargavan, Antoine Delignat-Lavaud [Microsoft Research], Cédric Fournet [Microsoft Research], Catalin Hritcu, Chantal Keller [Université Paris-Sud], Kenji Maillard, Guido Martínez, Gordon Plotkin, Samin Ishtiaq [Microsoft Research], Markulf Kohlweiss [Microsoft Research], Jonathan Protzenko [Microsoft Research], Tahina Ramananandro [Microsoft Research], Aseem Rastogi [Microsoft Research], Nikhil Swamy [Microsoft Research], Peng Wang [MIT], Santiago Zanella-Béguelin [Microsoft Research], Jean Karim Zinzindohoué.

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 Fω (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.

We published a paper on the design, implementation, and formal core of F* at POPL 2016 [20]. A first significant improvement on this design will appear at POPL 2017 under the name of "Dijkstra Monads for Free" [6]. Also significant work was put into extracting a subset of F* to C; we submitted a paper on this to PLDI 2017. F* is being developed as an open-source project at GitHub: https://github.com/FStarLang and the official webpage is at http://fstar-lang.org. We released several beta versions of the software this year.

7.5. Dependable Property-Based Testing

Participants: Maxime Dénès [Inria Sophia-Antipolis], Diane Gallois-Wong [ENS and Inria Paris], Catalin Hritcu, John Hughes [Chalmers University], Leonidas Lampropoulos [University of Pennsylvania], Zoe Paraskevopoulou [Princeton University], Benjamin Pierce [University of Pennsylvania], Li-Yao Xia [ENS and Inria Paris].

This year we finally released the Luck programming language for property-based generators (https://github.com/QuickChick/Luck); a paper on this is about to appear at POPL 2017 [18]. We also improved a previous case study on testing information-flow control mechanisms and published a journal paper on this at JCS [1]. Finally, we kept improving the QuickChick testing plugin for Coq (https://github.com/QuickChick/QuickChick), in particular by automatically producing generators from algebraic datatype definitions.

7.6. Micro-Policies and Secure Compilation

Participants: Arthur Azevedo de Amorim [University of Pennsylvania], André Dehon [University of Pennsylvania and Draper Labs], Catalin Hritcu, Yannis Juglaret, Boris Eng, Benjamin Pierce [University of Pennsylvania], Howard Shrobe [MIT], Stelios Sidiroglou-Douskos [MIT], Greg Sullivan [Draper Labs], Andrew Tolmach [Portland State University].

This year we obtained a new ERC Starting Grant on secure compilation using micro-policies; the grant will start in January 2017. Our work was focused on laying the foundations for this long-term research direction. Preliminary work on this appeared at CSF 2016 [17]. In addition, an improved version of our paper on micro-policies for information flow-control appeared at JFP [1]. Finally, we were part of Draper Labs' patent application on "Techniques for Metadata Processing", as developed jointly in the micro-policies project.

7.7. miTLS: Proofs for TLS 1.3

Participants: Karthikeyan Bhargavan, Chris Brzuska [Technical University of Hamburg], Cedric Fournet [Microsoft Research], Matthew Green [Johns Hopkins University], Markulf Kohlweiss [Microsoft Research], Santiago Zanella-Béguelin [Microsoft Research], Jean Karim Zinzindohoué.

transport layer security, cryptographic protocol, verified implementation, man-in-the-middle attack, impersonation attack

We actively participated in the design of TLS 1.3, and worked on a verified implementation of TLS 1.0-1.3 in F*, called miTLS. miTLS is being actively developed on GitHub and we have submitted a paper on our verified implementation of the TLS 1.3 record layer. We published a paper on our overall verification methodology in the IEEE Security and Privacy journal.

Many recent attacks on TLS, discovered by us and others, have relied on *downgrading* a TLS connection and forcing it to use obsolete cryptographic constructions, even if the client and server support and prefer to use modern cryptography. We wrote a paper that showed that such downgrade weaknesses also exist in other protocols such as IPsec, SSH, and ZRTP. We formalized a notion of *downgrade resilience* and showed how it can be achieved in different circumstances. In particular we proved that a new downgrade protection mechanism in TLS 1.3, which was proposed by us, prevents a large class of downgrade attacks. This paper appeared in IEEE S&P (Oakland) 2016 [7].

7.8. Attacks on obsolete cryptography

Participants: Karthikeyan Bhargavan, Gaëtan Leurent.

transport layer security, cryptographic protocol, man-in-the-middle attack, impersonation attack

At NDSS 2016, we published a paper [10] describing a new class of attacks on the use of weak hash functions in popular key exchange protocols such as TLS, IKE, and SSH. One of these attacks, called SLOTH, demonstrated a practical attack on MD5-based client authentication in TLS. We responsibly disclosed this vulnerability, which resulted in security updates in various web browsers and servers. For example, SLOTH-related updates were released for Firefox, Java, RedHat Linux, and for all websites hosted by the Akamai content delivery network.

At CCS 2016, we published a paper [9] that described an attack, called Sweet32, that affects protocols that use block ciphers with short 64-bit blocks, such as Triple-DES and Blowfish. When more than a certain amount of data is sent using such ciphers, the attacker can exploit ciphertext collisions to reconstruct the secret plaintext. We showed how this vulnerability affects TLS and OpenVPN connections. Our findings led to security advisories for OpenVPN, OpenSSL, and all Apple products.

7.9. HACL*: Verified cryptographic library

Participants: Karthikeyan Bhargavan, Jean Karim Zinzindohoué, Marina Polubelova, Benjamin Beurdouche, Jonathan Protzenko [Microsoft Research].

HACL* is a verified cryptographic library written in F*. It implements modern primitives, including elliptic curves like Curve25519, symmetric ciphers like Chacha20, and MAC algorithms like Poly1305. These primitives are then composed into higher-level constructions like Authenticated Encryption with Additional Data (AEAD) and the NaCl API. All the code in HACL* is verified for memory safety, side channel resistance, and where applicable, also for functional correctness and absence of integer overflow. HACL* code is used as the basis for cryptographic proofs for security in the miTLS project.

In CSF 2016, we published a paper on a library of elliptic curves written in F* and compiled to OCaml. This library included the first verified implementations for multiple curves: Curve25519, Curve448, and NIST P-256. However, our code was not very fast. More recently, we worked on Kremlin, a compiler from F* to C that generates code which is as fast as state-of-the-art cryptographic libraries written in C. We have submitted a paper on the Kremlin compiler and its use in HACL*. All our code is being actively and openly developed on GitHub.

7.10. Design and Verification of next-generation protocols: identity, blockchains, and messaging

Participants: Harry Halpin, George Danezis [University College London], Carmela Troncoso [IMDEA].

We began work on designing substantial modifications to existing protocols, verifying pre-standard protocols, or creating entirely new standards for new areas. In these areas the fundamental protocols are often unstandardized and controlled by a few large companies (such as the case of identity-based authorization in terms of Google and Facebook's use of OAuth) and new protocols (such as the incompatible space of protocols around secure messaging given by applications such as WhatsApp, Signal, Telegram, and Viber). In some cases, these protocols do not support basic features needed for standardization, such as decentralization and federation. Therefore, in the first half of 2017, Harry Halpin worked with colleagues at IMDEA and University College London in completing the first systematization of knowledge of decentralization, submitted to PETS 2017, and presented preliminary results in "The Responsibility of Open Standards" paper at the HotPETS 2016 workshop as well as in the First Monday journal.

One of the most important protocols in the entire Web is the OAuth protocol, yet it has suffered from a number of dangerous security and privacy issues. Previously formally analysed by Prosecco, one of the larger problems facing this widely deployed protocol is the lack of privacy. Whenever a user log-ins into a website via Google or Facebook Connect (their identity provider), and then authorizes the flow of data between that website and the identity provider. However, the identity provider then gains knowledge of the every single visit that their users make to other websites that request their data, in addition to the data that the identity provider stores itself. Using a new blind signature scheme based on Algebraic MACs, the new UnlimitID protocol makes the use of federated identity by a user at a website unlinkable to their identity provider, while still allowing websites to gain the advantage of single authenticated sign-on to a large identity to prevent spamming and abuse. This work was presented at the Workshop for Privacy in the Electronic Society at ACM CCS. Unlike previous work that requires substantial changes to both websites using OAuth and identity providers, by using the new W3C Web Crypto API (as analyzed by Halpin), this new protocol requires only changes to the identity provider and is do backwards-compatible with existing OAuth implementations. Microsoft has supported this work for possible future standardization in the OpenID Foundation.

In order to be decentralized, secure messaging requires an ability to discover key material and guarantee its integrity. Typically, today this is done via a single centralized and unstandardised service provider. In order to create an interoperable standard around secure messaging, key discovery needs to be decentralized. Blockchain-based approaches have been suggested in previous work in the security research community such as CONIKS, but have failed to take off due to the high deployment cost on centralized servers. We've designed a new protocol, ClaimChain, that builds on both existing work on blockchains while adding new optimizations and providing a decentralized logic based on Rivest and Lampson's SDSI to identify and discovery key material without a trusted third party. Joint work with CNRS to understand the social and economic considerations led to a publication in Internet Science and the existing design will be submitted to a top-notch security conference. Currently, we are discussing early use of this design with codebases used by secure messaging and email providers, and a security and privacy analysis of these codbases was published in CANS. Over the next year we plan for all of these protocols to have formally verified code for their cryptographic functionality and to present a design on how to integrate this work on key discovery into secure messaging with improved privacy and transcript consistency.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

8.1.1.1. AnaStaSec

Title: Static Analysis for Security Properties (ANR générique 2014.)

Other partners: Inria/Antique, Inria/Celtique, Airbus Operations SAS, AMOSSYS, CEA-LIST, TrustInSoft

Duration: January 2015 - December 2018.

Coordinator: Jérôme Féret, Inria Antique (France)

Participant: Bruno Blanchet

Abstract: The project aims at using automated static analysis techniques for verifying security and confidentiality properties of critical avionics software.

8.1.1.2. 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)

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.

8.1.1.3. SafeTLS

Title: SafeTLS: La se´curisation 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, Univesité de Rennes 1 (France)

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.

8.1.1.4. QuickChick

Title: QuickChick: Property-based Testing for Coq

Coordinator: Catalin Hritcu

Abstract: The goal of the project was to develop a property-based testing framework for Coq proofs. Catalin Hritcu was awarded an ANR Jeune Chercheur/Jeune Chercheuse grant to pursue this project, but he declined it in favour of his ERC Starting Grant SECOMP (described below.)

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. 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 Bhargavn, Inria

Abstract: 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.

8.2.1.2. ERC Starting Grant: SECOMP

Title: SECOMP: Efficient Formally Secure Compilers to a Tagged Architecture

Duration: Jan 2017 - December 2021

Coordinator: Catalin Hritcu, Inria

Abstract: This new ERC-funded project called SECOMP1 is aimed at leveraging emerging hardware capabilities for fine-grained protection to build the first, efficient secure compilers for realistic programming languages, both low-level (the C language) and high-level (F*, a dependently-typed ML variant). These compilers 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 compilers will 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.

8.2.1.3. NEXTLEAP

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

Abstract: NEXTLEAP aims to create, validate, and deploy protocols that can serve as pillars for a secure, trust-worthy, and privacy-respecting Internet. For this purpose NEXTLEAP will develop an interdisciplinary study of decentralisation that provides the basis on which these protocols cann be designed, working with sociologists to understand user needs. The modular specification of decentralized protocols, implemented as verified open-source software modules, will be done for both privacy-preserving secure federated identity as well as decentralized secure messaging services that hide metadata (e.g., who, when, how often, etc.).

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners

We have a range of long- and short-term collaborations with various universities and research labs. We summarize them by project:

- **F***: Microsoft Research (Cambdridge, Redmond), IMDEA (Madrid)
- **TLS analysis**: Microsoft Research (Cambridge), Johns Hopkins University, University of Michigan, University of Pennsylvania
- **Web Security**: Microsoft Research (Cambridge, Redmond), Imperial College (London)
- **Micro-Policies**: University of Pennsylvania, Portland State University

8.4. International Research Visitors

8.4.1. Visits of International Scientists

- Carmela Troncoso from IMDEA visited the group from 17-18th October and gave a seminar “Traffic Analysis - When Encryption is not Enough to Protect Privacy”

8.4.1.1. Internships

- Alejandro Aguirre: Apr 2016 until Aug 2016
- Abhishek Bichhawat: Sep 2016 until Dec 2016
- Diane Gallois-Wong: Mar 2016 until Aug 2016
- Ritobroto Maitra: May 2016 until Aug 2016
- Guido Martinez: Jan 2016 until Jun 2016

- Jianyang Pan: May 2016 until Aug 2016
- Marina Polubelova: Sep 2016 until Nov 2016
- Natalia Kulatova: May 2016 until Aug 2016
- Vinay Yogendra: May 2016 until Jul 2016

8.4.2. Visits to International Teams

- Bruno Blanchet, March 14 to June 10, 2016, Google, Mountain View.
- Catalin Hritcu, October to November 2016, Microsoft Research, Redmond, USA.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- Prosecco is organizing the 2nd IEEE European Symposium on Security and Privacy in Paris, | 26-28 April 2017. Catalin Hritcu is General Chair, Bruno Blanchet is Finance Chair, and Karthikeyan Bhargavan is Local arrangements Chair.
- Catalin Hritcu organized two Secure Compilation Meetings (SCM) at Inria Paris in (13 invited participants, 17–19 August 2016) and POPL (15 January 2017)

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Catalin Hritcu is PC member at POPL 2017
- Catalin Hritcu is PC member at POST 2017
- Catalin Hritcu was PC member at CSF 2016
- Catalin Hritcu was PC member at POST 2016
- Catalin Hritcu was PC member at ITP 2016
- Catalin Hritcu was PC member at cPP 2016
- Harry Halpin is PC member for ACM WWW 2017
- Harry Halpin was PC member for W3C Blockchains and the Web
- Karthikeyan Bhargavan is ERC member at POPL 2017
- Karthikeyan Bhargavan is PC member at IEEE S&P 2017
- Karthikeyan Bhargavan was PC member at IEEE S&P 2016
- Karthikeyan Bhargavan was PC member at ACM CCS 2016
- Karthikeyan Bhargavan was PC member at IEEE CSF 2016
- Karthikeyan Bhargavan was PC member at ACM PLAS 2016

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Associate Editor

- of the *International Journal of Applied Cryptography (IJACT)* – Inderscience Publishers: Bruno Blanchet

9.1.4. Invited Talks

- Karthikeyan Bhargavan gave an invited talk at EUROCRYPT 2016
- Karthikeyan Bhargavan gave an invited talk at the OAuth Workshop 2016
- Karthikeyan Bhargavan gave an invited talk at SSTIC 2016

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Master: Bruno Blanchet, Formal Methods, 9h equivalent TD, master M2 MIC, université Paris VII, France
- Master: Bruno Blanchet, Cryptographic protocols: formal and computational proofs, 31.5h equivalent TD, master M2 MPRI, université Paris VII, France
- Master: Karthikeyan Bhargavan, Cryptographic protocols: formal and computational proofs, 31.5h equivalent TD, master M2 MPRI, université Paris VII, France
- Master: Karthikeyan Bhargavan, Protocol Safety and Security, master ACN Telecom ParisTech et Ecole Polytechnique
- Undergraduate: Karthikeyan Bhargavan, INF421 and INF431: Programmation, Ecole Polytechnique
- Doctorat: Karthikeyan Bhargavan: Protecting TLS from legacy cryptography, l'École de printemps en codage et cryptographie, May 2016
- Master: Catalin Hritcu, Cryptographic protocols: formal and computational proofs, 31.5h equivalent TD, master M2 MPRI, université Paris VII, France
- Doctorat: Catalin Hritcu: F* course at Computer Aided Analysis of Cryptographic Protocols summer school, Bucharest, September 2016

9.2.2. Supervision

- PhD completed: Antoine Delignat-Lavaud
  On the Security of Authentication Protocols for the Web
  defended March 2016, supervised by Karthikeyan Bhargavan
- PhD in progress: Evmorfia-Iro Bartzia
  Machine-checked program verification for concrete cryptography,
  defence on February 15, 2017, supervised by Karthikeyan Bhargavan and Pierre-Yves Strub
- PhD in progress: Jean Karim Zinzindohoué
  Analyzing cryptographic protocols and their implementations,
  started September 2014, supervised by Karthikeyan Bhargavan
- PhD in progress: Nadim Kobeissi
  Analyzing cryptographic web applications,
  started February 2015, supervised by Karthikeyan Bhargavan
- PhD incomplete: Yannis Juglaret
  Micro-policies and Secure Compilation,
  started September 2015, interrupted September 2016, supervised by Catalin Hritcu

9.2.3. Juries

- Karthikeyan Bhargavan served on the PhD jury of Olivier Levillain
- Harry Halpin served on the PhD jury of Nikita Mazurov

9.3. Popularization

9.3.1. Seminars

- Karthikeyan Bhargavan: invited talks at SSTIC, EUROCRYPT, OAuth Workshop
- Bruno Blanchet: invited talks at John Mitchell's 60th birthday workshop, Stanford University, CA (May 2016), Facebook, Menlo Park, CA (May 2016), and at University of Oslo (Dec 2016).
- Catalin Hritcu: invited talks at CEA List, MSR Redmond, Inria Gallium, Secure Compilation Meeting, ERC, Inria Prosecco, MPI-SWS
- Harry Halpin: invited talks at NetFutures 2016 (April 2016), Trust in the Digital World (June 2016), Strategic Research Challenges in Privacy-Enhancing Technologies (July 2016), European Dialogue on Internet Governance (September 2016), Internet Governance Forum Tunis (October 2016), Keynote at International Workshop on Semantic Web, and Cryptodesign (November 2016).

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] A. Azevedo De Amorim, N. Collins, A. DeHon, D. Demange, C. Hriţcu, D. Pichardie, B. C. Pierce, R. Pollack, A. Tolmach.*A Verified Information-Flow Architecture*, in "Journal of Computer Security (JCS); Special Issue on Verified Information Flow Security", December 2016, vol. 24, n° 6, p. 689–734 [*DOI :* 10.3233/JCS-15784], https://hal.archives-ouvertes.fr/hal-01424797.

[2] K. Bhargavan, C. Fournet, M. Kohlweiss.*miTLS: Verifying Protocol Implementations against Real-World Attacks*, in "IEEE Security & Privacy", December 2016, vol. 14, n° 6, p. 18-25 [*DOI :* 10.1109/MSP.2016.123], https://hal.inria.fr/hal-01425964.

[3] 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° 1-2, p. 1 - 135 [*DOI :* 10.1561/3300000004], https://hal.inria.fr/hal-01423760.

[4] H. Halpin, A. Monnin.*The Decentralization of Knowledge: How Carnap and Heidegger influenced the Web*, in "First Monday", December 2016, vol. 21, n° 12 [*DOI :* 10.5210/fm.v21i12.7109], https://hal.archives-ouvertes.fr/hal-01397931.

[5] C. Hriţcu, L. Lampropoulos, A. Spector-Zabusky, A. A. d. Amorim, M. Dénès, J. Hughes, B. C. Pierce, D. Vytiniotis.*Testing Noninterference, Quickly*, in "Journal of Functional Programming (JFP); Special issue for ICFP 2013", April 2016, vol. 26, 62, e4 [*DOI :* 10.1017/S0956796816000058], https://hal.archives-ouvertes.fr/hal-01424796.

International Conferences with Proceedings

[6] 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)", Paris, France, ACM, 2017, p. 515-529 [*DOI :* 10.1145/3009837.3009878], https://hal.archives-ouvertes.fr/hal-01424794.

[7] K. Bhargavan, C. Brzuska, C. Fournet, M. Green, M. Kohlweiss, S. Zanella-Béguelin.*Downgrade Resilience in Key-Exchange Protocols*, in "IEEE Symposium on Security and Privacy (SP), 2016", San Jose, United States, May 2016 [*DOI :* 10.1109/SP.2016.37], https://hal.inria.fr/hal-01425962.

[8] K. BHARGAVAN, A. DELIGNAT-LAVAUD, C. FOURNET, A. GOLLAMUDI, G. GONTHIER, N. KOBEISSI, N. KULATOVA, A. RASTOGI, T. SIBUT-PINOTE, N. SWAMY, S. ZANELLA-BÉGUELIN.*Formal Verification of Smart Contracts: Short Paper*, in "ACM Workshop on Programming Languages and Analysis for Security", Vienna, Austria, October 2016 [*DOI :* 10.1145/2993600.2993611], https://hal.inria.fr/hal-01400469.

[9] K. BHARGAVAN, G. LEURENT.*On the Practical (In-)Security of 64-bit Block Ciphers: Collision Attacks on HTTP over TLS and OpenVPN*, 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.

[10] K. BHARGAVAN, G. LEURENT.*Transcript Collision Attacks: Breaking Authentication in TLS, IKE, and SSH*, in "Network and Distributed System Security Symposium – NDSS 2016", San Diego, United States, February 2016 [*DOI :* 10.14722/NDSS.2016.23418], https://hal.inria.fr/hal-01244855.

[11] K. BHARGAVAN, J. K. ZINZINDOHOUE, E.-I. BARTZIA.*A Verified Extensible Library of Elliptic Curves*, in "29th IEEE Computer Security Foundations Symposium (CSF)", Lisboa, Portugal, June 2016 [*DOI :* 10.1109/CSF.2016.28], https://hal.inria.fr/hal-01425957.

[12] B. BLANCHET, B. SMYTH.*Automated Reasoning for Equivalences in the Applied Pi Calculus with Barriers*, in "29th IEEE Computer Security Foundations Symposium (CSF'16)", Lisboa, Portugal, June 2016, p. 310 - 324 [*DOI :* 10.1109/CSF.2016.29], https://hal.inria.fr/hal-01423742.

[13] K. CAIRNS, H. HALPIN, G. STEEL.*Security Analysis of the W3C Web Cryptography API*, in "Proceedings of Security Standardisation Research (SSR)", Gaithersberg, United States, Lecture Notes in Computer Science (LNCS), Springer, December 2017, vol. 10074, p. 112 - 140 [*DOI :* 10.1007/978-3-319-49100-4_5], https://hal.inria.fr/hal-01426852.

[14] K. ERMOSHINA, F. MUSIANI, H. HALPIN.*End-to-End Encrypted Messaging Protocols: An Overview*, in "Third International Conference, INSCI 2016 - Internet Science", Florence, Italy, F. BAGNOLI, A. SATSIOU, I. STAVRAKAKIS, P. NESI, G. PACINI, Y. WELP, T. TIROPANIS, D. DIFRANZO (editors), Lecture Notes in Computer Science (LNCS), Springer, September 2016, vol. 9934, p. 244 - 254 [*DOI :* 10.1007/978-3-319-45982-0_22], https://hal.inria.fr/hal-01426845.

[15] H. HALPIN.*The Responsibility of Open Standards in the Era of Surveillance*, in "Hot Topics in Privacy Enchancing Technologies", Darmstadt, Germany, HotPETS, July 2016, https://hal.inria.fr/hal-01426848.

[16] 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", Vienna, Austria, WPES, October 2016, p. 139 - 142 [*DOI :* 10.1145/2994620.2994637], https://hal.inria.fr/hal-01426847.

[17] Y. JUGLARET, C. HRIŢCU, A. A. D. AMORIM, B. ENG, B. C. PIERCE.*Beyond Good and Evil: Formalizing the Security Guarantees of Compartmentalizing Compilation*, in "29th IEEE Symposium on Computer Security Foundations (CSF)", Lisabon, Portugal, IEEE Computer Society Press, 2016, p. 45–60 [*DOI :* 10.1109/CSF.2016.11], https://hal.archives-ouvertes.fr/hal-01424795.

[18] L. LAMPROPOULOS, D. GALLOIS-WONG, C. HRIŢCU, J. HUGHES, B. C. PIERCE, L.-Y. XIA.*Beginner's Luck: A Language for Random Generators*, in "44th ACM SIGPLAN Symposium on Principles of Program-

ming Languages (POPL)", Paris, France, ACM, 2017, p. 114-129 [*DOI :* 10.1145/3009837.3009868], https://hal.archives-ouvertes.fr/hal-01424793.

[19] E. SPARROW, H. HALPIN, K. KANEKO, R. POLLAN.*LEAP: A next-generation client VPN and encrypted email provider*, in "Proceedings of Cryptology and Network Security (CANS)", Milan, Italy, Lecture Notes in Computer Science (LNCS), Springer, November 2016, vol. 10052., p. 176 - 191 [*DOI :* 10.1007/978-3-319-48965-0_11], https://hal.inria.fr/hal-01426850.

[20] N. SWAMY, C. HRIŢCU, C. KELLER, A. RASTOGI, A. DELIGNAT-LAVAUD, S. FOREST, K. BHARGAVAN, C. FOURNET, P.-Y. STRUB, M. KOHLWEISS, J.-K. ZINZINDOHOUE, S. ZANELLA-BÉGUELIN.*Dependent Types and Multi-Monadic Effects in F**, in "43rd ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL)", St. Petersburg, Florida, United States, ACM, 2016, p. 256-270 [*DOI :* 10.1145/2837614.2837655], https://hal.archives-ouvertes.fr/hal-01265793.

Research Reports

[21] M. ABADI, B. BLANCHET, C. FOURNET.*The Applied Pi Calculus: Mobile Values, New Names, and Secure Communication*, ArXiv, September 2016, 110, https://hal.inria.fr/hal-01423924.

[22] K. BHARGAVAN, A. DELIGNAT-LAVAUD, N. KOBEISSI.*A Formal Model for ACME: Analyzing Domain Validation over Insecure Channels*, Inria Paris ; Microsoft Research Cambridge, November 2016, https://hal.inria.fr/hal-01397439.

[23] B. BLANCHET, B. SMYTH.*Automated reasoning for equivalences in the applied pi calculus with barriers*, Inria Paris, April 2016, n° RR-8906, 54, https://hal.inria.fr/hal-01306440.

References in notes

[24] *ARINC SPECIFICATION 823P1: DATALINK SECURITY, PART 1 – ACARS MESSAGE SECURITY*, December 2007.

[25] M. ABADI, B. BLANCHET.*Analyzing Security Protocols with Secrecy Types and Logic Programs*, in "Journal of the ACM", January 2005, vol. 52, n° 1, p. 102–146.

[26] 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° 3, p. 1–59.

[27] M. ABADI, C. FOURNET.*Mobile Values, New Names, and Secure Communication*, in "28th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL'01)", London, United Kingdom, ACM Press, January 2001, p. 104–115.

[28] J. BENGTSON, K. BHARGAVAN, C. FOURNET, A. D. GORDON, S. MAFFEIS.*Refinement types for secure implementations*, in "ACM Trans. Program. Lang. Syst.", 2011, vol. 33, n° 2, 8.

[29] K. BHARGAVAN, A. DELIGNAT-LAVAUD, S. MAFFEIS.*Language-Based Defenses Against Untrusted Browser Origins*, in "Proceedings of the 22th USENIX Security Symposium", 2013.

[30] K. BHARGAVAN, C. FOURNET, R. CORIN, E. ZALINESCU.*Verified Cryptographic Implementations for TLS*, in "ACM Transactions Inf. Syst. Secur.", March 2012, vol. 15, n° 1, 3:1.

[31] K. BHARGAVAN, C. FOURNET, A. D. GORDON.*Modular Verification of Security Protocol Code by Typing*, in "ACM Symposium on Principles of Programming Languages (POPL'10)", 2010, p. 445–456.

[32] K. BHARGAVAN, C. FOURNET, A. D. GORDON, N. SWAMY.*Verified Implementations of the Information Card Federated Identity-Management Protocol*, in "Proceedings of the ACM Symposium on Information, Computer and Communications Security (ASIACCS'08)", ACM Press, 2008, p. 123–135.

[33] 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.

[34] B. BLANCHET.*An Efficient Cryptographic Protocol Verifier Based on Prolog Rules*, in "14th IEEE Computer Security Foundations Workshop (CSFW'01)", 2001, p. 82–96.

[35] B. BLANCHET.*Automatic Verification of Correspondences for Security Protocols*, in "Journal of Computer Security", July 2009, vol. 17, n^{o} 4, p. 363–434.

[36] 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.

[37] J. CLULOW.*On the Security of PKCS#11*, in "CHES", 2003, p. 411-425.

[38] 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.

[39] S. DELAUNE, M. D. RYAN, B. SMYTH.*Automatic verification of privacy properties in the applied pi-calculus*, in "IFIPTM'08: 2nd Joint iTrust and PST Conferences on Privacy, Trust Management and Security", International Federation for Information Processing (IFIP), Springer, 2008, vol. 263, p. 263–278.

[40] 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.

[41] C. FOURNET, M. KOHLWEISS, P.-Y. STRUB.*Modular Code-Based Cryptographic Verification*, in "ACM Conference on Computer and Communications Security", 2011.

[42] 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.

[43] N. SWAMY, J. CHEN, C. FOURNET, P.-Y. STRUB, K. BHARGAVAN, J. YANG.*Secure distributed programming with value-dependent types*, in "16th ACM SIGPLAN international conference on Functional Programming", 2011, p. 266-278.

[44] N. SWAMY, J. CHEN, C. FOURNET, P.-Y. STRUB, K. BHARGAVAN, J. YANG.*Secure distributed programming with value-dependent types*, in "J. Funct. Program.", 2013, vol. 23, n^{o} 4, p. 402-451.

[45] 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.

Project-Team QUANTIC

QUANTum Information Circuits

IN COLLABORATION WITH: Laboratoire Pierre Aigrain

IN PARTNERSHIP WITH:
CNRS

Centre Automatique et Systèmes

Ecole normale supérieure de Paris

Mines ParisTech

Université Pierre et Marie Curie (Paris 6)

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:

1.1.11. - Quantum architectures
4.2. - Correcting codes
6. - Modeling, simulation and control
6.1. - Mathematical Modeling
6.1.1. - Continuous Modeling (PDE, ODE)
6.1.2. - Stochastic Modeling (SPDE, SDE)
6.1.3. - Discrete Modeling (multi-agent, people centered)
6.1.4. - Multiscale modeling
6.2. - Scientific Computing, Numerical Analysis & Optimization
6.2.1. - Numerical analysis of PDE and ODE
6.2.3. - Probabilistic methods
6.2.6. - Optimization
6.3.1. - Inverse problems
6.3.2. - Data assimilation
6.3.3. - Data processing
6.3.4. - Model reduction
6.4. - Automatic control
6.4.1. - Deterministic control
6.4.2. - Stochastic control
6.4.3. - Observability and Controlability
6.4.4. - Stability and Stabilization

Other Research Topics and Application Domains:

5.3. - Nanotechnology
5.4. - Microelectronics
6.5. - Information systems
9.8. - Privacy

1. Members

Research Scientists

Mazyar Mirrahimi [Team Leader, Inria, Senior Researcher, HDR]
Alain Sarlette [Inria, Researcher]
Benjamin Huard [CNRS, Researcher, HDR]

Faculty Members

Mallet François [Université Pierre et Marie Curie, Associate professor]
Zaki Leghtas [ENSM Paris, Associate professor]
Pierre Rouchon [ENSM Paris, Professor, HDR]

Technical Staff

Nicolas Didier [Inria]

PhD Students

Remi Azouit [ENSM Paris]
Gerardo Cardona Sanchez [ENSM Paris, from Nov 2016]
Joachim Cohen [Inria, until Oct 2016]
Pierre Six [ENSM Paris, until Nov 2016]
Lucas Verney [ENS Paris]
Lescanne Raphael [ENS Paris]
Markovic Danijela [ENS Paris]
Cottet Nathanael [ENS Paris]

Post-Doctoral Fellows

Zibo Miao [Inria, from Nov 2016]
Jezouin Sébastien [ENS Paris]

Administrative Assistant

Martine Verneuille [Inria]

Others

Francesca Carlotta Chittaro [Univ. Toulon, Associate professor, from Feb 2016 until Sep 2016]
Shantanu Mundhada [Yale Univ., PhD, from May 2016 until Jul 2016]
Pantita Palittapongarnpim [Univ. of Calgary, PhD, from Apr 2016 until Jul 2016]
Michel Sorine [Habilite]

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. Towards microwave quantum networks

The classical states of microwave radiation, are the so-called coherent states. They can be prepared by a commercial microwave generator (frequency $1\text{GHz} < f < 20\text{GHz}$) followed by thermalization to $k_B T \ll hf$ using a chain of attenuators anchored at various stages of a dilution refrigerator.

Owing to the strength of its coupling to superconducting circuits [53] or Rydberg atoms [70], microwave radiation can also be prepared in many possible non-classical states. Using a sequence of quanta exchanges between superconducting qubits and a microwave cavity, the direct preparation of an arbitrary superposition of Fock states has been demonstrated in 2009 [72] with about 90% fidelity up to 5 photons. Recently, the physicists at Yale university in collaboration with the theorists of QUANTIC team, demonstrated a superposition of classical states, or Schrödinger cat, with 100 photons on average, using the dispersive coupling to a transmon qubit [121].

An important class of states for quantum information processing with continuous variables is that of the Gaussian squeezed states [122]. These states can be seen as a coherent state for which the fluctuations on a quadrature are less than the zero point fluctuations. Of course, owing to Heisenberg uncertainty principle, this comes at the expense of larger fluctuations on the conjugated quadrature. In the optical domain, Gaussian light has been demonstrated and used with single and multimodes decades ago [122]. In the microwave domain, single mode squeezing of thermal noise had been demonstrated already in 1988 [127] but vacuum noise squeezing was only demonstrated in 2008 [50]. Since then, several groups have been able to generate single- and two-mode squeezing of microwave radiation, including us [57], [124], [88], [92], [59]. The two-mode squeezed states are of particular interest for quantum information processing, because they are maximally entangled for a given average number of quanta. In particular, the circuit developed by QUANTIC's experimentalists is able to directly generate two-mode squeezed states on separate transmission lines, at arbitrarily different frequencies [59].

In the perspective of a quantum network using microwave radiation, one needs a way to store and preserve microwave fields in nodes. Arguably, creating a memory for quantum systems able to preserve indefinitely a quantum state is the next big challenge on the road towards quantum computing [54], yet unrealized in any system. In a first step, we focus on a quantum node able to preserve a quantum state for a finite time.

In the optical domain, current implementations of quantum memories [112] rely mainly on two physical effects: the light deceleration due to electromagnetically induced transparency and the transfer of photonic quantum states onto collective atomic coherences (optical or spin). In the microwave domain, several quantum memories have emerged in the last years using spin ensembles [125], [78], [107], mechanical resonators [96], [97] or superconducting circuits [126], [123], such as our device described in [60].

All these microwave implementations have pros and cons. However, only two of them, the mechanical oscillator of the Lehnert group [97] and our device [60] have demonstrated entanglement between the memory and a propagating microwave mode. Specifically, our device consists in a 3D storage microwave cavity whose coupling to a transmission line is performed using an active superconducting circuit: the Josephson ring modulator. In the frequency conversion regime, it acts as a tunable coupler whose rate is solely controlled by the amplitude of a pump signal. In the parametric down-conversion regime, it acts as an entanglement generator, similarly to the mechanical version of the Boulder group. However, the inherently small coupling rate between the transmission line and the mechanical resonator in [97] makes our device [60] a much stronger candidate for a quantum node. Apart from this crucial possibility to generate entanglement, our device is similar to the implementation of Santa Barbara [126]. Both have demonstrated fast tuning (up to 30 MHz for Santa Barbara) with high catching efficiency and storage time of 4 μs. However we believe that two specificities make our route more promising. In their case it is a flux knob which allows tuning of the transparency of a 2D microwave cavity. The core of the device we propose is a 3D storage microwave, an architecture where there is plenty of room to improve the storage time and exceed this figure by orders of magnitude, even without quantum error correction [101]. Moreover the cavity transparency is controlled solely by the amplitude of a microwave tone, free of the complications of hysteresis inherent to fast flux tuning in a superconducting environment.

The quantum information protocols one can envision using the quantum node developed by QUANTIC's experimentalists gets a useful inspiration from what has been realized in the optical domain in the last 20 years. One of the most interesting protocols we would like to implement is the teleportation of a quantum state from the memory into a transmission line or another memory. In optics, this was performed already in 1998 for a coherent state [61], and more recently for a Schrödinger-cat-like state [79]. We could readily reproduce these experiments in the microwave regime. The deterministic teleportation of a superconducting quantum bit was realized only in 2013 [116] but no experiments have shown teleportation of a continuous variable state in the microwave domain up to now. Furthermore, none of the protocols needed for quantum information processing (entanglement distillation and dilution for instance) have ever been realized in the microwave domain with Gaussian states [122]. It is thus of great interest to investigate where the tools specific to superconducting circuits will allow us to go beyond what can be done in the optical domain. In particular, the microwave quantum limited amplifiers [104] developed by QUANTIC's experimentalists lead to unmatched heterodyne measurement efficiencies. Finally using a qubit as a Fock number resolved photocounter unleashes many scenarios in the preparation and manipulation by measurement of an entangled state [93].

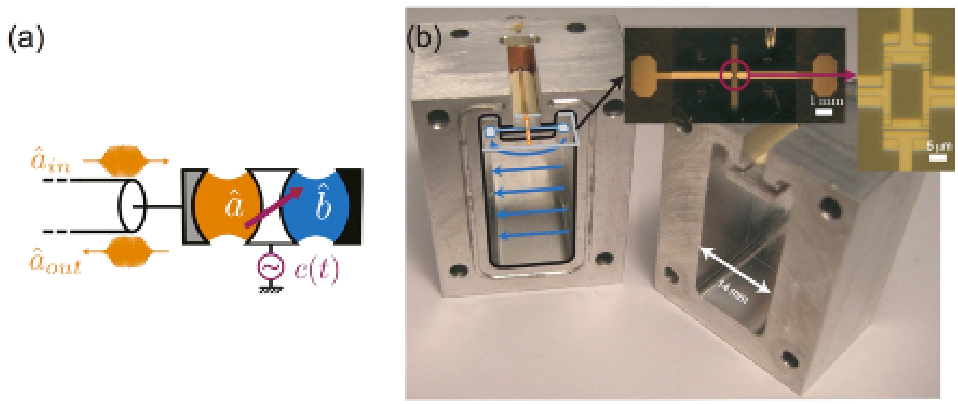

Figure 1. (a) Scheme of the quantum memory. A three-wave mixer is used as a controllable switch between a read/write cavity a and a long storage time cavity b *via the application of a control field c. (b) Picture of the first device. A 2D microstrip resonator on a Sapphire chip is dynamically coupled to a 3D aluminum cavity mode through antennas attached to a ring of 4 Josephson junctions.*

3.2. 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 [111], [115]. 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 [81]. 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 [76], [68] 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 [80]. Through a recent experimental work [121], 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 [117] has recently led to a first experimental realization of a full quantum error correcting code improving the coherence time of quantum information [6]. As shown in Figure 2, 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.

3.3. 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 the 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-demolotion* (QND) measurement has played a crucial role in understanding and resolving this difficulty [44]. In the context of cavity quantum electro-dynamics (cavity

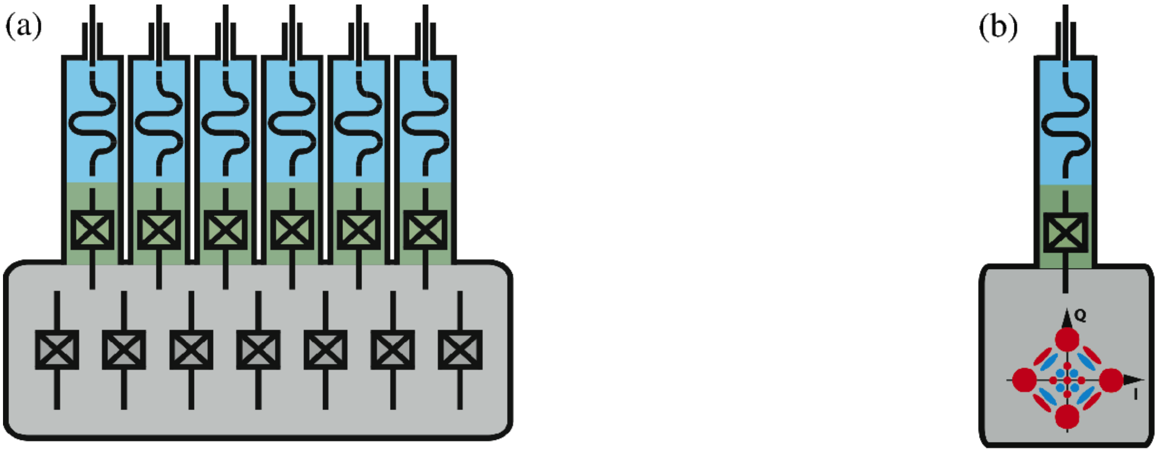

Figure 2. (a) A protected logical qubit consisting of a register of many qubits: here, we see a possible architecture for the Steane code [115] consisting of 7 qubits requiring the measurement of 6 error syndromes. In this sketch, 7 transmon qubits in a high-Q resonator and the measurement of the 6 error 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.

QED) with Rydberg atoms [70], a first experiment on continuous QND measurements of the number of microwave photons was performed by the group at Laboratoire Kastler-Brossel (ENS) [69]. Later on, this 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 [8] (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 [56], [35], [114], [37]. 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) [55], recent advances in quantum-limited amplifiers [104], [119] have opened doors to high-fidelity non-demolition measurements and real-time feedback for superconducting qubits [71]. 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 [119], [103], [46]. 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 [89].

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 [100] and the closely related coherent feedback [86] 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 [67], single-qubit state stabilization [91], and the creation [39] and stabilization [77], [85][9] 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 [67], [83], [51]. The experimental results based on these protocols have illustrated the efficiency of the approach [67][9]. Through these experiments, we exploit the strong dispersive interaction [109] between superconducting qubits and a single low-Q cavity mode playing the role of a dissipative reservoir. Applying some continuous-wave (cw) microwave drives with well-chosen fixed frequencies, amplitudes, and phases, we engineer an effective interaction Hamiltonian which evacuates entropy from the qubits when an eventual perturbation occurs: 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 [90], [106], [73][5].

3.4. 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 [105], [35], [113], [108], [114], [37][7] resulting from collaborations with the cavity quantum electrodynamics group of Laboratoire Kastler Brossel.

3.4.1. 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 [37]: 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 [128]; 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 [98]. 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 [71], [46].

3.4.2. 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 [40] and is related to quantum trajectories [48], [52]. A modern and mathematical exposure of the diffusive models is given in [38]. In [129] 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 [105], [113]. This stability result is extended to a large class of continuous-time filters in [36]. Further efforts are required to characterize asymptotic and exponential stability. Estimations of convergence rates are available only for quantum non-demolition measurements [41]. Parameter estimations based on measurement data of quantum trajectories can be formulated within such quantum filtering framework [62], [94].

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 [45] that post-selection statistics and "past quantum" state analysis [63] enhance sensitivity to parameters and could be interesting towards increasing the precision of an estimation.

3.4.3. 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 [100], [86]. 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 [7], [9] [67], 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 [99]). 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., [34][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 [7], [9] [67], [90] 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 [66], [65]. We will start with [110] and [102] where, based on a theorem due to Birkhoff [42], 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 [118] and by our recent generalized extension of the latter synchronizing interactions to quantum systems [87].

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 [95], 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 [120], [84] with experimental limitations. The characterization of Kraus maps to stabilize any types of states has also been established [43], 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.

3.4.3.1. 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 [70], [64]. 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 2b. 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 [82].

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 [47] for the adiabatic elimination of low-Q cavity). Contrarily 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 [58] and invariant manifold techniques [49] 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 [74] 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 [67][9] [83].

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

- Pierre Rouchon was a plenary speaker at 55th IEEE Conference on Decision and Control.

[0] Here the partiality means that no single quantum measurement is capable of providing the complete information on the state of the system.

- First demonstration of a quantum error correcting code extending the lifetime of a quantum bit: this experiment performed at Yale in collaboration with the team of Robert J. Schoelkopf realizes the hardware-efficient quantum error correction protocol that we had proposed a few years ago. This is the first experiment where a redundant encoding of quantum information, together with continuous measurements of an error syndrome and real-time closed-loop error corrections, extend the lifetime of the encoded information beyond the best physical part. This result was published in Nature [22].
- An experimental marriage of two central concepts of mechanics, the Schrödinger cat states and the entanglement, was realized in collaboration with the team of Robert J. Schoelkopf at Yale. Following our earlier theoretical proposals, an entangled Schrödinger cat state of light shared between two boxes (two high-Q cavities) were successfully achieved and measured. Experimental realization of such states of light were proposed more than 20 years ago and have important applications in quantum information processing. This result was published in Science [28] and has attracted important press converge around the world.
- First experimental demonstration of the quantum-state diffusion associated with spontaneous emission that triggered the field of quantum trajectories in the 1990s. This result was published in Phys. Rev. X [16]. This also led us to implement a first experimental demonstration of multi-input multi-output (MIMO) feedback in the quantum regime. This result was published in Phys. Rev. Lett. [15].

6. New Results

6.1. Observing Quantum State Diffusion by Heterodyne Detection of Fluorescence

Participants: Benjamin Huard, Mazyar Mirrahimi, Pierre Rouchon, Alain Sarlette, Pierre Six.

The results of this section were published in [16] and in [17].

Light emitted via fluorescence is associated with matter decaying in energy, and this light can be viewed as a probe that carries information about the state of its emitter. When this information is lost, the fragile quantum properties of the emitter are destroyed, a process known as decoherence. Using a superconducting qubit, we demonstrate how the sole measurement of fluorescence makes it possible to accurately track the quantum state in time. The observed evolution is erratic, which is expected based on the random backaction of measurements in quantum mechanics.

We continuously measure the amplitude of the fluorescence field emitted by a superconducting qubit using an amplifier close to the quantum limit; our measurements are obtained at cryogenic temperatures. From each fluorescence record, we can reconstruct a quantum trajectory, which is the succession of states the qubit occupies on a single relaxation event. We collect independent measurements of the qubit state at an arbitrary time during relaxation. These measurements follow the statistics that are expected from the quantum trajectories, thereby verifying the reconstructed quantum states. By repeating the experiment millions of times, we are able to determine the distribution of quantum trajectories. Strikingly, monitoring fluorescence can generate a superposition of states and counterintuitively lead to a temporary increase in the qubit excitation probability.

Our work provides an experimental demonstration of the quantum-state diffusion associated with spontaneous emission that triggered the field of quantum trajectories in the 1990s. We expect that our findings, which enlighten the correspondence between decoherence and measurement by the environment, will contribute to the progress of quantum error correction.

In a parallel work, we theoretically investigate statistical properties of the diffusion. In particular, we use a path integral formulation to determine the most likely trajectory during an evolution.

This work was made in collaboration with the team of Andrew Jordan at University of Rochester.

6.2. Using Spontaneous Emission of a Qubit as a Resource for Feedback Control

Participants: Nathanael Cottet, Benjamin Huard, Sebastien Jezouin, François Mallet, Pierre Rouchon, Alain Sarlette, Pierre Six.

The results of this section were published in [15].

We performed an experiment that demonstrates the permanent stabilization of any state of a superconducting qubit despite decoherence using a feedback scheme based on the information leaking out by the relaxation channel itself when the qubit spontaneously emits a photon.

At first sight, it may seem that using the detection of the photon that a qubit emits during a relaxation event cannot allow to protect an arbitrary quantum state from decoherence. First, it is very hard to collect efficiently the photons emitted by a two-level system. Second, the information contained in the emitted photon alone does not seem to be sufficient to correct the effect of relaxation and stabilize an arbitrary qubit state.

However, as we recently showed experimentally (see previous paragraph), it is now possible to measure the spontaneously emitted field using heterodyne detection, and reconstruct the quantum trajectory of a qubit. The information is therefore indeed useful and accessible!

Here, we go well beyond this previous work by not only decoding but also using the information contained in the spontaneously emitted field in real time. Specifically, we use the information contained in fluorescence to stabilize permanently any chosen state of the qubit by measurement feedback.

Stabilizing qubits by a feedback protocol based on the measurement of their relaxation channel had been proposed about 20 years ago by Hofmann and coworkers. They had claimed that it is possible to stabilize any state in the Southern hemisphere of the Bloch sphere. Wang and Wiseman revisited this problem 15 years ago and proposed a scheme that stabilizes any state of the Bloch sphere except the equator. In our work, we devise a new scheme that stabilizes any state, even on the equator! We are also the first ones to implement any such scheme experimentally.

The experiment itself covers several premieres, which are of wider interest to the quantum information and quantum control communities. First, we reach an unprecedented 35% of measurement efficiency for the spontaneously emitted photons out of a qubit (crucial parameter for feedback control). Second, this is the first multiple-input multiple-output feedback in the quantum regime. Finally, we devise a new feedback controller based on the ac-Stark effect to tune the qubit frequency as a function of one input analog signal.

6.3. Well-posedness and convergence of the Lindblad master equation for a quantum harmonic oscillator with multi-photon drive and damping

Participants: Remi Azouit, Pierre Rouchon, Alain Sarlette

The main motivation for this result was to finally treat in a rigorous way the convergence of a non-trivial infinite-dimensional system (harmonic oscillator Hilbert space) that is of relevance to physicists. The essential tools for this proof are the choice of an appropriate metric leading to contraction, and the Hille-Yosida theorem ensuring well-posedness of the problem. This could be a valuable basis towards a more general, yet easily invocable argument to treat the many other infinite-dimensional quantum dynamics which intuitively "should never escape towards infinite energies."

This result has been published in [13].

6.4. Quantum state tomography with non-instantaneous measurements, imperfections, and decoherence

Participants: Pierre Six, Alain Sarlette, Benjamin Huard, Pierre Rouchon

Tomography of a quantum state is usually based on positive operator-valued measure (POVM) and on their experimental statistics. Among the available reconstructions, the maximum-likelihood (MaxLike) technique is an efficient one. We propose an extension of this technique when the measurement process cannot be simply described by an instantaneous POVM. Instead, the tomography relies on a set of quantum trajectories and their measurement records. This model includes the fact that, in practice, each measurement could be corrupted by imperfections and decoherence, and could also be associated with the record of continuous-time signals over a finite amount of time. The goal is then to retrieve the quantum state that was present at the start of this measurement process. The proposed extension relies on an explicit expression of the likelihood function via the effective matrices appearing in quantum smoothing and solutions of the adjoint quantum filter. It allows to retrieve the initial quantum state as in standard MaxLike tomography, but where the traditional POVM operators are replaced by more general ones that depend on the measurement record of each trajectory. It also provides, aside the MaxLike estimate of the quantum state, confidence intervals for any observable. Such confidence intervals are derived, as the MaxLike estimate, from an asymptotic expansion of multi-dimensional Laplace integrals appearing in Bayesian Mean estimation. This work should allow much more accurate inference of the state achieved by some quantum experiment, before a non-instantaneous measurement process is performed to check its results – distinguishing the loss in fidelity truly incurred by the preparation process, from the loss in fidelity induced only by the benchmarking measurement process which would not be present in the final application. A validation is performed on two sets of experimental data: photon(s) trapped in a microwave cavity subject to quantum non-demolition measurements relying on Rydberg atoms, where we have collaborated with the group of Igor Dotsenko at the LKB, College de France; and the heterodyne fluorescence measurements of a superconducting qubit, with the experimentalists of the QUANTIC team.

This result has been published in [27].

6.5. Adiabatic elimination for open quantum systems with effective Lindblad master equations

Participants: Remi Azouit, Pierre Rouchon, Alain Sarlette

We consider an open quantum system described by a Lindblad-type master equation with two times-scales. The fast time-scale is strongly dissipative and drives the system towards a low-dimensional decoherence-free space. To perform the adiabatic elimination of this fast relaxation, we propose a geometric asymptotic expansion based on the small positive parameter describing the time-scale separation. This expansion exploits geometric singular perturbation theory and center-manifold techniques. We conjecture that, at any order, it provides an effective slow Lindblad master equation and a completely positive parameterization of the slow invariant sub-manifold associated to the low-dimensional decoherence-free space. By preserving complete positivity and trace, two important structural properties attached to open quantum dynamics, we obtain a reduced-order model that directly conveys a physical interpretation since it relies on effective Lindbladian descriptions of the slow evolution. At the first order, we derive simple formulae for the effective Lindblad master equation. For a specific type of fast dissipation, we show how any Hamiltonian perturbation yields Lindbladian second-order corrections to the first-order slow evolution governed by the Zeno-Hamiltonian. These results are illustrated on a composite system made of a strongly dissipative harmonic oscillator, the ancilla, weakly coupled to another quantum system.

This result has been published in [30].

6.6. Loss-tolerant parity measurement for distant quantum bits

Participants: Mazyar Mirrahimi, Alain Sarlette

We propose a scheme to measure the parity of two distant qubits, while ensuring that losses on the quantum channel between them does not destroy coherences within the parity subspaces. This last property is a new and essential feature towards using repeated parity measurements in realistic physical conditions. It is achieved thanks to the use of cat states for the probe field that interacts with the two remote qubits. We show how this allows to stabilize highly entangled states between distant qubits, with the current state-of-the-art circuit QED capabilities. Highly entangled states are envisioned as a fundamental building block of the so-called modular quantum computing architecture, so their stabilization, i.e rapid availability, can be viewed as a major step towards enabling such technology.

This result has been submitted as a journal paper [23].

6.7. Holonomic quantum control with continuous variable systems

Participants: Mazyar Mirrahimi

In a collaboration with the team of Liang Jiang at Yale University we propose a scheme to realize a set of universal gates on protected cat-qubits. Universal computation of a quantum system consisting of superpositions of well-separated coherent states of multiple harmonic oscillators can be achieved by three families of adiabatic holonomic gates. The first gate consists of moving a coherent state around a closed path in phase space, resulting in a relative Berry phase between that state and the other states. The second gate consists of "colliding" two coherent states of the same oscillator, resulting in coherent population transfer between them. The third gate is an effective controlled-phase gate on coherent states of two different oscillators. Such gates should be realizable via reservoir engineering of systems that support tunable nonlinearities, such as trapped ions and circuit QED.

This result has been published in [11].

6.8. A Schrodinger cat living in two boxes

Participants: Mazyar Mirrahimi

Quantum superpositions of distinct coherent states in a single-mode harmonic oscillator, known as cat states, have been an elegant demonstration of Schrodinger's famous cat paradox. Here, in a collaboration with the team of Robert Schoelkopf at Yale university, we realize a two-mode cat state of electromagnetic fields in two microwave cavities bridged by a superconducting artificial atom, which can also be viewed as an entangled pair of single-cavity cat states. We present full quantum state tomography of this complex cat state over a Hilbert space exceeding 100 dimensions via quantum nondemolition measurements of the joint photon number parity. The ability to manipulate such multicavity quantum states paves the way for logical operations between redundantly encoded qubits for fault-tolerant quantum computation and communication.

This result has been published in [28].

6.9. Extending the lifetime of a quantum bit with error correction in superconducting circuits

Participants: Zaki Leghtas, Mazyar Mirrahimi

Quantum error correction (QEC) can overcome the errors experienced by qubits and is therefore an essential component of a future quantum computer. To implement QEC, a qubit is redundantly encoded in a higher-dimensional space using quantum states with carefully tailored symmetry properties. Projective measurements of these parity-type observables provide error syndrome information, with which errors can be corrected via simple operations. The break-even point of QEC at which the lifetime of a qubit exceeds the lifetime of the constituents of the system has so far remained out of reach. Although previous works have demonstrated elements of QEC, they primarily illustrate the signatures or scaling properties of QEC codes rather than test the capacity of the system to preserve a qubit over time. Here, in a collaboration with the team of Robert Schoelkopf at Yale University, we demonstrate a QEC system that reaches the break-even point by suppressing the natural errors due to energy loss for a qubit logically encoded in superpositions of Schrodinger-cat states of a superconducting resonator. We implement a full QEC protocol by using real-time feedback to encode, monitor naturally occurring errors, decode and correct. As measured by full process tomography, without any post-selection, the corrected qubit lifetime is 320 microseconds, which is longer than the lifetime of any of the parts of the system: 20 times longer than the lifetime of the transmon, about 2.2 times longer than the lifetime of an uncorrected logical encoding and about 1.1 longer than the lifetime of the best physical qubit (Fock states of the resonator). Our results illustrate the benefit of using hardware-efficient qubit encodings rather than traditional QEC schemes. Furthermore, they advance the field of experimental error correction from confirming basic concepts to exploring the metrics that drive system performance and the challenges in realizing a fault-tolerant system.

This result has been published in [22].

6.10. Robust Concurrent Remote Entanglement Between Two Superconducting Qubits

Participants: Zaki Leghtas

Entangling two remote quantum systems that never interact directly is an essential primitive in quantum information science and forms the basis for the modular architecture of quantum computing. When protocols to generate these remote entangled pairs rely on using traveling single-photon states as carriers of quantum information, they can be made robust to photon losses, unlike schemes that rely on continuous variable states. However, efficiently detecting single photons is challenging in the domain of superconducting quantum circuits because of the low energy of microwave quanta. Here, in a collaboration with the team of Michel Devoret at Yale University, we report the realization of a robust form of concurrent remote entanglement based on a novel microwave photon detector implemented in the superconducting circuit quantum electrodynamics platform of quantum information. Remote entangled pairs with a fidelity of 0.57 are generated at 200 Hz. Our experiment opens the way for the implementation of the modular architecture of quantum computation with superconducting qubits.

This work was published in [21].

6.11. Planar Multilayer Circuit Quantum Electrodynamics

Participants: Zaki Leghtas

Experimental quantum information processing with superconducting circuits is rapidly advancing, driven by innovation in two classes of devices, one involving planar microfabricated (2D) resonators, and the other involving machined three-dimensional (3D) cavities. In a collaboration with the team of Michel Devoret at Yale University, we demonstrate that circuit quantum electrodynamics can be implemented in a multilayer superconducting structure that combines 2D and 3D advantages. We employ standard microfabrication techniques to pattern each layer, and rely on a vacuum gap between the layers to store the electromagnetic energy. Planar qubits are lithographically defined as an aperture in a conducting boundary of the resonators. We demonstrate the aperture concept by implementing an integrated, two-cavity-mode, one-transmon-qubit system.

This work was published in [19].

6.12. Theory of remote entanglement via quantum-limited phase-preserving amplification

Participants: Zaki Leghtas

In a collaboration with the teams of Steven Girvin and Michel Devoret at Yale University, we show that a quantum-limited phase-preserving amplifier can act as a which-path information eraser when followed by heterodyne detection. This “beam splitter with gain” implements a continuous joint measurement on the signal sources. As an application, we propose heralded concurrent remote entanglement generation between two qubits coupled dispersively to separate cavities. Dissimilar qubit-cavity pairs can be made indistinguishable by simple engineering of the cavity driving fields providing further experimental flexibility and the prospect for scalability. Additionally, we find an analytic solution for the stochastic master equation, a quantum filter, yielding a thorough physical understanding of the nonlinear measurement process leading to an entangled state of the qubits. We determine the concurrence of the entangled states and analyze its dependence on losses and measurement inefficiencies.

This work was published in [26].

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. Emergences-Ville de Paris program, QuMotel project

This project, entitled “Quantum memory for microwaves: towards quantum error correction and quantum state teleportation” and led by François Mallet, started on september 2013 and ran till september 2016. It was composed of the members of the QUANTIC project-team. In this project we worked on the development of a decoherence free quantum memory with the tools of circuit quantum electrodynamics. This crucial device is still missing in any implementations of quantum information processing. It aims at capturing, in an efficient manner, the quantum information encoded by flying photons, protect this information over long times, and release it on demand towards a desired channel. The realization of this memory is based on a high quality factor cavity connected to a superconducting circuit performing three-wave mixing. We will entangle the memory state with a propagating microwave signal, then use it to perform quantum teleportation from one memory to another, generate Schrödinger cat states in the memory and realize quantum error correction protocols in order to stabilize a cat state in the memory for an arbitrary time.

7.2. National Initiatives

7.2.1. ANR project GEARED

This three-year collaborative ANR project, entitled “Reservoir engineering quantum entanglement in the microwave domain” and coordinated by Mazyar Mirrahimi, started on October 2014. The participants of the project are Mazyar Mirrahimi, François Mallet and Benjamin Huard (QUANTIC project-team), Daniel Esteve and Fabien Portier (Quantronics group, CEA Saclay), Nicolas Roch and Olivier Buisson (Institut Neel, Grenoble). This project deals with robust generation of entanglement as a key resource for quantum information processing (quantum simulation, computation and communication). The entangled states are difficult to generate and sustain as interaction with a noisy environment leads to rapid loss of their unique quantum properties. Through Geared we intend to investigate different complementary approaches to master the entanglement of microwave photons coupled to quantum superconducting circuits.

7.2.2. 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 to purchase the experimental equipment to build a new experiment based at ENS.

7.3. European Initiatives

7.3.1. Collaborations with Major European Organizations

Partner 1: University of Padova

Alain Sarlette has been pursued a fruitful collaboration with the group of Francesco Ticozzi on “dynamical systems aspects of quantum systems”. A novel line of work in the direction of quantum thermalization and quantum random walks has been explored, in the framework of the PhD of S. Apers (Ghent University) supervised by A. Sarlette. Further joint work for the future is planned about among others generalized Markovian feedback and, reservoir engineering, and linear Lyapunov functions for quantum systems. F. Ticozzi has visited us for one week.

Partner 2: Ghent University.

A. Sarlette is collaborating with applied mathematicians interested in quantum control at UGent (Dirk Aeyels, Lode Wylleman, Gert De Cooman) in the framework of thesis co-supervisions. One PhD student is co-supervised with Dirk Aeyels in the framework of Belgian Inter-University Attraction Poles “Dynamical Systems, Control and Optimization” network 2013-2017. A second PhD student is also co-supervised with Dirk Aeyels in the framework of Chinese Scholarship Council and Flanders Research Fund grant “Developing control mechanisms to counter biases and drifts in coordination”, 2013-2016. Finally, benefiting from a UGent starting grant on “Coordination control algorithms inspired from nonlinear PDEs and lattices”, 2013-2017, Alain Sarlette also supervises a third PhD student at Ghent University.

7.4. International Initiatives

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

TAQUILLA: is an Inria associate team (between Quantic team and Yale university) with principal Inria investigator, Mazyar Mirrahimi, and principal Yale investigator Michel Devoret. In this framework we had many exchanges between Inria and Yale in 2016. Shantanu Mundhada from Yale visited Inria for 2 months. Nicolas Didier and Lucas Verney visited Yale for 3 months, and Joachim Cohen for 3 weeks.

Pierre Rouchon is a participant to the Inria associate Team CDSS with principal Inria investigator, François Dufour of the Inria Team Project CQFD on the topic "Control of dynamic systems subject to stochastic jumps".

7.4.2. Inria International Partners

7.4.2.1. Informal International Partners

Partner 1: University of Yale

The long-term collaborations with the teams of Michel H. Devoret, Robert J. Schoelkopf, Liang Jiang and Steven M. Girvin, enforced through a two year sabbatical visit of Mazyar Mirrahimi at Yale university, have led to a set of contributions ranging from the theoretical analysis and performance optimization of ongoing experiments on weak quantum measurements [71] and preparation of non-classical field states through single photon Kerr effect [75] to the design of new experiments on single qubit cooling [67] and stabilization of maximally entangled states of superconducting qubits [9] by reservoir engineering techniques. Through these collaborations, Zaki Leghtas and Mazyar Mirrahimi have introduced a new direction for hardware-efficient universal quantum computation [81], [90]. These theoretical proposals have already led to groundbreaking experiments [5], [6], [10]. This collaboration is partially formalized through the Taquilla associate team.

Partner 2: University of SaoPaulo and Federal University of Santa Catarina

Pierre Rouchon is collaborating with P. S. Pereira da Silva (Escola Politecnica, PTC, University of SaoPaulo, Brazil) and H. B. Silveira Federal (University of Santa Catarina (UFSC), Florianopolis, Brazil) on the system theory problems behind the experiment on the feedback stabilization of the photon box.

7.5. International Research Visitors

7.5.1. Visits of International Scientists

Francesca Chittaro from Université de Toulon made a 6-month sabbatical visit (February-July 2016) working on adiabatic elimination for composite quantum systems. Preliminary results have been submitted to the IFAC World Congress 2017 [32].

P. S. Pereira da Silva (Escola Politécnica, PTC, University of SaoPaulo, Brazil) made a 3-week visit (June 27 to July 15) to investigate with Mazyar Mirrahimi and Pierre Rouchon controllability issues on composite quantum systems.

7.5.1.1. Internships

In the framework of the Inria-MITACS program, Pantita Palittapongarnpim, student in the group of Barry Sanders at University of Calgary, visited QUANTIC for a period of 4 months working on optimal control methods for photon-number parity measurements.

In the framework of TAQUILLA associate team, Shantanu Mundhada, student in the group of Michel Devoret at Yale University, visited QUANTIC for a period of 2 months working on circuit designs for high-order non-linear quantum dissipation.

Partner: University of Calgary

In the framework of the Inria-MITACS program, Pantita Palittapongarnpim, student in the group of Barry Sanders visited QUANTIC for a period of 4 months working on optimal control methods for photon-number parity measurements.

7.5.2. Visits to International Teams

7.5.2.1. Research Stays Abroad

In the framework of TAQUILLA associate team, Mazyar Mirrahimi spent four months in the Quantronics Laboratory of Michel H. Devoret and in the Rob Schoelkopf Lab at Yale University. Also, in this same framework Nicolas Didier and Lucas Verney spent three months and Joachim Cohen three weeks in the same group.

Pierre Rouchon was invited to give a one-week visit and several lectures on modelling and control of open-quantum systems at Zhejiang University (Hangzhou, China), College of control and Engineering (28 May – 7 June 2016).

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Journal

8.1.1.1. Member of the Editorial Boards

Mazyar Mirrahimi is an associate editor of System and Control Letters.

Mazyar Mirrahimi is a guest editor for the journal "Quantum Science and Technology" (Institute Of Physics, 2016), Special number on “Quantum coherent feedback and quantum reservoir engineering”.

8.1.1.2. Reviewer - Reviewing Activities

Benjamin Huard served as a referee for Nature, Science, Physical Review Letters and other physics journals.

Zaki Leghtas served as a referee for Physical Review Letters and Physical Review X.

Mazyar Mirrahimi served as a referee for Nature and Physical Review Journals.

Pierre Rouchon has been a reviewer for several automatic control and dynamical systems journals and conferences.

Alain Sarlette has been a reviewer for several automatic control and dynamical systems journals and conferences.

8.1.2. Invited Talks

Benjamin Huard, Aug 2016 Formulating and Finding Higher-Order Interference Workshop, Perimeter Institute, Canada.

Benjamin Huard, Jun 2016 4th International Workshop on Frontiers in Quantum Optics and Quantum Information, Beijing Computational Research Center, Beijing, China.

Benjamin Huard, May 2016 Physics department seminar, Chalmers University, Sweden.

Benjamin Huard, May 2016 Statistical Mechanics of Quantum Dynamics, Mariehamn, Finland.

Benjamin Huard, May 2016 Workshop Non-equilibrium thermodynamic phenomena and problems of mesoscopic physics, Aalto University, Finland.

Benjamin Huard, Mar 2016 Mini-Colloque “Rencontre du Non Linéaire” 2016, Paris, France.

Benjamin Huard, Jan 2016 Conference SCALEQIT 2016, Delft, Netherlands.

Zaki Leghtas, Oct 2016, Karlsruhe Institute of Technology, Germany.

Mazyar Mirrahimi, Dec 2016, Yale University, USA.

Mazyar Mirrahimi, Dec 2016, Tutorial at Conference GDR Physique Mesoscopique, Aussois, France.

Mazyar Mirrahimi, Nov 2016, UC Berkeley, USA.

Mazyar Mirrahimi, Oct 2016, University of Pennsylvania, USA.

Mazyar Mirrahimi, June 2016, Journées Scientifiques Inria, Inria Rennes, France.

Mazyar Mirrahimi, July 2016, Tutorial in summer school "Stochastic Methods in Quantum Mechanics", Autrans, France.

Mazyar Mirrahimi, April 2016, "Quantum and Nano Control" workshop, Institute of Mathematics and its Applications of Minneapolis, USA.

Mazyar Mirrahimi, Feb 2016, Institut Néel, Grenoble, France.

Mazyar Mirrahimi, Jan 2016, Conference SCALEQIT 2016, Delft, Netherlands.

Pierre Rouchon, Dec 2016, IEEE Conference on Decision and Control, Las Vegas, USA.

Pierre Rouchon, April 2016, "Quantum and Nano Control" workshop, Institute of Mathematics and its Applications of Minneapolis, USA.

Pierre Rouchon, June 2016, Conference "Nonlinear Partial Differential Equations and Applications" in the honor of Jean-Michel Coron for his 60th birthday, IHP, Paris.

Alain Sarlette, May 2016, workshop on quantum dynamics and control, IHP, Paris.

Alain Sarlette, July 2016, summer school "Stochastic Methods in Quantum Mechanics", Autrans, France.

8.1.3. Scientific Expertise

Benjamin Huard was a member of the ANR CES30 committee in 2016.

Mazyar Mirrahimi is a member of the Technical Committee on "Distributed Parameter Systems" in IFAC (International Federation of Automatic Control).

Pierre Rouchon was "president du comité d'experts" for the HCERES evaluation in January 2016 of the "Laboratoire des Sciences du numériques de Nantes (LS2N)" .

Pierre Rouchon acts as panel member for the panel PE1-Mathematics in the ERC Advanced Grant 2016 evaluation.

Pierre Rouchon is a member of the "Conseil Scientifique du Conservatoire National des Arts et Metiers" since 2014.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Benjamin Huard has given a course (20 hours) entitled "Quantum fluctuations and measurement" in the Master ICFP (International Centre for Fundamental Physics) of ENS Paris, UPMC, Paris 7 and Orsay.

Benjamin Huard has coordinated an experimental project (40 hours) entitled "Measuring the quantum of conductance across an atom" at ENS Paris for Physics students in L3.

Zaki Leghtas taught a course on Quantum Mechanics at Paris Sciences et Lettres (40 hours).

Zaki Leghtas taught a course on Quantum Mechanics and Statistical Physics at Mines ParisTech (12 hours).

Zaki Leghtas taught a course on Complex Analysis at Mines ParisTech (10 hours).

Mazyar Mirrahimi and Pierre Rouchon have given a course (20 hours) entitled "UE : Analyse et contrôle de systèmes quantiques " in the "Master de sciences et technologies, mention mathématiques et applications, Université Pierre et Marie Curie".

Mazyar Mirrahimi has given a graduate-level course (15 hours) entitled "Quantum Control" at Yale University.

Mazyar Mirrahimi has given a 4-hour tutorial on "Quantum measurement and feedback" at the summer school "Stochastic Methods in Quantum Mechanics", Autrans, France.

Pierre Rouchon and Alain Sarlette gave a one week half-time course (15hours) on feedback control of quantum systems at the Elgersburg schools for Mathematical System Theory, 29/2/2016 - 4/3/2016, Germany.

Alain Sarlette has given a master course on "Probabilistic robotics" at Ghent University (30 hours).

8.2.2. Supervision

PhD in progress: Danijela Markovic, ENS. “Quantum information protocols with microwave quantum optics”. Sept 2014. (advisors: Benjamin Huard and François Mallet).

PhD in progress: Nathanaël Cottet. ENS. “Quantum heat engines based on superconducting circuits”. Sept 2015. (advisor: Benjamin Huard).

PhD in progress: Quentin Ficheux. ENS. “Thermodynamics of quantum information”. Sept 2015. (advisors: Benjamin Huard and Zaki Leghtas).

PhD in progress: Raphael Lescanne. ENS. “Engineering Multi-Photon Dissipation In Superconducting Circuits For Quantum Error Correction”. September 2016. (advisors: Zaki Leghtas and Benjamin Huard).

PhD in progress: Rémi Azouit. Mines Paristech. “Quantum circuits, Input/Output theory and adiabatic elimination”. Sept 2014. (advisor: Pierre Rouchon).

PhD in progress: Gerardo Cardona. Mines ParisTech. "Beyond static gains in analog quantum feedback control". Nov 2016 (advisors: Pierre Rouchon and Alain Sarlette).

PhD in progress: Alain Sarlette is co-supervising 3 PhD students with his former institution UGent (Simon Apers, Zhifei Zhang, Arash Farnam). Simon Apers is working on (quantum) network algorithms accelerations and intends to address other quantum control questions.

PhD: Joachim Cohen. ENS. “Autonomous quantum error correction with superconducting circuits”. Nov 2013 (advisor: Mazyar Mirrahimi), His defense is programmed for Feb 2017.

PhD in progress: Lucas Verney. ENS. “Robust quantum information processing with superconducting circuits”. Sept 2016. (advisors: Zaki Leghtas and Mazyar Mirrahimi).

PhD: Noad Hamze El Badaoui. He has defended his PhD thesis on December 2, 2016. His thesis entitled "Dynamique et estimation paramétrique pour les gyroscopes laser à milieu amplificateur gazeux" was under the supervision of Philippe Martin and Pierre Rouchon.

PhD: Pierre Six. He has defended his PhD thesis on November 22, 2016. His thesis entitled "Estimation d’état et de paramètres pour les systèmes quantiques ouverts" was under the supervision of Pierre Rouchon.

8.2.3. Juries

Benjamin Huard was a member of the PhD defense committees of Katrina Sliwa (Yale University, USA), Antoine Tilloy (ENS Paris, France), Philip Krantz (Chalmers, Sweden), Kristinn Juliusson (CEA Saclay, France), Yehan Liu (Yale University, USA), Pierre Six (Mines ParisTech, France) and of the HdR committee of Caglar Girit (Collège de France, France).

Mazyar Mirrahimi was a member the PhD defense committees of Zhan Shi (Reviewer, University of New South Wales, Australia), Shakib Daryanoush (Reviewer, University of Griffith, Australia), Ying Fu (Reviewer, Université Paris Dauphine), Kristinn Juliusson (CEA Saclay, France).

Alain Sarlette was a member of the PhD defense committee of Bram Vervisch (Ghent University, Belgium).

8.3. Popularization

Mazyar Mirrahimi gave interviews for radios, newspapers, magazines and websites (France Culture, Le Monde, La Recherche, Silicon, Industrie and Technologies).

Pierre Rouchon was invited by the "Département de Mathématiques Appliquées de l'Ecole Polytechnique" to give a talk entilted "Dynamique et contrôle des systèmes: du classique au quantique' for the students of Ecole Polytechnique (April 21, 2016).

Alain Sarlette is answering questions about quantum control and quantum computing on the website "ik-heb-een-vraag.be" where Flemish layman can ask questions to scientific experts.

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] L. BRETHEAU, P. CAMPAGNE-IBARCQ, E. FLURIN, F. MALLET, B. HUARD.*Quantum dynamics of an electromagnetic mode that cannot have N photons*, in "Science", May 2015, vol. 348, p. 776-779 [*DOI :* 10.1126/SCIENCE.1259345], https://hal.archives-ouvertes.fr/hal-01154446.

[3] P. CAMPAGNE-IBARCQ, S. JEZOUIN, N. COTTET, P. SIX, L. BRETHEAU, F. MALLET, A. SARLETTE, P. ROUCHON, B. HUARD.*Using Spontaneous Emission of a Qubit as a Resource for Feedback Control*, in "Physical Review Letters", August 2016, vol. 117, 060502, https://hal.inria.fr/hal-01395591.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] C. WANG, Y. GAO, P. REINHOLD, R. HEERES, N. OFEK, K. CHOU, C. AXLINE, M. REAGOR, J. BLUMOFF, K. SLIWA, L. FRUNZIO, S. GIRVIN, L. JIANG, M. MIRRAHIMI, M. H. DEVORET, R. J. SCHOELKOPF.*A Schrodinger cat living in two boxes*, in "Science", 2016, vol. 352, 5.

Publications of the year

Articles in International Peer-Reviewed Journal

[11] V. V. ALBERT, C. SHU, S. KRASTANOV, C. SHEN, R.-B. LIU, Z.-B. YANG, R. J. SCHOELKOPF, M. MIRRAHIMI, M. H. DEVORET, L. JIANG.*Holonomic Quantum Control with Continuous Variable Systems*, in "Physical Review Letters", 2016, vol. 116, n° 14, https://hal.inria.fr/hal-01240208.

[12] S. APERS, A. SARLETTE.*Accelerating consensus by spectral clustering and polynomial filters*, in "IEEE Transactions on Control of Network Systems", 2016, conditionally accepted, https://hal.inria.fr/hal-01093939.

[13] R. AZOUIT, A. SARLETTE, P. ROUCHON.*Well-posedness and convergence of the Lindblad master equation for a quantum harmonic oscillator with multi-photon drive and damping*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2016, vol. 22, n° 4, p. 1353-1369 [*DOI :* 10.1051/COCV/2016050], https://hal.inria.fr/hal-01395585.

[14] A. BROWET, J. HENDRICKX, A. SARLETTE.*Incompatibility boundaries for properties of community partitions*, in "IEEE Transactions on Network Science and Engineering", 2016, conditionally accepted, https://hal.inria.fr/hal-01395589.

[15] P. CAMPAGNE-IBARCQ, S. JEZOUIN, N. COTTET, P. SIX, L. BRETHEAU, F. MALLET, A. SARLETTE, P. ROUCHON, B. HUARD.*Using Spontaneous Emission of a Qubit as a Resource for Feedback Control*, in "Physical Review Letters", August 2016, vol. 117, 060502 [*DOI :* 10.1103/PHYSREVLETT.117.060502], https://hal.inria.fr/hal-01395591.

[16] 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.

[17] A. N. JORDAN, A. N. CHANTASRI, P. ROUCHON, B. N. HUARD.*Anatomy of fluorescence: quantum trajectory statistics from continuously measuring spontaneous emission*, in "Quantum Studies: Mathematics and Foundations", May 2016, vol. 3, p. 237 - 263 [*DOI :* 10.1007/S40509-016-0075-9], https://hal-mines-paristech.archives-ouvertes.fr/hal-01403635.

[18] P. MARTIN, L. ROSIER, P. ROUCHON.*On the reachable states for the boundary control of the heat equation*, in "Applied Mathematics Research eXpress", 2016, vol. 2, p. 181-216, https://hal.archives-ouvertes.fr/hal-01206378.

[19] Z. MINEV, K. SERNIAK, I. POP, Z. LEGHTAS, K. M. SLIWA, M. HATRIDGE, L. FRUNZIO, R. J. SCHOELKOPF, M. H. DEVORET.*Planar multilayer circuit quantum electrodynamics*, in "Physical Review Applied", 2016, vol. 5, 044021 [*DOI :* 10.1103/PHYSREVAPPLIED.5.044021], https://hal-mines-paristech.archives-ouvertes.fr/hal-01403600.

[20] M. MIRRAHIMI.*Cat-qubits for quantum computation*, in "Comptes Rendus Physique", 2016, vol. 17, n^{o} 7, p. 778 - 787 [*DOI :* 10.1016/J.CRHY.2016.07.011], https://hal.inria.fr/hal-01400975.

[21] A. NARLA, S. H. SHANKAR, M. H. HATRIDGE, Z. LEGHTAS, K. M. SLIWA, E. ZALYS-GELLER, O. MUNDHADA, W. H. PFAFF, L. H. FRUNZIO, R. H. SCHOELKOPF, M. H. DEVORET.*Robust Concurrent Remote Entanglement Between Two Superconducting Qubits*, in "Physical Review X", 2016, vol. 6, 031036 [*DOI :* 10.1103/PHYSREVX.6.031036], https://hal-mines-paristech.archives-ouvertes.fr/hal-01403587.

[22] N. OFEK, A. PETRENKO, R. W. HEERES, P. REINHOLD, Z. LEGHTAS, B. VLASTAKIS, Y. LIU, L. FRUNZIO, S. M. GIRVIN, L. JIANG, M. MIRRAHIMI, M. H. DEVORET, R. J. SCHOELKOPF.*Demonstrating Quantum Error Correction that Extends the Lifetime of Quantum Information*, in "Nature", 2016, vol. 536, p. 441-445 [*DOI :* 10.1038/NATURE18949], https://hal-mines-paristech.archives-ouvertes.fr/hal-01403595.

[23] A. SARLETTE, M. MIRRAHIMI.*Loss-tolerant parity measurement for distant quantum bits*, in "(Physical Review A)", 2016, under review for PRA Rapid Communications, https://hal.inria.fr/hal-01395590.

[24] A. SARLETTE.*Adding a single state memory optimally accelerates symmetric linear maps*, in "IEEE Transactions on Automatic Control", 2016, vol. 61, n^{o} 11, 15, 3533, https://hal.inria.fr/hal-01093907.

[25] H. SILVEIRA, P. PEREIRA DA SILVA, P. ROUCHON.*Quantum gate generation for systems with drift in U(n) using Lyapunov-LaSalle techniques*, in "International Journal of Control", 2016 [*DOI :* 10.1080/00207179.2016.1161830], https://hal-mines-paristech.archives-ouvertes.fr/hal-01403629.

[26] M. SILVERI, E. ZALYS-GELLER, M. J. HATRIDGE, Z. LEGHTAS, M. H. DEVORET, S. M. GIRVIN.*Theory of remote entanglement via quantum-limited phase-preserving amplification*, in "Physical Review A", 2016, vol. 93, 062310 [*DOI :* 10.1103/PHYSREVA.93.062310], https://hal-mines-paristech.archives-ouvertes.fr/hal-01403603.

[27] P. SIX, P. CAMPAGNE-IBARCQ, I. DOTSENKO, A. SARLETTE, B. HUARD, P. ROUCHON.*Quantum state tomography with noninstantaneous measurements, imperfections, and decoherence*, in "Physical Review A", 2016, vol. 93, 12109 [*DOI :* 10.1103/PHYSREVA.93.012109], https://hal.inria.fr/hal-01395584.

[28] C. WANG, Y. Y. GAO, P. REINHOLD, R. W. HEERES, N. OFEK, K. CHOU, C. AXLINE, M. REAGOR, J. BLUMOFF, M. SLIWA, L. FRUNZIO, M. GIRVIN, L. JIANG, M. MIRRAHIMI, M. H. DEVORET, R. J.

SCHOELKOPF.*A Schrodinger cat living in two boxes*, in "Science", 2016, vol. 352, n^{o} 6289, p. 1087 - 1091 [*DOI :* 10.1126/SCIENCE.AAF2941], https://hal.inria.fr/hal-01399479.

International Conferences with Proceedings

[29] Z. ZHANG, Z. LING, A. SARLETTE.*Integral control on nonlinear spaces: two extensions*, in "IFAC conference on Nonlinear Control Systems (NOLCOS)", Monterey, United States, International Federation of Automatic Control (IFAC), August 2016, https://hal.inria.fr/hal-01395593.

Conferences without Proceedings

[30] R. AZOUIT, A. SARLETTE, P. ROUCHON.*Adiabatic elimination for open quantum systems with effective Lindblad master equations*, in "55th Conference on Decision and Control (CDC 2016)", Las Vegas, NV, United States, December 2016, https://hal.archives-ouvertes.fr/hal-01398460.

Other Publications

[31] S. APERS, A. SARLETTE, F. TICOZZI.*Fast Mixing with Quantum Walks vs. Classical Processes*, January 2017, Quantum Information Processing (QIP) 2017, Poster, https://hal.inria.fr/hal-01395592.

[32] R. AZOUIT, F. CHITTARO, A. SARLETTE, P. ROUCHON.*Structure-preserving adiabatic elimination for open bipartite quantum systems* *, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01394422.

[33] K. JULIUSSON, S. BERNON, X. ZHOU, V. SCHMITT, H. LE SUEUR, P. BERTET, D. VION, M. MIRAHIMI, P. ROUCHON, D. ESTEVE.*Manipulating Fock states of a harmonic oscillator while preserving its linearity*, July 2016, working paper or preprint, https://hal-cea.archives-ouvertes.fr/cea-01346391.

References in notes

[34] S. ATTAL, A. JOYE, C.-A. PILLET (editors). *Open Quantum Systems III: Recent Developments*, Springer, Lecture notes in Mathematics 1880, 2006.

[35] 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.

[36] 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.

[37] 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.

[38] A. BARCHIELLI, M. GREGORATTI.*Quantum Trajectories and Measurements in Continuous Time: the Diffusive Case*, Springer Verlag, 2009.

[39] 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.

[40] V. BELAVKIN.*Quantum stochastic calculus and quantum nonlinear filtering*, in "Journal of Multivariate Analysis", 1992, vol. 42, n^o 2, p. 171–201.

[41] 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.

[42] G. BIRKHOFF.*Extensions of Jentzch's theorem*, in "Trans. Amer. Math. Soc.", 1957, vol. 85, p. 219–227.

[43] S. BOLOGNANI, F. TICOZZI.*Engineering stable discrete-time quantum dynamics via a canonical QR decomposition*, in "IEEE Trans. Autom. Control", 2010, vol. 55.

[44] V. BRAGINSKI, F. KHALILI.*Quantum Measurements*, Cambridge University Press, 1992.

[45] 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.

[46] 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.

[47] H. CARMICHAEL.*Statistical Methods in Quantum Optics 2: Non-Classical Fields*, Spinger, 2007.

[48] H. CARMICHAEL.*An Open Systems Approach to Quantum Optics*, Springer-Verlag, 1993.

[49] J. CARR.*Application of Center Manifold Theory*, Springer, 1981.

[50] M. CASTELLANOS-BELTRAN, K. IRWIN, G. HILTON, L. VALE, K. LEHNERT.*Amplification and squeezing of quantum noise with a tunable Josephson metamaterial*, in "Nature Physics", 2008, vol. 4, 928.

[51] J. COHEN, M. MIRRAHIMI.*Dissipation-induced continuous quantum error correction for superconducting circuits*, in "Phys. Rev. A", 2014, vol. 90, 062344.

[52] 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.

[53] M. DEVORET, S. GIRVIN, R. SCHOELKOPF.*Circuit-QED: How strong can the coupling between a Josephson junction atom and a transmission line resonator be?*, in "Annalen der Physik", 2007, vol. 16, 767.

[54] M. DEVORET, R. SCHOELKOPF.*Superconducting Circuits for Quantum Information: An Outlook*, in "Science", 2013, vol. 339, p. 1169–1174.

[55] M. H. DEVORET, A. WALLRAFF, J. MARTINIS.*Superconducting Qubits: A Short Review*, 2004, arXiv:cond-mat/0411174.

[56] 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.

[57] C. EICHLER, D. BOZYIGIT, C. LANG, M. BAUR, L. STEFFEN, J. FINK, S. FILIPP, A. WALLRAFF.*Observation of Two-Mode Squeezing in the Microwave Frequency Domain*, in "Phys. Rev. Lett.", 2011, vol. 107, 113601.

[58] N. FENICHEL.*Geometric singular perturbation theory for ordinary differential equations*, in "J. Diff. Equations", 1979, vol. 31, p. 53–98.

[59] E. FLURIN, N. ROCH, F. MALLET, M. H. DEVORET, B. HUARD.*Generating Entangled Microwave Radiation Over Two Transmission Lines*, in "Phys. Rev. Lett", 2012, vol. 109, 183901.

[60] E. FLURIN, N. ROCH, J.-D. PILLET, F. MALLET, B. HUARD.*Superconducting quantum node for entanglement and storage of microwave radiation*, in "submitted, preprint at arXiv:1401.5622", 2014.

[61] A. FURUSAWA, J. SORENSEN, S. BRAUNSTEIN, C. FUCHS, H. KIMBLE, E. POLZIK.*Unconditional quantum teleportation*, in "Science", 1998, vol. 282, 706.

[62] J. GAMBETTA, H. M. WISEMAN.*State and dynamical parameter estimation for open quantum systems*, in "Phys. Rev. A", September 2001, vol. 64, n° 4, 042105, http://link.aps.org/doi/10.1103/PhysRevA.64.042105.

[63] S. GAMMELMARK, B. JULSGAARD, K. MÖLMER.*Past Quantum States of a Monitored System*, in "Phys. Rev. Lett.", October 2013, vol. 111, n° 16, 160401, http://link.aps.org/doi/10.1103/PhysRevLett.111.160401.

[64] C. GARDINER, P. ZOLLER.*Quantum Noise*, third, Springer, 2010.

[65] S. GAUBERT, Z. QU.*Checking the strict positivity of Kraus maps is NP-hard*, in "arXiv:1402.1429", 2014.

[66] 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° 8, p. 2902–2948, http://www.sciencedirect.com/science/article/pii/S0022039614000424.

[67] 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.

[68] D. GOTTESMAN, A. KITAEV, J. PRESKILL.*Encoding a qubit in an oscillator*, in "Phys. Rev. A", 2001, vol. 64, 012310.

[69] 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.

[70] S. HAROCHE, J.-M. RAIMOND.*Exploring the Quantum: Atoms, Cavities and Photons*, Oxford University Press, 2006.

[71] 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.

[72] M. HOFHEINZ, H. WANG, M. ANSMANN, R. BIALCZAK, E. LUCERO, M. NEELEY, A. O'CONNELL, D. SANK, J. WENNER, J. MARTINIS, A. CLELAND.*Synthesizing arbitrary quantum states in a superconducting resonator*, in "Nature", 2009, vol. 459, 546.

[73] 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.

[74] T. KATO.*Perturbation Theory for Linear Operators*, Springer, 1966.

[75] G. KIRCHMAIR, B. VLASTAKIS, Z. LEGHTAS, S. NIGG, H. PAIK, E. GINOSSAR, M. MIRRAHIMI, L. FRUNZIO, S. GIRVIN, R. SCHOELKOPF.*Observation of quantum state collapse and revival due to the single-photon Kerr effect*, in "Nature", 2013, vol. 495, p. 205–209.

[76] E. KNILL, R. LAFLAMME, G. MILBURN.*A scheme for efficient quantum computation with linear optics*, in "Nature", 2001, vol. 409, 46.

[77] 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.

[78] Y. KUBO, I. DINIZ, A. DEWES, V. JACQUES, A. DREAU, J.-F. ROCH, A. AUFFEVES, D. VION, D. ESTEVE, P. BERTET.*Storage and retrieval of a microwave field in a spin ensemble*, in "Phys. Rev. A", 2012, vol. 85, 012333.

[79] N. LEE, H. BENICHI, Y. TAKENO, S. TAKEDA, J. WEBB, E. HUNTINGTON, A. FURUSAWA.*Teleportation of Nonclassical Wave Packets of Light*, in "Science", 2011, vol. 332, 330.

[80] 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.

[81] 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.

[82] Z. LEGHTAS, A. SARLETTE, P. ROUCHON.*Adiabatic passage and ensemble control of quantum systems*, in "J. Phys. B", 2011, vol. 44, 154017.

[83] 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.

[84] J.-S. LI, N. KHANEJA.*Ensemble control of Bloch equations*, in "IEEE Trans. Autom. Control", 2009, vol. 54, p. 528–536.

[85] 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.

[86] S. LLOYD.*Coherent quantum feedback*, in "Phys. Rev. A", 2000, vol. 62, 022108.

[87] L. MAZZARELLA, A. SARLETTE, F. TICOZZI.*Consensus for quantum networks: from symmetry to gossip iterations*, in "IEEE Trans. Automat. Control", 2014, in press.

[88] E. MENZEL, R. D. CANDIA, F. DEPPE, P. EDER, L. ZHONG, M. IHMIG, M. HAEBERLEIN, A. BAUST, E. HOFFMANN, D. BALLESTER, K. INOMATA, T. YAMAMOTO, Y. NAKAMURA, E. SOLANO, A. MARX, R. GROSS.*Path Entanglement of Continuous-Variable Quantum Microwaves*, in "Phys. Rev. Lett.", 2012, vol. 109, 250502.

[89] 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.

[90] M. MIRRAHIMI, Z. LEGHTAS, 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 J. Phys.", 2014, vol. 16, 045014.

[91] 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.

[92] K. MURCH, S. WEBER, K. BECK, E. GINOSSAR, I. SIDDIQI.*Reduction of the radiative decay of atomic coherence in squeezed vacuum*, in "Nature", 2013, vol. 499, 62.

[93] A. NARLA, S. SHANKAR, M. HATRIDGE, Z. LEGHTAS, K. SLIWA, E. ZALYS-GELLER, S. MUNDHADA, W. PFAFF, L. FRUNZIO, R. SCHOELKOPF, M. DEVORET.*Robust Concurrent Remote Entanglement Between Two Superconducting Qubits*, in "Phys. Rev. X", 2016, vol. 6, 031036.

[94] 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.

[95] H. NURDIN, M. JAMES, I. PETERSEN.*Coherent quantum LQG control*, in "Automatica", 2009, vol. 45, p. 1837–1846.

[96] T. PALOMAKI, J. HARLOW, J. TEUFEL, R. SIMMONDS, K. LEHNERT.*Coherent state transfer between itinerant microwave fields and a mechanical oscillator*, in "Nature", 2013, vol. 495, 210.

[97] T. PALOMAKI, J. TEUFEL, R. SIMMONDS, K. LEHNERT.*Entangling Mechanical Motion with Microwave Fields*, in "Science", 2013, vol. 342, 710.

[98] 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.

[99] D. PETZ.*Monotone Metrics on matrix spaces*, in "Linear Algebra and its Applications", 1996, vol. 244, p. 81–96.

[100] 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.

[101] M. REAGOR, H. PAIK, G. CATELANI, L. SUN, C. AXLINE, E. HOLLAND, I. POP, N. MASLUK, T. BRECHT, L. FRUNZIO, M. H. DEVORET, L. GLAZMAN, R. SCHOELKOPF.*Reaching 10 ms single photon lifetimes for superconducting aluminum cavities*, in "Applied Physics Letters", 2013, vol. 102, 192604.

[102] 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.

[103] 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.

[104] 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.

[105] 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.

[106] 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.

[107] S. SAITO, X. ZHU, R. AMSUSS, Y. MATSUZAKI, K. KAKUYANAGI, T. SHIMO-OKA, N. MIZUOCHI, K. NEMOTO, W. MUNRO, K. SEMBA.*Towards Realizing a Quantum Memory for a Superconducting Qubit: Storage and Retrieval of Quantum States*, in "Phys. Rev. Lett.", 2013, vol. 111, 107008.

[108] 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.

[109] 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.

[110] 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.

[111] P. SHOR.*Scheme for reducing decoherence in quantum memory*, in "Phys. Rev. A", 1995, vol. 52, p. 2493–2496.

[112] C. SIMON, M. AFZELIUS, J. APPEL, A. B. DE LA GIRODAY, S. J. DEWHURST, N. GISIN, C. Y. HU, F. JELEZKO, S. KROLL, J. H. MULLER, J. NUNN, E. S. POLZIK, J. G. RARITY, H. D. RIEDMATTEN, W. ROSENFELD, A. J. SHIELDS, N. SKOLD, R. M. STEVENSON, R. THEW, I. WALMSLEY, M. WEBER, H. WEINFURTER, J. WRACHTRUP, R. YOUNG.*Quantum memories*, in "European Physical Journal D", 2012, vol. 58, p. 1–22.

[113] 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.

[114] A. SOMARAJU, M. MIRRAHIMI, P. ROUCHON.*Approximate stabilization of infinite dimensional quantum stochastic system*, in "Reviews in Mathematical Physics", 2013, vol. 25, 1350001.

[115] A. STEANE.*Error Correcting Codes in Quantum Theory*, in "Phys. Rev. Lett", 1996, vol. 77, n^{o} 5.

[116] L. STEFFEN, Y. SALATHE, M. OPPLIGER, P. KURPIERS, M. BAUR, C. LANG, C. EICHLER, G. PUEBLA-HELLMANN, A. FEDOROV, A. WALLRAFF.*Deterministic quantum teleportation with feed-forward in a solid state system*, in "Nature", 2013, vol. 500, 319.

[117] 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.

[118] J. TSITSIKLIS.*Problems in decentralized decision making and computation*, in "PhD Thesis, MIT", 1984.

[119] 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.

[120] L. VIOLA, E. KNILL, S. LLOYD.*Dynamical decoupling of open quantum system*, in "Phys. Rev. Lett.", 1999, vol. 82, p. 2417-2421.

[121] 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.

[122] C. WEEDBROOK, S. PIRANDOLA, R. GARCIA-PATRON, N. CERF, T. RALPH, J. SHAPIRO, S. LLOYD.*Gaussian quantum information*, in "Reviews of Modern Physics", 2012, vol. 84, n^{o} 2, p. 621–669.

[123] J. WENNER, Y. YIN, Y. CHEN, R. BARENDS, B. CHIARO, E. JEFFREY, J. KELLY, A. MEGRANT, J. MUTUS, C. NEILL, P. O'MALLEY, P. ROUSHAN, D. SANK, A. VAINSENCHER, T. WHITE, A. KOROTKOV, A. CLELAND, J. MARTINIS.*Catching Shaped Microwave Photons with 99.4% Absorption Efficiency*, in "Physical Review Letters", 2014, vol. 112, 210501.

[124] C. WILSON, G. JOHANSSON, A. POURKABIRIAN, M. SIMOEN, J. R. JOHANSSON, T. DUTY, F. NORI, P. DELSING.*Observation of the dynamical Casimir effect in a superconducting circuit*, in "Nature", 2012, vol. 479, 376.

[125] H. WU, R. GEORGE, A. ARDAVAN, J. WESENBERG, K. MÖLMER, D. SCHUSTER, R. SCHOELKOPF, K. ITOH, J. MORTON, G. BRIGGS.*Storage of multiple coherent microwave excitations in an electron spin ensemble*, in "Phys. Rev. Lett.", 2010, vol. 105, 140503.

[126] Y. YIN, Y. CHEN, D. SANK, P. O'MALLEY, T. WHITE, R. BARENDS, J. KELLY, E. LUCERO, M. MARIANTONI, A. MEGRANT, C. NEILL, A. VAINSENCHER, J. WENNER, A. KOROTKOV, A. CLELAND, J. MARTINIS.*Catch and Release of Microwave Photon States*, in "Phys. Rev. Lett.", 2013, vol. 110, 107001.

[127] B. YURKE, P. KAMINSKY, R. MILLER, E. WHITTAKER, A. SMITH, A. SILVER, R. SIMON.*Observation of 4.2-K equilibrium-noise squeezing via a Josephson-parametric amplifier*, in "Phys. Rev. Lett.", 1988, vol. 60, 764.

[128] 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.

[129] R. VAN HANDEL.*The stability of quantum Markov filters*, in "Infin. Dimens. Anal. Quantum Probab. Relat. Top.", 2009, vol. 12, p. 153–172.

Project-Team RAP

Networks, Algorithms and Probabilities

RESEARCH CENTER
Paris

THEME
Networks and Telecommunications

Table of contents

Project-Team RAP

Creation of the Project-Team: 2004 February 01, end of the Project-Team: 2016 December 31

Keywords:

Computer Science and Digital Science:

1.3. - Distributed Systems
6.1.2. - Stochastic Modeling (SPDE, SDE)
6.1.4. - Multiscale modeling
7.10. - Network science

Other Research Topics and Application Domains:

1.1. - Biology
1.1.2. - Molecular biology
1.1.10. - Mathematical biology
6. - IT and telecom
6.5. - Information systems

1. Members

Research Scientists

Philippe Robert [Team leader, Inria, Senior Researcher, HDR]
Nicolas Broutin [Inria, Researcher, HDR]
Christine Fricker [Inria, Researcher]

PhD Students

Renaud Dessalles [INRA]
Sarah Eugene [Inria, until Aug 2016, granted by FP7 BANGMD-ERCSKIPPERAD project]
Wen Sun [Inria]
Guilherme Thompson [Inria, granted by Brazilian Grant]
Veronica Quintuna Rodriguez [ORANGE Labs]

Post-Doctoral Fellow

Davit Martirosyan [Inria, from Sep 2016]

Visiting Scientist

Ravi Rasendra Mazumdar [Inria]

Administrative Assistant

Nelly Maloisel [Inria]

Others

Yousra Chabchoub [ISEP, Associate Professor]
Hanene Mohamed [Univ. Paris X, Associate Professor]
Thanh-Huy Nguyen [Inria, Intern, from Mar 2016 until Jul 2016]
Othmane Safsafi [ENS Paris, PhD Student]

2. Overall Objectives

2.1. Overall Objectives

The research team RAP (Networks, Algorithms and Communication Networks) was created in 2004 on the basis of a long standing collaboration between engineers at Orange Labs in Lannion and researchers from Inria Paris. The initial objective was to formalize and expand this fruitful collaboration.

At Orange Labs in Lannion, the members of the team are experts in the analytical modeling of communication networks as well as on some of the operational aspects of network management concerning traffic measurements on ADSL networks, for example.

At Inria Paris, the members of RAP have a recognized expertise in modeling methodologies applied to stochastic models of communication networks.

RAP also has the objective of developing new fundamental tools to investigate *probabilistic* models of complex communication networks. We believe that mathematical models of complex communication networks require a deep understanding of general results on stochastic processes. The two fundamental domains targeted are:

1. Design and analysis of algorithms for communication networks.
2. Analysis of scaling methods for Markov processes: fluid limits and functional limit theorems.

From the very beginning, it has been decided that RAP would focus on a number of particular issues over a period of three or four years. The general goal of the collaboration with Orange Labs is to develop, analyze and optimize algorithms for communication networks. The design of algorithms to allocate resources in large distributed systems is currently investigated in the framework of this collaboration:

3. Research Program

3.1. Scaling of Markov Processes

The growing complexity of communication networks makes it more difficult to apply classical mathematical methods. For a one/two-dimensional Markov process describing the evolution of some network, it is sometimes possible to write down the equilibrium equations and to solve them. The key idea to overcome these difficulties is to consider the system in limit regimes. This list of possible renormalization procedures is, of course, not exhaustive. The advantages of these methods lie in their flexibility to various situations and to the interesting theoretical problems they raised.

A fluid limit scaling is a particularly important means to scale a Markov process. It is related to the first order behavior of the process and, roughly speaking, amounts to a functional law of large numbers for the system considered.

A fluid limit keeps the main characteristics of the initial stochastic process while some second order stochastic fluctuations disappear. In “good” cases, a fluid limit is a deterministic function, obtained as the solution of some ordinary differential equation. As can be expected, the general situation is somewhat more complicated. These ideas of rescaling stochastic processes have emerged recently in the analysis of stochastic networks, to study their ergodicity properties in particular.

3.2. Design and Analysis of Algorithms

Data Structures, Stochastic Algorithms

The general goal of the research in this domain is of designing algorithms to analyze and control the traffic of communication networks. The team is currently involved in the design of algorithms to allocate bandwidth in optical networks and also to allocate resources in large distributed networks. See the corresponding sections below.

The team also pursues analysis of algorithms and data structures in the spirit of the former Algorithms team. The team is especially interested in the ubiquitous divide-and-conquer paradigm and its applications to the design of search trees, and stable collision resolution protocols.

3.3. Structure of random networks

This line of research aims at understanding the global structure of stochastic networks (connectivity, magnitude of distances, etc) via models of random graphs. It consists of two complementary foundational and applied aspects of connectivity.

RANDOM GRAPHS, STATISTICAL PHYSICS AND COMBINATORIAL OPTIMIZATION. The connectivity of usual models for networks based on random graphs models (Erdős–Rényi and random geometric graphs) may be tuned by adjusting the average degree. There is a *phase transition* as the average degree approaches one, a *giant* connected component containing a positive proportion of the nodes suddenly appears. The phase of practical interest is the *supercritical* one, when there is at least a giant component, while the theoretical interest lies at the *critical phase*, the break-point just before it appears.

At the critical point there is not yet a macroscopic component and the network consists of a large number of connected component at the mesoscopic scale. From a theoretical point of view, this phase is most interesting since the structure of the clusters there is expected (heuristically) to be *universal*. Understanding this phase and its universality is a great challenge that would impact the knowledge of phase transitions in all high-dimensional models of *statistical physics* and *combinatorial optimization*.

RANDOM GEOMETRIC GRAPHS AND WIRELESS NETWORKS. The level of connection of the network is of course crucial, but the *scalability* imposes that the underlying graph also be *sparse*: trade offs must be made, which required a fine evaluation of the costs/benefits. Various direct and indirect measures of connectivity are crucial to these choices: What is the size of the overwhelming connected component? When does complete connectivity occur? What is the order of magnitude of distances? Are paths to a target easy to find using only local information? Are there simple broadcasting algorithms? Can one put an end to viral infections? How much time for a random crawler to see most of the network?

NAVIGATION AND POINT LOCATION IN RANDOM MESHES. Other applications which are less directly related to networks include the design of improved navigation or point location algorithms in geometric meshes such as the Delaunay triangulation build from random point sets. There the graph model is essentially fixed, but the constraints it imposes raise a number of challenging problems. The aim is to prove performance guarantees for these algorithms which are used in most manipulations of the meshes.

4. New Results

4.1. Random Graphs

Participant: Nicolas Broutin.

Self-similar real trees defined as fixed-points [15]: Random trees that are fixed points of some random decompositions are ubiquitous: the essential building blocks of the scaling limits of graphs, but also various other trees associated to combinatorial models are such trees. We study a general class of fixed-points equations in spaces of measure metric spaces that yield such objects, and study the existence/uniqueness of the fixed-points in the natural spaces of interest. We also obtain geometric information such as fractal dimension or estimates about the degrees directly from the equations. This is joint work with Henning Sulzbach.

4.2. Resource Allocation in Large Data Centres

Participants: Christine Fricker, Philippe Robert, Guilherme Thompson, Veronica Quintuna Rodriguez.

Efficient resource allocation in large data centers has become crucial matter since the expansion in volume and in variety of the internet based services and applications. Everyday examples, such as Video-on-Demand and Cloud Computing are part of this change in the internet environment, bringing new perspectives and challenges with it. Resource pooling (gathering resources to avoid idleness) and resource decentralization (to bring the service “closer" to the user) are too an important topic in service design, specially because of the inherent dichotomy presented in this discussion. Understanding and assessing the performance of such systems ought enable to better resource management and, consequently, better quality of service.

Currently, most systems operate under decentralized policies due to the complexity of managing data exchange on large scale. In such systems, customer demands are served respecting their initial service requirements (a certain video quality, amount of memory or processing power etc.) until the system reaches saturation, which then leads to the blockage of subsequent customer demands. Strategies that rely on the scheduling of tasks are often not suitable to address this load balancing problem as the users expect instantaneous service usage in real time applications, such as video transmission and elastic computation. Our research goal is to understand and redesign its algorithms in order to develop decentralized schemes that can improve global performance using local instantaneous information. This research is made in collaboration with Fabrice Guillemin, from Orange Labs.

In a first approach to this problem, we examined offloading schemes in fog computing context, where one data centers are installed at the edge of the network. We analyze the case with one data center close to user which is backed up by a central (usually bigger) data center. In [10], when a request arrives at an overloaded data center, it is forwarded to the other data center with a given probability, in order to help dealing with saturation and reducing the rejection of requests. In [17], we studied another scheme, where requests are systematically forwarded by the small data to a larger one, but with some trunk reservation to ensure service performance in the second one. We have been able to demonstrate the behavior and performance of these systems, using the invariant distribution of a random walks in the quarter plane, and obtaining explicit expressions for both schemes. Those two papers shed some light in the effectiveness of this fog computing design, by investigating two basic and intuitive policies, whose advantages can now be compared.

In [11] and [16], we investigated allocation schemes which consist in reducing the bandwidth of arriving requests to a minimal value. In the first, this process is initiated when the system is saturated and in the second when the system is close to saturation. We analyzed the effectiveness of such a downgrading policies. In the case of downgrading at saturation, we were able to find an explicit expression of the key performance metrics when two types of customers share a resource and type two asks for the double of resources compared to type one. And, for the second case, we could show that if the system is correctly designed then we can stop losing clients. We developed a mathematical model which allows us to predict system behavior under such a policy and calculate the optimal threshold (in the same scale as the resource) after which downgrading should be initiated. We proved the existence of a unique equilibrium point, around which we have been able to determine the probability a customer receives service at requested quality. We have also shown that system blockage becomes indeed negligible. This policy finds a natural application in the framework of video streaming services and other real time applications. Notably, we are able to derive explicit and simple expressions for many aspects of this system, giving special predictability the outcome of such policy.

Recently, we started to investigate the framework of network function virtualization, another emergent stream stream of research in resource allocation. We start by considering the execution of Virtualized Network Functions (VNFs) in data centers whose capacities are limited and service execution time is constrained by telecommunication protocols. Virtualization practices play a crucial role in the evolution of telecommunications network archiÂtectures, since the service providers can reduce the investment on the edge and share resource more efficiently. Macrofunctions are virtuzatized into micro ones and treat individually. Through simulations and basic mathematical models, we aroused the discussion of three different prioritization policies and their *trade-offs*. The have shown that in for parallelizable macrofunctions (i.e. no order of execution), the greedy algorithm ensures the best performance in terms of execution delay. For chained ones, macrofunctions whose microfunctions need to be run in a certain order, this algorithm is not suitable, the Round Robin and the Dedicated Core policies perform with the same level.

With these results in mind, we have extend our research towards more complex systems, investigating the behaviour of multiple resource systems (such as a Cloud environment, where computational power is provided using unities of CPU and GB of RAM). We analyzed cooperation between data centers offering multiple resources and under imbalanced loads, a problem that naturally arises from the decentralization of resources. Again, we consider instantaneous service. By forwarding some clients across the system, we could design a policy that is allows cooperation between system and preserves service quality at both data centers. We consider two types of demands asking for two types of resources; particularly, type one clients demand more of type one resource (and symmetrically for type two). We have shown that under our forwarding scheme, which offloads clients requiring most of the saturated resource locally at each data center, we can eliminate losses (in a well design system). Some other interesting properties that can help systems designers are as well derived, such as the minimum threshold for the sustainability of such scheme and the offloading rates. A document is being written to further publication.

4.3. Ressource allocation in vehicle sharing systems

Participants: Christine Fricker, Hanene Mohamed, Thanh-Huy Nguyen.

Vehicle sharing systems are becoming an urban mode of transportation, and launched in many cities, as Velib' and Autolib' in Paris. One of the major issues is the availability of the resources: vehicles or free slots to return them. These systems became an important topic in Operation Research and now the importance of stochasticity on the system behavior is commonly admitted. The problem is to understand the system behavior and how to manage these systems in order to provide both resources to users.

Our stochastic model is the first one taking into account the finite number of spots at the stations.

Equivalence of ensembles We used limit local theorems to obtain the asymptotic stationary joint distributions of several station states when the system is large (both numbers of stations and bikes), in the case of finite capacities of the stations. This gives the asymptotic independence property for node states. This widely extends the existing results on heterogeneous bike-sharing systems.

Load balancing policies. Recently we investigated some load balancing algorithms for stochastic networks to improve the bike sharing system behavior. We focus on the choice of the least loaded station among two to return the bike. In real systems, this choice is local. Thus the main challenge is to deal with the choice between two neighboring stations.

For that, a set of N queues, with a local choice policy, is studied. When a customer arrives at queue i, he joins the least loaded queue between queues i and $i+1$. When the load tends to zero, we obtain an asymptotic for the stationary distribution of the number of customers at a queue. It allows to compare local choice, no choice and choice between two chosen at random.

For a bike-sharing homogeneous model, we study a deterministic cooperation between the stations, two by two. Analytic results are achieved in an homogeneous bike-sharing model. They concern the limit as the system is large, the so-called mean-field limit, and its equilibrium point. Results on performance mainly involve an original closed form expression of the stationary blocking probability in the classical join-the-shortest-queue model. These results are compared by simulations with the policy where the users choose the least loaded station between two stations to return close to their destination. It turns out that, because of randomness, the choice between two neighbours gives better performance than grouping stations two by two.

Bike-sharing model with waiting In real systems, if the customer does not find the resource (a bike or an place to return), he can either leave, or search in a neighbouring station, or wait. We extend a basic model to take into account waiting.

4.4. Scaling Methods

Participants: Philippe Robert, Wen Sun.

4.4.1. Fluid Limits in Wireless Networks

This is a collaboration with Amandine Veber (CMAP, École Polytechnique). The goal is to investigate the stability properties of wireless networks when the bandwidth allocated to a node is proportional to a function of its backlog: if a node of this network has x requests to transmit, then it receives a fraction of the capacity proportional to $\log(1+x)$, the logarithm of its current load. This year we completed the analysis of a star network topology with multiple nodes. Several scalings were used to describe the fluid limit behaviour.

4.4.2. *Large Unreliable Stochastic Networks*

The reliability of a large distributed system is studied. The framework is a system where files have several copies on different servers. When one of these servers breaks down, all copies stored on it are lost. These copies can be retrieved afterwards if there is another copy of the same files stored on other servers. In the case where no other copy of a given file is present in the system, it definitely lost. We study two math model on this problem.

In the first model, it is assumed that the duplication process is local, any server has a capacity to make copies to another server, but the capacity can only be used for the copies present on this server. We have studied the asymptotic behavior of this system, i.e. the number of servers is large, via mean field methods. We have shown that asymptotically, the load of each server can be described by a non-linear Markov process. This limiting process can also give an exponential decay of the number of files. This is a joint work with Reza Aghajani, Brown University.

In the second model, two policies for the reassignment of files are studied. It is assumed that each server has a neighborhood, that consists of a set of servers in the system. When a server breaks down, it restarts immediately but empty. Copies on it are reassigned to other servers in the neighborhood, following “Random Choice” (RC) policy or “Power of choices” (PoC) policy.

- (RC) Each copy join a server in the neighborhood at random.
- (PoC) Each copy chooses several servers in the neighborhood at random, and joins the least loaded one.

The asymptotic behaviors of these two policies are investigated through mean field models. We have show that when the number of servers getting large, the load of each server can be approached by a linear (resp. non-linear) Markov process for RC (resp. PoC) policy. The equilibrium distributions of these asymptotic processes are also given. This is a joint work with Inria/UPMC Team Regal.

4.5. Stochastic Models of Biological Networks

Participants: Renaud Dessalles, Sarah Eugene, Philippe Robert, Wen Sun.

4.5.1. *Stochastic Modelling of self-regulation in the protein production system of bacteria.*

This is a collaboration with Vincent Fromion from INRA Jouy-en-Josas, which started in December 2013.

In prokaryotic cells (e.g. E. Coli. or B. Subtilis) the protein production system has to produce in a cell cycle (i.e. less than one hour) more than 10^6 molecules of more than 2500 kinds, each having different level of expression. The bacteria uses more than 67% of its resources to the protein production. Gene expression is a highly stochastic process: bacteria sharing the same genome, in a same environment will not produce exactly the same amount of a given protein. Some of this stochasticity can be due to the system of production itself: molecules, that take part in the production process, move freely into the cytoplasm and therefore reach any target in the cell after some random time; some of them are present in so much limited amount that none of them can be available for a certain time; the gene can be deactivated by repressors for a certain time, etc. We study the integration of several mechanisms of regulation and their performances in terms of variance and distribution. As all molecules tends to move freely into the cytoplasm, it is assumed that the encounter time between a given entity and its target is exponentially distributed.

4.5.1.1. Feedback model

We have also investigated the production of a single protein, with the transcription and the translation steps, but we also introduced a direct feedback on it: the protein tends to bind on the promoter of its own gene, blocking therefore the transcription. The protein remains on it during an exponential time until its detachment caused by thermal agitation.

The mathematical analysis aims at understanding the nature of the internal noise of the system and to quantify it. We tend to test the hypothesis usually made that such feedback permits a noise reduction of protein distribution compared to the "open loop" model. We have made the mathematical analysis of the model (using a scaling to be able to have explicit results), it appeared that reduction of variance compared to an "open loop" model is limited: the variance cannot be reduced for more than 50%.

We proposed another possible effect of the feedback loop: the return to equilibrium is faster in the case of a feedback model compared to the open loop model. Such behaviour can be beneficial for the bacteria to change of command for a new level of production of a particular protein (due, for example, to a radical change in the environment) by reducing the respond time to reach this new average. This study has been mainly performed by simulation and it has been shown that the feedback model can go 50% faster than the open loop results.

4.5.1.2. Models with Cell Cycle

Usually, classical models of protein production do not explicitly represent several aspects of the cell cycle: the volume variations, the division and the gene replication. Yet these aspects have been proposed in literature to impact the protein production. We have therefore proposed a series of "gene-centered" models (that concentrates on the production of only one type of protein) that integrates successively all the aspects of the cell cycle. The goal is to obtain a realistic representation of the expression of one particular gene during the cell cycle. When it was possible, we analytically determined the mean and the variance of the protein concentration using Marked Poisson Point Process framework.

We based our analysis on a simple model where the volume changes across the cell cycle, and where only the mechanisms of protein production (transcription and translation) are represented. The variability predicted by this model is usually assimilated to the "intrinsic noise" (i.e. directly due to the protein production mechanism itself). We then add the random segregation of compounds at division to see its effect on protein variability: at division, every mRNA and every protein has an equal chance to go to either of the two daughter cells. It appears that this division sampling of compounds can add a significant variability to protein concentration. This effect directly depends on the relative variance (Fano factor) of the protein concentration: this effect is stronger as the relative variance is low. The dependence on the relative variance can be explained by considering a simplified model. With parameters deduced from real experimental measures, we estimate that the random segregation of compounds can double the variability of the genes with the lowest relative variance.

Finally, we integrate the gene replication to the model: at some point in the cell cycle, the gene is replicated, hence doubling the transcription rate. We are able to give analytical expressions for the mean and the variance of protein concentration at any moment of the cell cycle; it allows to directly compare the variance with the previous model with division. We show that gene replication has little impact on the protein variability: an environmental state decomposition shows that the part of the variance due to gene replication represents only at most 2% of the total variability predicted by the model.

In the end, these results are compared to the real experimental measure of protein variability. It appears that the models with cell cycle presented above tend to underestimate the protein variability especially for highly expressed proteins.

4.5.1.3. Multi-protein Model

In continuation of the previous models, we propose a model that still considers the division and the gene replication but which also integrates the sharing of common resources: the different genes are in competition for the limited quantity of RNA-polymerases and ribosomes in order to produce the mRNAs and proteins. The goal is to examine if fluctuations in the availability of these macromolecules have an important impact on the protein variability, as it has been suggested in literature. As the model considers the interaction between the different protein productions, one needs to represent all the genes of the bacteria altogether: it is therefore a multi-protein model.

As this model is too complex to be studied analytically, we have developed a procedure to estimate the parameters so that they correspond to real experimental measures. We then perform simulations in order to determine the variance of each protein and compare them with the one predicted by the models with cell cycle previously presented. It appears that the common sharing of RNA-polymerases and ribosomes has a limited impact on the protein production: for most of proteins the variance increases of at most 10%.

Finally, we have investigated other possible sources of variability by presenting other simulations that integrate some specific aspects: variability in the production of RNA-polymerases and ribosomes, uncertainty in the division and DNA replication decisions, etc. None of the considered aspects seems to have a significant impact on the protein variability.

4.5.2. Stochastic Modelling of Protein Polymerization

This is a collaboration with Marie Doumic, Inria MAMBA team.

The first part of our work focuses on the study of the polymerization of protein. This phenomenon is involved in many neurodegenerative diseases such as Alzheimer's and Prion diseases, e.g. mad cow. In this context, it consists in the abnormal aggregation of proteins. Curves obtained by measuring the quantity of polymers formed in in vitro experiments are sigmoids: a long lag phase with almost no polymers followed by a fast consumption of all monomers. Furthermore, repeating the experiment under the same initial conditions leads to somewhat identical curves up to translation. After having proposed a simple model to explain this fluctuations, we studied a more sophisticated model, closer to the reality. We added a conformation step: before being able to polymerize, proteins have to misfold. This step is very quick and remains at equilibrium during the whole process. Nevertheless, this equilibrium depends on the polymerization which is happening on a slower time scale. The analysis of these models involves stochastic averaging principles.

We have also investigated a more detailed model of polymerisation by considering the the evolution of the number of polymers with different sizes $(X_i(t))$ where $X_i(t)$ is the number of polymers of size i at time t. By assuming that the transitions rates are scaled by a large parameter N, it has been shown that, in the limit, the process $(X_i^N(t))$ is converging to the solution of Becker-Döring equations as N goes to infinity. For another model including nucleation, we have given an asymptotic description of the lag time at the first and second order. These results are obtained in particular by proving stochastic averaging theorems.

The second part concerns the study of telomeres. This work is made in collaboration with Zhou Xu, Teresa Teixeira, from IBCP in Paris.
In eukaryotic cells, at each mitosis, chromosomes are shortened, because the DNA polymerase is not able to duplicate one ending of the chromosome. To prevent loss of genetic information- which could be catastrophic for the cell-chromosomes are equipped with telomeres at their endings. These telomeres do not contain any genetic information; they are a repetition of the sequence T-T-A-G-G-G thousands times. At each mitosis, there is therefore a loss of telomere. As it has a finite length, when the telomeres are too short, the cell cannot divide anymore: they enter in replicative senescence. Our model tries to captures the two phases of the shortening of telomeres: first, the initial state of the cells, when the telomerase is still active to repair the telomeres. Second, when the telomerase is inhibited, we try to estimate the senescence threshold, when the replication of the cells stops. See [8].

5. Bilateral Contracts and Grants with Industry

5.1. Bilateral Contracts with Industry

- Contrat de recherche externalisé avec ORANGE SA "Scheduling Global OS". Duration three years 2014-2016.
- *Christine Fricker* is the leader of PGMO project “Systèmes de véhicules en libre-service: Modélisation, Analyse et Optimisation” with G-Scop (CNRS lab, Grenoble) and Ifsttar. From 1 to 3 years. From 1/10/2013 to 30/9/2016.
- PhD grant CJS (Contrat Jeune Scientifique) Frontières du vivant of INRA for Renaud Dessalles.
- PhD grant from Fondation Sciences Mathématiques de Paris for Wen Sun.

- PhD grant from Brazilian Government for Guilherme Thompson.
- CELTIC+ Contract “SENDATE”.

6. Partnerships and Cooperations

6.1. International Research Visitors

6.1.1. Visits of International Scientists

- Louigi Addario-Berri (McGill)
- Gabor Lugosi (ICREA and Pompeu Fabra)
- Christina Goldschmidt (Oxford)
- Hui He (Beijing Normal University)
- Elie de Panafieu (Bell Labs France)
- Minmin Wang (Buenos Aires)
- Neil Walton (UK)

6.1.1.1. Internships

- Thanhh-Huy, Nguyen, M1 student at Polytechnique (Avril-July 2016). Internship on“Resource allocation in vehicle-sharing systems”.

6.1.2. Visits to International Teams

- *Nicolas Broutin* has visited the mathematics department of the University of Bath, the School of Computer Science at McGill University, the computer science laboratory in Bordeaux (LaBRI) and the NYU-ECNU institute for mathematical sciences at NYU Shanghai.
- *Philippe Robert* (17-27 April 2016) and *Wen Sun* (25-29 April 2016) have visited the Department of Applied Mathematics at Brown University.

7. Dissemination

7.1. Promoting Scientific Activities

7.1.1. Scientific events selection

7.1.1.1. Member of the editorial boards

- *Nicolas Broutin* is member of the steering committee of the international meeting on analysis of algorithms (AofA).
- *Philippe Robert* is Associate Editor of the Book Series “Mathématiques et Applications” edited by Springer Verlag and Associate Editor of the journal “Queueing Systems, Theory and Applications”.

7.1.2. Conferences

- *Nicolas Broutin* has given lectures at annual meeting of the ALEA working group of the GDR-IM, the workshop GRAAL that took place at IHES, at the workshop organized in honor of Svante Janson’s 60’s birthday in Sweden, and the workshop on geometric networks at the Symposium on Computational Geometry that was held in Einhoven. He gave a plenary lecture at the conference on analysis of algorithms that was held in Strobl in June. He has also exposed his results at the probability seminar in Grenoble, the seminar of algorithmic and combinatorial geometry at IHP, and at the probability seminar of the NYU-ECNU Math institute in Shanghai.

  Nicolas Broutin has given the following lectures:

- "Almost optimal sparsification of random geometric graphs" on Jan 30 at the Journées Présage at the Computer Science department of the University of Marne-la-Vallée,
- "The scaling limit of inhomogeneous random graphs" on June 29 at the Paris-Bath workshop that took place at the IHP.
- "Combinatorial structures and fixed-points arguments" on Nov 1 at NYU-ECNU Mathematics institute in Shanghai
- "Random graphs and their scaling limits" on Nov 24 at the Hong Kong University of Science and Technology

He has visited the mathematics department of Beijing Normal University (China), the computer science department of McGill University (Canada), the Mathematics Institute at the Hong-Kong University of Science and Technology, the NYU-ECNU institute for mathematical sciences at NYU Shanghai (China).

- *Renaud Dessalles* has given talks at the conference “Les probabilités de demain” (May 2016) and “Journées Modélisation Aléatoire et Statistique” (August 2016).
- *Christine Fricker* and GT have given talks at the Sigmetrics and MAMA workshop, Juan-les-Pins,14-18 /6/2016 and at the 2nd ECQT (European Conference on Queueing Theory) Toulouse, 18-20/7/2016.
- *Veronica Quintuna Rodriguez* gave a talk at the conference “17th International Network Strategy and Planning Symposium” in september (Montreal).
- *Philippe Robert* gave a talk at the department of mathematics of Brown University (April), at the stochastic networks conference in San Diego, and at the conference “Modern Problems in Theoretical and Applied Probability” in Novosibirsk in August.

7.1.3. Scientific expertise

- *Christine Fricker* is member of the jury of agrégation.
- *Philippe Robert* is member of the scientific council of EURANDOM.

7.2. Teaching - Supervision - Juries

7.2.1. Teaching

Master : *Nicolas Broutin* Master Parisien de Recherche en Informatique (MPRI), in the course 2.15 on Analysis of Algorithms.

Master: *Nicolas Broutin* , Analysis of Algorithms, NYU Shanghai.

Master: *Philippe Robert* , Master, Probabilités et Applications, UPMC.

7.2.2. Juries

- *Philippe Robert* has been member of the jury of PhD defense of Sarah Eugène, Claudio Imbrenda and Tanguy Cabana.

8. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] 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.

Articles in International Peer-Reviewed Journal

[2] L. ADDARIO-BERRY, N. BROUTIN, C. GOLDSCHMIDT, G. MIERMONT.*The scaling limit of the minimum spanning tree of the complete graph*, in "Annals of Probability", 2016, 78, to appear, https://hal.inria.fr/hal-00773360.

[3] N. BROUTIN, O. DEVILLERS, R. HEMSLEY.*Efficiently navigating a random Delaunay triangulation*, in "Random Structures and Algorithms", 2016, vol. 49, n^o 1, p. 95–136 [*DOI :* 10.1002/RSA.20630], https://hal.inria.fr/hal-00940743.

[4] N. BROUTIN, L. DEVROYE, G. LUGOSI.*Almost optimal sparsification of random geometric graphs*, in "The Annals of Applied Probability : an official journal of the institute of mathematical statistics", 2016, vol. 26, p. 3078-3109 [*DOI :* 10.1214/15-AAP1170], https://hal.archives-ouvertes.fr/hal-01056127.

[5] N. BROUTIN, J.-F. MARCKERT.*A new encoding of coalescent processes. Applications to the additive and multiplicative cases*, in "Probability Theory and Related Fields", 2016, vol. 166, n^o 1, p. 515–552 [*DOI :* 10.1007/S00440-015-0665-1], https://hal.inria.fr/hal-01092562.

[6] N. BROUTIN, M. WANG.*Cutting down* $\mathbf{p}$*-trees and inhomogeneous continuum random trees*, in "Bernoulli", 2016, 44 pages, 6 figures, https://hal.archives-ouvertes.fr/hal-01056125.

[7] 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.

[8] S. EUGENE, T. BOURGERON, Z. XU.*Effects of initial telomere length distribution on senescence onset and heterogeneity*, in "Journal of Theoretical Biology", January 2017, vol. 413, 8, https://hal.inria.fr/hal-01378596.

[9] 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.

[10] C. FRICKER, F. GUILLEMIN, P. ROBERT, G. R. THOMPSON.*Analysis of an offloading scheme for data centers in the framework of fog computing*, in "ACM Transactions on Modeling and Performance Evaluation of Computing Systems", September 2016, vol. 1, n^o 4, 18 [*DOI :* 10.1145/2950047], https://hal.inria.fr/hal-01179206.

[11] C. FRICKER, F. GUILLEMIN, P. ROBERT, G. R. THOMPSON.*Analysis of Downgrading for Resource Allocation*, in "Performance Evaluation Review", September 2016, vol. 44, n^o 2, 3 [*DOI :* 10.1145/3003977.3003986], https://hal.inria.fr/hal-01406664.

[12] P. ROBERT, J. TOUBOUL.*On the dynamics of random neuronal networks*, in "Journal of Statistical Physics", September 2016 [*DOI :* 10.1007/S10955-016-1622-9], https://hal.inria.fr/hal-01075242.

[13] W. SUN, M. FEUILLET, P. ROBERT.*Analysis of Large Unreliable Stochastic Networks*, in "Annals of Applied Probability", October 2016, vol. 26, n^o 5, 41 [*DOI :* 10.1214/15-AAP1167], https://hal.inria.fr/hal-01203096.

Other Publications

[14] R. AGHAJANI, P. ROBERT, W. SUN.*A Large Scale Analysis of Unreliable Stochastic Networks*, August 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01359208.

[15] N. BROUTIN, H. SULZBACH.*Self-similar real trees defined as fixed-points and their geometric properties*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01384309.

[16] C. FRICKER, F. GUILLEMIN, P. ROBERT, G. R. THOMPSON.*Allocation Schemes of Ressources with Downgrading*, April 2016, working paper or preprint, https://hal.inria.fr/hal-01301291.

[17] F. GUILLEMIN, G. R. THOMPSON.*Analysis of a trunk reservation policy in the framework of fog computing*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01378533.

[18] P. ROBERT.*Jacques Neveu et les modèles probabilistes de réseaux*, August 2016, working paper or preprint, https://hal.inria.fr/hal-01382215.

[19] P. ROBERT, A. VEBER.*A Scaling Analysis of a Star Network with Logarithmic Weights*, September 2016, working paper or preprint, https://hal.inria.fr/hal-01377703.

Project-Team REGAL

Large-Scale Distributed Systems and Applications

IN COLLABORATION WITH: Laboratoire d'informatique de Paris 6 (LIP6)

IN PARTNERSHIP WITH:
CNRS

Université Pierre et Marie Curie (Paris 6)

RESEARCH CENTER
Paris

THEME
Distributed Systems and middleware

Table of contents

Project-Team REGAL

Creation of the Project-Team: 2005 July 01

Keywords:

Computer Science and Digital Science:

1.1.1. - Multicore
1.1.6. - Cloud
1.1.7. - Peer to peer
1.1.9. - Fault tolerant systems
1.1.13. - Virtualization
1.3. - Distributed Systems
1.6. - Green Computing
2.6. - Infrastructure software
2.6.1. - Operating systems
2.6.2. - Middleware
2.6.3. - Virtual machines
3.1.3. - Distributed data
3.1.8. - Big data (production, storage, transfer)
7.1. - Parallel and distributed algorithms

Other Research Topics and Application Domains:

4.5. - Energy consumption
6.4. - Internet of things
8.2. - Connected city
9.2.3. - Video games
9.4.1. - Computer science

1. Members

Research Scientists

Mesaac Makpangou [Inria, Researcher, HDR]
Marc Shapiro [Inria, Senior Researcher, HDR]

Faculty Members

Pierre Sens [Team leader, Univ. Paris VI, Professor, HDR]
Luciana Bezerra Arantes [Univ. Paris VI, Associate Professor]
Swan Dubois [Univ. Paris VI, Associate Professor]
Jonathan Lejeune [Univ. Paris VI, Associate Professor, since Sep 2016]
Sebastien Monnet [Univ. Paris VI, Associate Professor, until Sep 2016]
Franck Petit [Univ. Paris VI, Professor, HDR]
Julien Sopena [Univ. Paris VI, Associate Professor]

Technical Staff

Salvatore Pileggi [Inria, unded by FP7 SyncFree, until Sep 2016]
Tyler Crain [Inria, post-doc funded by FP7 SyncFree, until Sep 2016]

PhD Students

Sébastien Bouchard [Inria, from Oct 2016]

Marjorie Bournat [Univ. Paris VI]
Michael Damien Carver [Magency, under CIFRE grant]
Joao Paulo de Araujo [Univ. Paris VI]
Lyes Hamidouche [Magency, under CIFRE grant]
Denis Jeanneau [Univ. Paris VI]
Mahsa Najafzadeh [Inria, until May 2016]
Vinh Tao Thanh [Scality, under CIFRE grant]
Alejandro Tomsic [Inria]
Guillaume Turchini [Univ. Paris VI]
Gauthier Voron [Univ. Paris VI]
Bassirou Ngom [Univ. Cheikh Anta Diop de Dakar / Univ. Paris VI, until Sep 2016]

Visiting Scientist
Rostom Mennour [Univ. Constantine]

Administrative Assistant
Helene Milome [Inria]

2. Overall Objectives

2.1. Overall Objectives

The research of the Regal team addresses the theory and practice of *Computer Systems,* including multicore computers, clusters, networks, peer-to-peer systems, cloud computing systems, and other communicating entities such as swarms of robots. It addresses the challenges of communicating, sharing information, and computing correctly 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.

Regal is a joint research team between LIP6 and Inria Paris.

3. Research Program

3.1. Research rationale

The research of Regal addresses both theoretical and practical issues of *Computer Systems*, i.e., its goal is a dual expertise in theoretical and experimental research. Our approach is a “virtuous cycle” of algorithm design triggered by issues with real systems, which we prove correct and evaluate theoretically, and then eventually implement and test experimentally.

Regal’s major challenges comprise communication, sharing of information, and correct execution in large-scale and/or highly dynamic computer systems. While Regal’s historically focused in static distributed systems, since some years ago we have covered a larger spectrum of distributed computer systems: multicore computers, clusters, mobile networks, peer-to-peer systems, cloud computing systems, and other communicating entities such as swarms of robots. This holistic approach allows the handling of related problems at different levels. Among such problems we can highlight communication between cores, consensus, fault detection, scalability, search and diffusion of information, allocation resource, replication and consistency of shared data, dynamic content distribution, and multi-core concurrent algorithms.

Computer Systems is a rapidly evolving domain, with strong interactions with industry and modern computer systems, which are increasingly distributed. Ensuring persistence, availability, and consistency of data in a distributed setting is a major requirement: the system must remain correct despite slow networks, disconnection, crashes, failures, churn, and attacks. Easiness of use, performance, and efficiency are equally fundamental. However, these requirements are somewhat conflicting, and there are many algorithmic and engineering trade-offs, which often depend on specific workloads or usage scenarios. At the same time, years of research in distributed systems are now coming to fruition, and are being used by millions of users of web systems, peer-to-peer systems, gaming and social applications, or cloud computing. These new usages bring new challenges of extreme scalability and adaptation to dynamically-changing conditions, where knowledge of the system state might only be partial and incomplete. Therefore, the scientific challenges of the distributed computing systems listed above are subject to additional trade-offs which include scalability, fault tolerance, dynamics, and virtualization of physical infrastructure. 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.

In in particular, Regal focuses on three key challenges:

- the adaptation of algorithms to the new dynamics of distributed systems;
- data management on extreme large configurations;
- the adaptation of execution support to new multi-core architectures.

We should emphasize that these challenges are complementary: the two first challenges aim at building new distributed algorithms and strategies for large and dynamic distributed configurations whereas the last one focusses on the scalability of internal OS mechanisms.

4. Highlights of the Year

4.1. Highlights of the Year

- We initiate a collaboration with ICL Lab (University of Tennessee) to study failure detection in Exascale computing. We designed and evaluated a new robust failure detector. This result is published at SC 2016 [26].

5. New Software and Platforms

5.1. Antidote

FUNCTIONAL DESCRIPTION

Antidote is the flexible cloud database platform currently under development in the SyncFree European project. Antidote aims to be both a research platform for studying replication and consistency at the large scale, and an instrument for exploiting research results. The platform supports replication of CRDTs, in and between sharded (partitioned) data centres (DCs). The current stable version supports strong transactional consistency inside a DC, and causal transactional consistency between DCs. Ongoing research includes support for explicit consistency [23] , for elastic version management, for adaptive replication, for partial replication, and for reconfigurable sharding.

- Participants: Tyler Crain, Marc Shapiro and Alejandro Tomsic
- Contact: Tyler Crain
- URL: https://github.com/SyncFree/

5.2. G-DUR

FUNCTIONAL DESCRIPTION

A large family of distributed transactional protocols have a common structure, called Deferred Update Replication (DUR). DUR provides dependability by replicating data, and performance by not re-executing transactions but only applying their updates. Protocols of the DUR family differ only in behaviors of few generic functions. Based on this insight, we offer a generic DUR middleware, called G-DUR, along with a library of finely-optimized plug-in implementations of the required behaviors.

- Participants: Marc Shapiro and Masoud Saeida Ardekani
- Contact: Marc Shapiro
- URL: https://github.com/msaeida/jessy

5.3. NumaGIC

FUNCTIONAL DESCRIPTION

NumaGiC is a version of the HotSpot garbage collector (GC) adapted to many-core computers with very large main memories. In order to maximise GC throughput, it manages the trade-off between memory locality (local scans) and parallelism (work stealing) in a self-balancing manner. Furthemore, the collector features several memory placement heuristics that improve locality.

- Participants: Lokesh Gidra, Marc Shapiro, Julien Sopena and Gaël Thomas
- Contact: Lokesh Gidra
- URL: http://gforge.inria.fr/projects/transgc/

6. New Results

6.1. Distributed Algorithms for Dynamic Networks and Fault Tolerance

Participants: Luciana Bezerra Arantes [correspondent], Sébastien Bouchart, Marjorie Bournat, Swan Dubois, Denis Jeanneau, Mohamed Hamza Kaaouachi, Sébastien Monnet, Franck Petit [correspondent], 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.

We address both system dynamic and fault tolerance through various aspects: (1) Fault Detection, (2) Self-Stabilization, and (3) Dynamic System Design. Our approach covers the whole spectrum from theory to experimentation. We design algorithms, prove them correct, implement them, and evaluate them within simulation plateforms.

6.1.1. Failure detection

Since 2013, we address both theoretical and practical aspects of failure detector. The failure detector (FD) abstraction has been used to solve agreement problems in asynchronous systems prone to crash failures, but so far it has mostly been used in static and complete networks. FDs are distributed oracles that provide processes with unreliable information on process failures, often in the form of a list of trusted process identities. In 2016 we obtain the following results.

We propose in [31] a new failure detector that expresses the confidence with regard to the system as a whole. Similarly to a reputation approach, it is possible to indicate the relative importance of each process of the system, while a threshold offers a degree of flexibility for failures and false suspicions. Performance evaluation results, based on real PlanetLab traces, confirm the degree of flexible of the failure detector. By logically organizing nodes in a distributed hypercube, denoted VCube, which dynamically re-organizes itself in case of node failures, detected by a hierarchical perfect failure, we have proposed a autonomic distributed quorum algorithm [35]. By replacing the perfect failure detector by another one that offers eventual strong completeness, we have presented in [33] a second autonomic reliable broadcast protocol.

In the context of large networks, we propose Internet Failure Detector Service (IFDS) [16] for processes running in the Internet on multiple autonomous systems. The failure detection service is adaptive, and can be easily integrated into applications that require configurable QoS guarantees. The service is based on monitors which are capable of providing global process state information through a SNMP MIB. Monitors at different networks communicate across the Internet using Web Services. The system was implemented and evaluated for monitored processes running both on single LAN and on PlanetLab. Experimental results are presented, showing the performance of the detector, in particular the advantages of using the self-tuning strategies to address the requirements of multiple concurrent applications running on a dynamic environment.

Finally, in collaboration with ICL Lab. (University of Tennessee), we study failure detection in the context of ExaScale computing. We designed and evaluated a new robust failure detector, able to maintain and distribute the correct list of alive resources within proven and scalable bounds. The detection and distribution of the fault information follow different overlay topologies that together guarantee minimal disturbance to the applications. A virtual observation ring minimizes the overhead by allowing each node to be observed by another single node, providing an unobtrusive behavior. The propagation stage is using a non-uniform variant of a reliable broadcast over a circulant graph overlay network, and guarantees a logarithmic fault propagation. Extensive simulations, together with experiments on the Titan ORNL supercomputer, show that the algorithm performs extremely well, and exhibits all the desired properties of an Exascale-ready algorithm. This work has been published at SC 2016 conference [26].

6.1.2. Self-Stabilization

Regardless its initial state, a *self-stabilizing* system has the ability to reach a correct behavior in finite time. Self-stabilization is a generic paradigm to tolerate transient faults (*i.e.*, faults of finite duration) in distributed systems. Self-stabilization is also a suitable approach to design reliable solutions for dynamic systems. Results obtained in this area by Regal members in 2016 follow.

In [8], we address the ability to maintain distributed structures at large scale. Among the many different structures proposed in this context, The prefix tree structure is a good candidate for indexing and retrieving information. One weakness of using such a distributed structure stands in its poor native fault tolerance, leading to the use of preventive costly mechanisms such as replication. We focus on making tries self-stabilizing over such platforms, and propose a self-stabilizing maintenance algorithm for a prefix tree using a message passing model. The proof of self-stabilization is provided, and simulation results are given, to better capture its performances.

In [4], we propose a silent self-stabilizing leader election algorithm for bidirectional connected identified networks of arbitrary topology. Written in the locally shared memory model, it assumes the distributed unfair daemon, *i.e.*, the most general scheduling hypothesis of the model. Our algorithm requires no global knowledge on the network (such as an upper bound on the diameter or the number of processes). We show that its stabilization time is in $\Theta(n^3)$ steps in the worst case, where n is the number of processes. Its memory requirement is asymptotically optimal, *i.e.*, $\Theta(\log n)$ bits per processes. Its round complexity is of the same order of magnitude — *i.e.*, $\Theta(n)$ rounds — as the best existing algorithms designed with similar settings. To the best of our knowledge, this is the first asynchronous self-stabilizing leader election algorithm for arbitrary identified networks that is proven to achieve a stabilization time polynomial in steps. By contrast, we show that the previous best existing algorithms stabilize in a non polynomial number of steps in the worst case.

A *snap-stabilizing* protocol, regardless of the initial configuration of the system, guarantees that it always behaves according to its specification. In [9], we consider the locally shared memory model. In this model, we propose a snap-stabilizing Propagation of Information with Feedback (PIF) protocol for rooted networks of arbitrary topology. Then, we use the proposed PIF protocol as a key module in the design of snap-stabilizing solutions for some fundamental problems in distributed systems, such as Leader Election, Reset, Snapshot, and Termination Detection. Finally, we show that in the locally shared memory model, snap-stabilization is as expressive as self-stabilization by designing a universal transformer to provide a snap-stabilizing version of any protocol that can be (automatically) self-stabilized. Since by definition, a snap-stabilizing algorithm is self-stabilizing, self- and snap-stabilization have the same expressiveness in the locally shared memory model.

In [6], we address the *committee coordination problem*: A committee consists of a set of professors and committee meetings are synchronized, so that each professor participates in at most one committee meeting at a time. We propose two snap-stabilizing distributed algorithms for the committee coordination. They are enriched with some desirable properties related to concurrency, (weak) fairness, and a stronger synchronization mechanism called 2-Phase Discussion. Existing work in the literature has shown that (1) in general, fairness cannot be achieved in committee coordination, and (2) it becomes feasible if each professor waits for meetings infinitely often. Nevertheless, we show that even under this latter assumption, it is impossible to implement a fair solution that allows maximal concurrency. Hence, we propose two orthogonal snap-stabilizing algorithms, each satisfying 2-phase discussion, and either maximal concurrency or fairness.

6.1.3. Dynamic Distributed Systems

In [19], we introduce the notion of *gradually stabilizing* algorithm as any self-stabilizing algorithm with the following additional feature: if at most τ*dynamic steps*—a dynamic step is a step containing topological changes—occur starting from a legitimate configuration, it first quickly recovers to a configuration from which a minimum quality of service is satisfied and then gradually converges to stronger and stronger safety guarantees until reaching a legitimate configuration again. We illustrate this new property by proposing a gradually stabilizing unison algorithm, that consists in synchronizing logical clocks locally maintained by the processes.

The next results consider highly dynamic distributed systems modelled by time-varying graphs (TVGs). In [7], we first address proof of impossibility results that often use informal arguments about convergence. We provide a general framework that formally proves the convergence of the sequence of executions of any deterministic algorithm over TVGs of any convergent sequence of TVGs. Next, we focus of the weakest class of long-lived TVGs, *i.e.*, the class of TVGs where any node can communicate any other node infinitely often. We illustrate the relevance of our result by showing that no deterministic algorithm is able to compute various distributed covering structure on any TVG of this class. Namely, our impossibility results focus on the eventual footprint, the minimal dominating set and the maximal matching problems.

We also study the k-set agreement problem, a generalization of the consensus problem where processes can decide up to k different values. Very few papers have tackled this problem in dynamic networks. Exploiting the formalism of TVGs, we propose in [11] a new quorum-based failure detector for solving k-set agreement in dynamic networks with asynchronous communications. We present two algorithms that implement this new failure detector using graph connectivity and message pattern assumptions. We also provide an algorithm for solving k-set agreement using our new failure detector.

Finally, in [22], 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.

6.2. Large scale data distribution

Participants: Luciana Arantes [correspondent], Rudyar Cortes, Mesaac Makpangou, Sébastien Monnet, Pierre Sens.

The proliferation of GPS-enabled devices leads to the massive generation of geotagged data sets recently known as Big Location Data. It allows users to explore and analyse data in space and time, and requires an architecture that scales with the insertions and location-temporal queries workload from thousands to millions of users. Most large scale key-value data storage solutions only provide a single one-dimensional index which does not natively support efficient multidimensional queries. In 2016, we propose GeoTrie [29], a scalable architecture built by coalescing any number of machines organized on top of a Distributed Hash Table. The key idea of our approach is to provide a distributed global index which scales with the number of nodes and provides natural load balancing for insertions and location-temporal range queries. We assess our solution using the largest public multimedia data set released by Yahoo! which includes millions of geotagged multimedia files.

We also propose ECHO [10], a novel and lightweight solution that efficiently supports range queries over a ring-like Distributed Hash Table (DHT) structure. By implementing a tree-based index structure and an effective query routing strategy, ECHO provides low-latency and low-overhead query searches by exploiting the Tabu Search principle. Load balancing is also improved reducing the traditional bottleneck problems arising in upper level nodes of tree-based index structures such as PHT. Furthermore, ECHO copes with DHT churn problems as its index exploits logical information as opposed to static reference cache approaches or replication techniques. The performance evaluation results obtained using PeerSim simulator show that ECHO achieves efficient performance compared other solutions such as the PHT strategy and its optimized version which includes a query cache.

6.3. Consistency protocols

Participants: Marc Shapiro [correspondent], Tyler Crain, Mahsa Najafzadeh, Marek Zawirski, Alejandro Tomsic.

6.3.1. Static Reasoning About Consistency, and associated tools

Large-scale distributed systems often rely on replicated databases that allow a programmer to request different data consistency guarantees for different operations, and thereby control their performance. Using such databases is far from trivial: requesting stronger consistency in too many places may hurt performance, and requesting it in too few places may violate correctness. To help programmers in this task, we propose the first proof rule for establishing that a particular choice of consistency guarantees for various operations on a replicated database is enough to ensure the preservation of a given data integrity invariant. Our rule is modular: it allows reasoning about the behaviour of every operation separately under some assumption on the behaviour of other operations. This leads to simple reasoning, which we have automated in an SMT-based tool. We present a nontrivial proof of soundness of our rule and illustrate its use on several examples.

The intuition was presented at EuroSys 2015 [47]. We present the full theory and proofs in the POPL 2016 paper "'Cause I'm Strong Enough: Reasoning about Consistency Choices in Distributed Systems" [30]. The proof procedure and tool are described in PaPoC 2016 paper "The CISE Tool: Proving Weakly-Consistent Applications Correct" [34] and a YouTube video [48]. It is also the focus of Mahsa Najafzadeh's PhD thesis [3].

6.3.2. Scalable consistency protocols

Developers of cloud-scale applications face a difficult decision of which kind of storage to use, summarised by the CAP theorem. Currently the choice is between classical CP databases, which provide strong guarantees but are slow, expensive, and unavailable under partition; and NoSQL-style AP databases, which are fast and available, but too hard to program against. We present an alternative: Cure provides the highest level of guarantees that remains compatible with availability. These guarantees include: causal consistency (no ordering anomalies), atomicity (consistent multi-key updates), and support for high-level data types (developer friendly API) with safe resolution of concurrent updates (guaranteeing convergence). These guarantees minimise the anomalies caused by parallelism and distribution, thus facilitating the development of applications. This paper presents the protocols for highly available transactions, and an experimental evaluation showing that Cure is able to achieve scalability similar to eventually- consistent NoSQL databases, while providing stronger guarantees.

This work is published under the title "Cure: Strong semantics meets high availability and low latency" at ICDCS 2016 [18].

6.3.3. Lightweight, correct causal consistency

Non-Monotonic Snapshot Isolation (NMSI), a variant of the widely deployed Snapshot Isolation (SI), aims at improving scalability by relaxing snapshots. In contrast to SI, NMSI snapshots are causally consistent, which allows for more parallelism and a reduced abort rate.

This work documents the design of PhysiCS-NMSI, a transactional protocol implementing NMSI in a partitioned data store. It is the first protocol to rely on a single scalar taken from a physical clock for tracking causal dependencies and building causally consistent snapshots. Its commit protocol ensures atomicity and the absence of write-write conflicts. Our PhysiCS-NMSI approach increases concurrency and reduces abort rate and metadata overhead as compared to state-of-art systems.

The paper "PhysiCS-NMSI: efficient consistent snapshots for scalable snapshot isolation" is published at PaPoC 2016 [36].

6.3.4. Reconciling consistency and scalability

Geo-replicated storage systems are at the core of current Internet services. Unfortunately, there exists a fundamental tension between consistency and performance for offering scalable geo-replication. Weakening consistency semantics leads to less coordination and consequently a good user experience, but it may introduce anomalies such as state divergence and invariant violation. In contrast, maintaining stronger consistency precludes anomalies but requires more coordination. This paper discusses two main contributions to address this tension. First, RedBlue Consistency enables blue operations to be fast (and weakly consistent) while the remaining red operations are strongly consistent (and slow). We identify sufficient conditions for determining when operations can be blue or must be red. Second, Explicit Consistency further increases the space of operations that can be fast by restricting the concurrent execution of only the operations that can break application-defined invariants. We further show how to allow operations to complete locally in the common case, by relying on a reservation system that moves coordination off the critical path of operation execution.

The paper "Geo-Replication: Fast If Possible, Consistent If Necessary" is published in the IEEE CS Data Engineering Bulletin of March 2016 [5].

6.3.5. Consistency in 3D

Comparisons of different consistency models often try to place them in a linear strong-to-weak order. However this view is clearly inadequate, since it is well known, for instance, that Snapshot Isolation and Serialisability are incomparable. In the interest of a better understanding, we propose a new classification, along three dimensions, related to: a total order of writes, a causal order of reads, and transactional composition of multiple operations. A model may be stronger than another on one dimension and weaker on another. We believe that this new classification scheme is both scientifically sound and has good explicative value. We presents the three-dimensional design space intuitively.

This work was presented as an invited keynote paper at Concur 2016 [17].

6.3.6. Scalable consistency protocols

Collaborative text editing systems allow users to concurrently edit a shared document, inserting and deleting elements (e.g., characters or lines). There are a number of protocols for collaborative text editing, but so far there has been no precise specification of their desired behavior, and several of these protocols have been shown not to satisfy even basic expectations. This work provides a precise specification of a replicated list object, which models the core functionality of replicated systems for collaborative text editing. We define a strong list specification, which we prove is implemented by an existing protocol, as well as a weak list specification, which admits additional protocol behaviors.

A major factor determining the efficiency and practical feasibility of a collaborative text editing protocol is the space overhead of the metadata that the protocol must maintain to ensure correctness. We show that for a large class of list protocols, implementing either the strong or the weak list specification requires a metadata overhead that is at least linear in the number of elements deleted from the list. The class of protocols to which this lower bound applies includes all list protocols that we are aware of, and we show that one of these protocols almost matches the bound.

This work is published at PODC 2016 [21].

6.3.7. Highly-responsive CRDTs for group editing

Group editing is a crucial feature for many end-user applications. It requires high responsiveness, which can be provided only by optimistic replication algorithms, which come in two classes: classical Operational Transformation (OT), or more recent Conflict-Free Replicated Data Types (CRDTs).

Typically, CRDTs perform better on *downstream* operations, i.e., when merging concurrent operations than OT, because the former have logarithmic complexity and the latter quadratic. However, CRDTs are often less responsive, because their *upstream* complexity is linear. To improve this, this paper proposes to interpose an auxiliary data structure, called the *identifier data structure* in front of the base CRDT. The identifier structure ensures logarithmic complexity and does not require replication or synchronization. Combined with a block-wise storage approach, this approach improves upstream execution time by several orders of magnitude, with negligeable impact on memory occupation, network bandwidth, and downstream execution performance.

This work is published at ACM Group 2016 [27].

6.4. Memory management for multicores

Participants: Antoine Blin, Damien Carver, Maxime Lorrillere, Sébastien Monnet, Julien Sopena [correspondent].

Regal co-advises with Whisper team the PhD of Antoine Blin. The thesis focusses on modern complex embedded systems that involve a mix of real-time and best effort applications. The recent emergence of low-cost multicore processors raises the possibility of running both kinds of applications on a single machine, with virtualization ensuring isolation. Nevertheless, memory contention can introduce other sources of delay, that can lead to missed deadlines. We first investigated the source of memory contention for the Mibench benchmark in a paper published at ETYS 2016 [25]. Then, in a paper published at ECRTS 2016 [24], we present a combined offline/online memory bandwidth monitoring approach. Our approach estimates and limits the impact of the memory contention incurred by the best-effort applications on the execution time of the real-time application. Using our approach, the system designer can limit the overhead on the real-time application to under 5% of its expected execution time, while still enabling progress of the best-effort applications.

Another memory management challenge for multi-cores is the fragmentation induced by the virtualized environments. Previously, we proposed Puma (for Pooling Unused Memory in Virtual Machines) which allows I/O intensive applications running on top of VMs to benefit of large caches. This was realized by providing a remote caching mechanism that provides the ability for any VM to extend its cache using the memory of other VMs located either in the same or in a different host. This work was defended by Maxime Lorrillere in April 2016 [2].

More recently, we study the memory arbitration between containers. In the Damien Carver's PhD thesis (started in October 2015), we are designing ACDC (Advanced Consolidation for Dynamic Containers), a kernel-level mechanisms that automatically provides more memory to the most active containers.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. Joint industrial PhD with Renault

- Renault, 2014-2016, 45 000 euros. The purpose of this contract is to develop solutions for running a mix of real-time and best-effort applications on a small embedded multicore architecture. Our goal is to optimize the usage of the processor resource. The PhD of Antoine Blin is supported by a CIFRE fellowship with Renault.

7.1.2. Joint industrial PhD with Scality SA: CRDTs for Large-Scale Storage Systems

This year, we continued the joint CIFRE (industrial PhD) research of Tao Thanh Vinh, with the French start-up company Scality, as described above (under “Large-Scale File Systems”).

The objective of this research is to design new algorithms for file and block storage systems, considering both the issues of scaling the file naming tree to a very large size, and the issue of conflicting updates to files or to the name tree, in the case of high latency or disconnected work.

A new CIFRE agreement with Scality is awaiting the agreement from ANRT and will start ASAP. The PhD student is Dimitrios Vasilas, and his topic is “Scalable indexing for large-scale distributed storage systems.”

7.1.3. Joint industrial PhDs: data sharing in mobile networks and automatic resizing of shared I/O caches, with Magency

Magency organizes large events during which participants can use mobile devices to access related data and interact together.

The thesis of Lyes Hamidouche concerns efficient data sharing among a large number of mobile devices. Magency brings traces captured during real events (data accesses and user mobility). We are jointly working on the design of algorithms allowing a large number of mobile devices to efficiently access remote data.

Magency also runs servers. A server is used before an event in order to be prepared and tested, and then, during the event to serve the numerous mobile devices accesses. Many servers are run on a single physical machine using containers. Using this configuration, the memory is partitioned, leading to poor performances for applications that need a large amount of memory for caching purpose. In the context of Damien Carver’s PhD thesis, we are designing kernel-level mechanisms that automatically give more memory to the most active containers, leveraging the expertise acquired during Maxime Lorrillere’s PhD thesis.

7.1.4. EMR CREDIT, with Thales

Franck Petit and Swan Dubois participate to the creation of the EMR (Equipe Mixte de Recherche) *CREDIT*, (Compréhension, Représentation et Exploitation Des Interactions Temporelles) between LIP6/UPMC and Thales.

Nowadays, networks are the field of temporal interactions that occur in many settings networks, including security issues. The amount and the speed of such interactions increases everyday. Until recently, the dynamics of these objects was little studied due to the lack of appropriate tools and methods. However, it becomes crucial to understand the dynamics of these interactions. Typically, how can we detect failures or attacks in network traffic, fraud in financial transactions, bugs or attacks traces of software execution. More generally, we seek to identify patterns in the dynamics of interactions. Recently, several different approaches have been proposed to study such interactions. For instance, by merging all interactions taking place over a period (e.g. one day) in a graph that are studied thereafter (evolving graphs). Another approach was to built meta-objects by duplicating entities at each unit of time of their activity, and by connecting them together.

The goal of the EMR is to join both teams of LIP6 and Thales on these issues. More specifically, we hope to make significant progress on security issues such as anomaly detection. This requires the use of a formalism sufficiently expressive to formulate complex temporal properties. Recently, a vast collection of concepts, formalisms, and models has been unified in a framework called Time-Varying Graphs. We want to pursuit that way. In the short run, the challenges facing us are: (1) refine the model to capture some interaction patterns, (2) design of algorithms to separate sequences of interactions, (3) Identify classes of entities playing a particular role in the dynamics, such as bridges between communities, or sources and sinks.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. Labex SMART - (2012–2019)

Members: ISIR (UPMC/CNRS), LIP6 (UPMC/CNRS), LIB (UPMC/INSERM), LJLL (UPMC/CNRS), LTCI (Institut Mines-Télécom/CNRS), CHArt-LUTIN (Univ. Paris 8/EPHE), L2E (UPMC), 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 Regal is responsible of the axe "Autonomic Distributed Environments for Mobility."

The project involves a PhD grant of 100 000 euros over 3 years.

8.1.2. ESTATE - (2016–2020)

Members: LIP6 (Regal, 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 Regal.

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.

8.1.3. RainbowFS - (2016–2020)

Members: LIP6 (Regal, project leader), Scality SA, CNRS-LIG, Télécom Sud-Paris.

Funding: is funded by ANR (PRC) for a total of 919 534 euros, of which 359 554 euros for Regal.

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.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. SyncFree

Title: Large-scale computation without synchronisation

Programm: FP7

Duration: October 2013 – December 2016

Coordinator: Inria

Partners:

Basho Technologies (United Kingdom)

Faculdade de Ciencias E Tecnologiada Universidade Nova de Lisboa (Portugal)

Koc University (Turkey)

Rovio Entertainment OY (Finland)

Trifork AS (Denmark)

Université Catholique de Louvain (Belgium)

Technische Universitaet Kaiserslautern (Germany)

Erlang Solutions Ltd (United Kingdom).

Inria contact: Marc Shapiro

The goal of SyncFree is to enable large-scale distributed applications without global synchronisation, by exploiting the recent concept of Conflict-free Replicated Data Types (CRDTs). CRDTs allow unsynchronised concurrent updates, yet ensure data consistency. This revolutionary approach maximises responsiveness and availability; it enables locating data near its users, in decentralised clouds.

Global-scale applications, such as virtual wallets, advertising platforms, social networks, online games, or collaboration networks, require consistency across distributed data items. As networked users, objects, devices, and sensors proliferate, the consistency issue is increasingly acute for the software industry. Current alternatives are both unsatisfactory: either to rely on synchronisation to ensure strong consistency, or to forfeit synchronisation and consistency altogether with ad-hoc eventual consistency. The former approach does not scale beyond a single data centre and is expensive. The latter is extremely difficult to understand, and remains error-prone, even for highly-skilled programmers.

SyncFree avoids both global synchronisation and the complexities of ad-hoc eventual consistency by leveraging the formal properties of CRDTs. CRDTs are designed so that unsynchronised concurrent updates do not conflict and have well-defined semantics. By combining CRDT objects from a standard library of proven datatypes (counters, sets, graphs, sequences, etc.), large-scale distributed programming is simpler and less error-prone. CRDTs are a practical and cost-effective approach.

The SyncFree project will develop both theoretical and practical understanding of large-scale synchronisation-free programming based on CRDTs. Project results will be new industrial applications, new application architectures, large-scale evaluation of both, programming models and algorithms for large-scale applications, and advanced scientific understanding.

8.2.1.2. 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.

8.3. International Initiatives

8.3.1. Inria International Labs

Inria Chile

Associate Team involved in the International Lab:

8.3.1.1. ARMADA

Title: hARnessing MAssive DAta flows

International Partner (Institution - Laboratory - Researcher):

Universidad Tecnica Federico Santa Maria (Chile) - Department of Computer Science (Department of Comput) - Xavier Bonnaire

Start year: 2014

See also: http://web.inria-armada.org

The ARMADA project aims at designing and implementing a reliable framework for the management and processing of massive dynamic dataflows. The project is two-pronged: fault-tolerant middleware support for processing massive continuous input, and a redundant storage service for mutable data on a massive scale.

8.3.2. Participation in Other International Programs

8.3.2.1. CNRS-Inria-FAP's

Title: Autonomic and Scalable Algorithms for Building Resilient Distributed Systems

International Partner (Institution - Laboratory - Researcher):

Universida de Federal do Paraná (UFPR), Brazil, Prof. Elias Duarte

Duration: 2015–2017

In the context of autonomic computing systems that detect and diagnose problems, self-adapting themselves, the VCube (Virtual Cube), proposed by Prof. Elias Duarte , is a distributed diagnosis algorithm that organizes the system nodes on a virtual hypercube topology. VCube has logarithmic properties: when all nodes are fault-free, processes are virtually connected to form a perfect hypercube; as soon as one or more failures are detected, links are automatically reconnected to remove the faulty nodes and the resulting topology, connecting only fault-free nodes, keeps the logarithmic properties. The goal of this project is to exploit the autonomic and logarithmic properties of the VCube by proposing self-adapting and self-configurable services.

8.3.2.2. Capes-Cofecub

Title: CHOOSING - Cooperation on Hybrid cOmputing clOuds for energy SavING

French Partners: Paris XI (LRI), Regal, LIG, SUPELEC

International Partners (Institution - Laboratory - Researcher):

Universidade de São Paulo - Instituto de Matemática e Estatística - Brazil, Unicamp - Instituto de Computação - Brazil

Duration: 2014–2018

The cloud computing is an important factor for environmentally sustainable development. If, in the one hand, the increasing demand of users drive the creation of large datacenters, in the other hand, cloud computing's "multitenancy" trait allows the reduction of physical hardware and, therefore, the saving of energy. Thus, it is imperative to optimize the energy consumption corresponding to the datacenter's activities. Three elements are crucial on energy consumption of a cloud platform: computation (processing), storage and network infrastructure. Therefore, the aim of this project is to provide different techniques to reduce energy consumption regarding these three elements. Our work mainly focuses on energy saving aspects based on virtualization, i.e., pursuing the idea of the intensive migration of classical storage/processing systems to virtual ones. We will study how different organizations (whose resources are combined as hybrid clouds) can cooperate with each other in order to minimize the energy consumption without the detriment of client requirements or quality of service. Then, we intend to propose efficient algorithmic solutions and design new coordination mechanisms that incentive cloud providers to collaborate.

8.3.2.3. 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.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

Ajoy Kumar Datta

Date: May 2016 - June 2016

Institution: University of Nevada, Las Vegas (USA)

João Barreto

Date: April 2016 - September 2016

Institution: Instituto Superior Técnico, Lisbon, INESC-ID (Portugal)

8.4.1.1. Internships

Alvarez Colombo Santiago Javier

Date: Jul 2015 - Jan 2016

Institution: Universidad de Buenos Aires (Argentina)

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Swan Dubois: Co-Chair of the first Workshop DGDC, collocated with DISC 2016

Franck Petit: General Chair of the 17^{th} International Symposium on Stabilization, Safety, and Security of Distributed Systems (SSS 2016), ed. LNCS, Lyon, France.

9.1.1.2. Member of the Organizing Committees

Swan Dubois: Member of the Organizing Committee and Publicity Chair of the 30^{th} International Symposium on DIStributed Computing (DISC 2016)

9.1.2. Scientific Events Selection

9.1.2.1. Member of Steering Committee

- Marc Shapiro, Steering Committee of OPODIS 2013–2016.

9.1.2.2. Chair of Conference Program Committees

Pierre Sens: Chair of Distributed Systems and Networks track, 28th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD 2016), Los Angeles, USA.

9.1.2.3. Member of the Conference Program Committees

Luciana Arantes: IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS 2016), International Workshop on Recent Advances in the Dependability Assessment of Complex System (Radiance 2016), IEEE International Symposium on Network Computing and Applications (NCA 2016), 28th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD 2016), 7th Latin-American Symposium on Dependable Computing (LADC 2016), 18th International Conferences on High Performance Computing and Communications (HPCC 2016), 11th International Conference on Green, Pervasive and Cloud Computing (GPC-2016), 35th International Conference of the Chilean Computer Science Society (SCCC 2016).

Swan Dubois: 30^{th} International Symposium on DIStributed Computing (DISC 2016), 17^{th} International Symposium on Stabilization, Safety, and Security of Distributed Systems (SSS 2016), 18^{th}*Workshop on Advances in Parallel and Distributed Computational Models* (APDCM 2016), 18^{eme} Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (ALGOTEL 2016)

Franck Petit: 18^{th}*Workshop on Advances in Parallel and Distributed Computational Models* (APDCM 2016)

Pierre Sens: 27th International Symposium on Software Reliability Engineering (ISSRE 2016), 12th European Dependable Computing Conference (EDCC 2016) 30th IEEE International Parallel & Distributed Processing Symposium (IPDPS 2016), IEEE International Symposium on Network Computing and Applications (NCA 2016).

Marc Shapiro: Int. Conf. on Architectural Support for Programming Languages and Systems (ASPLOS 2016), Int. Conf. on Principles of Distributed Systems (OPODIS 2016), Selection committee of Enterprise Research Challenge Contest (ERCC) at Int. Symp. on Stabilization, Safety, and Security of Dist. Sys. (SSS 2016), W. on Planet-Scale Distributed Systems (W-PSDS 2016).

9.1.2.4. Reviewer

Swan Dubois: 20th International Conference on Principles of Distributed Systems (OPODIS 2016), 4th International Conference on Networked Systems (NETYS 2016)

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Franck Petit: co-editor of the proceedings of the 18th International Symposium on Stabilization, Safety, and Security of Distributed Systems, LNCS 10083.

Pierre Sens: Associate editor of International Journal of High Performance Computing and Networking (IJHPCN)

9.1.3.2. Reviewer - Reviewing Activities

Swan Dubois: ACM Transactions on Storage (TOS), SIAM Journal on Computing (SICOMP)

9.1.4. Invited Talks

Pierre Sens. Resource Management in Large Scale Distributed Systems, Dec. 2016, University of Santiago

Marc Shapiro. KTH, Stockholm, Sweden, Nov. 2016. AntidoteDB: a Planet-Scale Available Transactional Database With Strong Semantics

Marc Shapiro. Paris-Descartes University, Nov. 2016. Just-Right Consistency

Nuno Preguiça. Invariant-Preserving Applications for Weakly-consistent Replicated Databases. OPODIS PC workshop, October 2016, Lugano, Switzerland.

Marc Shapiro. Consistency in 3D. OPODIS PC workshop, October 2016, Lugano, Switzerland.

Marc Shapiro. "Consistency in 3D" [17]. Invited talk at Int. Conf. on Concurrency Theory (CONCUR), Québec, Canada, Aug. 2016.

Marc Shapiro. "Just-Right Consistency: Static analysis for minimal synchronisation." Technical talk at Critéo, Paris, France, March 2016.

Marc Shapiro. Keynote Speach at Laboratoire d'Informatique de Grenoble. "Just-Right Consistency: Static analysis for minimal synchronisation." March 2016.

Marc Shapiro, "Reconciling performance and safety in the Antidote widely-geo-replicated database," invited talk at GDR RSD and ASF Winter School 2016, le Pleynet les Sept Laux, France, March 2016.

Marc Shapiro. Invited talk at Inria/EPFL Workshop "Consistency in three dimensions (It's the invariants, stupid)." Rennes, Jan. 2016.

9.1.5. Scientific Expertise

Franck Petit, 2016: Expertise for HCERES

9.1.6. Research Administration

Franck Petit, 2014-2018: Deputy Director of LIP6

Franck Petit, since 2012: Member of the Executive Committee of Labex SMART, Co-Chair (with P. Sens) of Track 4, Autonomic Distributed Environments for Mobility.

Franck Petit, 2016–2020: Coordinator, ANR project ESTATE.

Pierre Sens, since 2016: Member of Section 6 of the national committee for scientific research CoNRS

Pierre Sens, since 2012: Member of the Executive Committee of Labex SMART, Co-Chair (with F. Petit) of Track 4, Autonomic Distributed Environments for Mobility.

Pierre Sens, 2012–2016: Member of the scientific council of UPMC

Pierre Sens, since 2015, officer at scientific research vice presidency UPMC

Pierre Sens, since 2014: Member of Steering Committee of International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD).

Marc Shapiro, since 2009: Member of Scientific Advisory Board of the Center of Informatics and Information Technology (CITI), then Lab. for Informatics and Comp. Sc. (NOVA-LINCS) Universidade Nova de Lisboa, Portugal.

Marc Shapiro, since 2012: Member ERC (European Research Council) Consolidator Grants panel PE6.

Marc Shapiro, since 2013: member of the board of *Société Informatique de France*. Vice-Chair for Research since 2016.

Marc Shapiro, since 2016: Member of Advisory Board of IEEE Computer Society Special Technical Community on Operating Systems (STCOS).

Marc Shapiro, since 2014: Member of Steering Committee of Int. Conf. on Principles of Distributed Systems (OPODIS).

Marc Shapiro, 2016–2020: Coordinator, ANR project RainbowFS.

Marc Shapiro, 2013–2016: Coordinator, FP7 project SyncFree.

Julien Sopena, Member of "Directoire des formations et de l'insertion professionnelle" of UPMC Sorbonne Universités, France

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Julien Sopena is Member of “Directoire des formations et de l’insertion professionnelle” of UPMC Sorbonne Universités, France

Master: Julien Sopena is responsible of Computer Science Master’s degree in Distributed systems and applications (in French, SAR), UPMC Sorbonne Universités, France

Master: Luciana Arantes, Swan Dubois, Sébastien Monnet, Franck Petit, Pierre Sens, Julien Sopena, Advanced distributed algorithms, M2, UPMC Sorbonne Universités, France

Master: Jonathan Lejeune, Designing Large-Scale Distributed Applications, M2, UPMC Sorbonne Universités, France

Master: Maxime Lorrillere, Julien Sopena, Linux Kernel Programming, M1, UPMC Sorbonne Universités, France

Master: Luciana Arantes, Sébastien Monnet, Pierre Sens, Julien Sopena, Operating systems kernel, M1, UPMC Sorbonne Universités, France

Master: Luciana Arantes, Swan Dubois, System distributed Programming, M1, UPMC Sorbonne Universités, France

Master: Luciana Arantes, Swan Dubois, Franck Petit, Distributed Algorithms, M1, UPMC Sorbonne Universités, France

Master: Sébastien Monnet, Julien Sopena, Client-server distributed systems, M1, UPMC Sorbonne Universités, France

Licence: Pierre Sens, Luciana Arantes, Julien Sopena, Principles of operating systems, L3, UPMC Sorbonne Universités, France

Licence: Swan Dubois, Initiation to operating systems, L3, UPMC Sorbonne Universités, France

Licence: Swan Dubois, Franck Petit, Advanced C Programming, L2, UPMC Sorbonne Universités, France

Licence: Swan Dubois, Sébastien Monnet, Introduction to operating systems, L2, UPMC Sorbonne Universités, France

Licence: Mesaac Makpangou, C Programming Language, 27 h, L2, UPMC Sorbonne Universités, France

Ingénieur 4ème année : Marc Shapiro, Introduction aux systèmes d’exploitation, 26 h, M1, Polytech UPMC Sorbonne Universités, France.

9.2.2. Supervision

PhD: Mohamed Hamza Kaaouachi, “Autonomic Distributed Environments for Mobility”, UPMC/Chart-LUTIN (Labex SMART), Franck Petit, Swan Dubois, and Francois Jouen (Chart), 01/12/2016.

PhD: Maxime Lorrillere, “A kernel cooperative cache for virtualized environments”, UPMC, Sébastien Monnet, Julien Sopena, Pierre Sens, 02/04/2016.

PhD: Mahsa Najafzadeh, UPMC, funded by Inria competitive grant (Cordi-S), since Nov. 2012, Marc Shapiro, 22 April 2016.

PhD in Progress: Joao Paulo de Araujo, “L’exécution efficace d’algorithmes distribués dans les réseaux véhiculaires”, funded by CNPq (Brésil), since Nov.2015, Pierre Sens and Luciana Arantes.

PhD in progress : Antoine Blin, “Execution of real-time applications on a small multicore embedded system”, since April 2012, Gilles Muller (Whisper) and Julien Sopena, CIFRE Renault

PhD in progress: Sébastien Bouchard, “Gathering with faulty robots”, UPMC, since Oct. 2016, Swan Dubois, Franck Petit, Yoann Dieudonné (University of Picardy Jules Verne)

PhD in progress: Marjorie Bournat, “Exploration with robots in dynamic networks”, UPMC, since Sep. 2015, Swan Dubois, Franck Petit, Yoann Dieudonné (University of Picardy Jules Verne)

PhD in progress: Damien Carver, “HACHE : HorizontAl Cache cHorEgraphy - Toward automatic resizing of shared I/O caches.”, UPMC, CIFRE, since Jan. 2015, Sébastien Monnet, Pierre Sens, Julien Sopena, Dimitri Refauvelet (Magency).

PhD in Progress: Florent Coriat, “Géolocalisation et routage en situation de crise” since Sept 2014, UPMC, Anne Fladenmuller (NPA-LIP6) and Luciana Arantes.

PhD in progress: Rudyar Cortes,“Un Environnement à grande échelle pour le traitement de flots massifs de données,” UPMC, funded by Chile government, since Sep. 2013, Olivier Marin, Luciana Arantes, Pierre Sens.

PhD in progress: Lyes Hamidouche, “Data replication and data sharing in mobile networks”, UPMC, CIFRE, since Nov. 2014, Sébastien Monnet, Pierre Sens, Dimitri Refauvelet (Magency).

PhD in progress: Denis Jeanneau,“Problèmes d’accord et détecteurs de défaillances dans les réseaux dynamique,” UPMC, funded by Labex Smart, since Oct. 2015, Luciana Arantes, Pierre Sens.

PhD in progress: Yoann Péron, “Development of an adaptive recommendation system”, UPMC/Makazi, since May 2014, Franck Petit, Patrick Gallinari, Matthias Oehler (Makazi).

PhD in progress: Alejandro Z. Tomsic, UPMC, funded by SyncFree, since Feb. 2014, Marc Shapiro.

PhD in progress: Tao Thanh Vinh, UPMC, CIFRE, since Feb. 2014, Marc Shapiro, Vianney Rancurel (Scality).

PhD in progress: Gauthier Voron, “Big-Os : un OS pour les grands volumes de données,”, UPMC, since Sep. 2014, Gaël Thomas, Pierre Sens.

9.2.3. PhD committees

Franck Petit was the reviewer of:

- Mohamed Khaled, PhD UPJV (Advisors: F. Levé and V. Villain)

Franck Petit was Chair of:

- Quentin Bramas, UPMC (Advisor: S. Tixeuil)
- Fadwa Boubekeur, UPMC (Advisor: L. Blin)

Pierre Sens was the reviewer of:

- S. Chiky, HDR ISEP, Paris
- P. Li, PhD Bordeaux (Advisors: R. Namyst, E. Brunnet)
- J. Olivares, PhD Rennes 1 (Advisor: A-M Kermarrec)
- J. Pastor, PhD ENM, Nantes (Advisors: A. Lebre, F. Desprez)
- C. Sauvagnat, PhD IRIT, Toulouse (Advisor: M. Kaaniche)

Pierre Sens was Chair of

- C. Concolato, HDR Telecom Paris
- T. Ecarot, PhD, Telecom SupParis, UPMC (Advisor: D. Zeghlache)

Marc Shapiro was on the following committees:

- Kirill Bogdanov (Advisors: Dejan Kostic, Jeff Maguire), licentiate thesis, KTH, Stockholm, Nov. 2016.
- Brice Nédélec (Advisors: Pascal Molli, Achour Mostefaoui), PhD thesis, Nantes, oct. 2016.

9.3. Popularization

Julien Sopena animated an activity during the Science Festival 2016 at UPMC

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] M. H. KAAOUACHI.*A distributed approach to covering problems in highly dynamic systems*, UPMC - Paris 6 Sorbonne Universités, January 2016, https://tel.archives-ouvertes.fr/tel-01289153.

[2] M. LORRILLERE.*A kernel cooperative cache for virtualized environments*, Université Pierre et Marie Curie - Paris VI, February 2016, https://hal.inria.fr/tel-01273367.

[3] M. NAJAFZADEH.*The Analysis and Co-design of Weakly-Consistent Applications*, Université Pierre et Marie Curie, April 2016, https://hal.inria.fr/tel-01351187.

Articles in International Peer-Reviewed Journal

[4] K. ALTISEN, A. COURNIER, S. DEVISMES, A. DURAND, F. PETIT.*Self-Stabilizing Leader Election in Polynomial Steps*, in "Information and Computation", 2016, http://hal.upmc.fr/hal-01347471.

[5] V. BALEGAS, C. LI, M. NAJAFZADEH, D. PORTO, A. CLEMENT, S. DUARTE, C. FERREIRA, J. GEHRKE, J. LEITÃO, N. PREGUIÇA, R. RODRIGUES, M. SHAPIRO, V. VAFEIADIS.*Geo-Replication: Fast If Possible, Consistent If Necessary*, in "IEEE Data Engineering Bulletin", March 2016, vol. 39, n° 1, 12, https://hal.inria.fr/hal-01350652.

[6] B. BONAKDARPOUR, S. DEVISMES, F. PETIT.*Snap-Stabilizing Committee Coordination*, in "Journal of Parallel and Distributed Computing", January 2016, vol. 87, p. 26-42 [*DOI :* 10.1016/J.JPDC.2015.09.004], http://hal.upmc.fr/hal-01347461.

[7] N. BRAUD-SANTONI, S. DUBOIS, M. H. KAAOUACHI, F. PETIT.*The Next 700 Impossibility Results in Time-Varying Graphs*, in "International Journal of Networking and Computing", 2016, vol. 6, n° 1, p. 27-41, https://hal.inria.fr/hal-01344422.

[8] E. CARON, A. K. DATTA, C. TEDESCHI, F. PETIT.*Self-Stabilizing Prefix Tree Based Overlay Networks*, in "International Journal of Foundations of Computer Science", 2016, vol. 27, n° 5, p. 607–630 [*DOI :* 10.1142/S0129054116500192], http://hal.upmc.fr/hal-01347457.

[9] A. COURNIER, A. K. DATTA, S. DEVISMES, F. PETIT, V. VILLAIN.*The expressive power of snap-stabilization*, in "Theoretical Computer Science", May 2016, vol. 626, p. 40–66 [*DOI :* 10.1016/J.TCS.2016.01.036], http://hal.upmc.fr/hal-01292988.

[10] N. HIDALGO, L. ARANTES, P. SENS, X. BONNAIRE.*ECHO: Efficient Complex Query over DHT overlays*, in "Journal of Parallel and Distributed Computing", February 2016, vol. 88, p. 31-45 [*DOI :* 10.1016/J.JPDC.2015.10.007], https://hal.archives-ouvertes.fr/hal-01340589.

[11] D. JEANNEAU, T. RIEUTORD, L. ARANTES, P. SENS.*Solving k-Set Agreement Using Failure Detectors in Unknown Dynamic Networks*, in "IEEE Transactions on Parallel and Distributed Systems", 2016, https://hal.inria.fr/hal-01354864.

[12] S. F. PILEGGI.*Probabilistic Semantics*, in "Procedia Computer Science", 2016, vol. 80, p. 1834–1845 [*DOI :* 10.1016/J.PROCS.2016.05.472], http://hal.upmc.fr/hal-01327353.

[13] E. ROSAS, N. HIDALGO, V. GIL COSTA, M. MARIN, C. BONACIC, H. SENGER, L. ARANTES, C. MARCONDES, O. MARIN.*Survey on Simulation for Mobile Ad-Hoc Communication for Disaster Scenarios*, in "Journal of Computer Science and Technology", 2016, vol. 31, n^o 2 [*DOI :* 10.1007/S11390-016-1630-X], http://hal.upmc.fr/hal-01354787.

[14] H. SENGER, V. GIL-COSTA, L. ARANTES, C. A. MARCONDES, M. MARIN, L. M. SATO, F. A. B. DA SILVA.*BSP Cost and Scalability Analysis for MapReduce Operations*, in "Concurrency and Computation: Practice and Experience", June 2016, vol. 28, n^o 8, p. 2503-2527 [*DOI :* 10.1002/CPE.3628], https://hal.inria.fr/hal-01254275.

[15] D. SERRANO, S. BOUCHENAK, Y. KOUKI, F. ALVARES DE OLIVEIRA JR., T. LEDOUX, J. LEJEUNE, J. SOPENA, L. ARANTES, P. SENS.*SLA guarantees for cloud services*, in "Future Generation Computer Systems", January 2016, vol. 54, p. 233–246 [*DOI :* 10.1016/J.FUTURE.2015.03.018], https://hal.archives-ouvertes.fr/hal-01162654.

[16] R. C. TURCHETTI, E. P. DUARTE, L. ARANTES, P. SENS.*A QoS-configurable failure detection service for internet applications*, in "Journal of Internet Services and Applications", 2016, vol. 7, n^o 9, 14 [*DOI :* 10.1186/S13174-016-0051-Y], https://hal.inria.fr/hal-01405120.

Invited Conferences

[17] M. SHAPIRO, M. SAEIDA ARDEKANI, G. PETRI.*Consistency in 3D*, in "Int. Conf. on Concurrency Theory (CONCUR) 2016", Québec, Canada, LIPICS (editor), Int. Conf. on Concurrency Theory (CONCUR) 2016, August 2016, vol. CONCUR 2016, 15, https://hal.inria.fr/hal-01350668.

International Conferences with Proceedings

[18] D. D. AKKOORATH, A. TOMSIC, M. BRAVO, Z. LI, T. CRAIN, A. BIENIUSA, N. PREGUIÇA, M. SHAPIRO.*Cure: Strong semantics meets high availability and low latency*, in "Int. Conf. on Distributed Computing Systems (ICDCS 2016)", Nara, Japan, IEEE, June 2016, p. 405-414 [*DOI :* 10.1109/ICDCS.2016.98], https://hal.inria.fr/hal-01350558.

[19] K. ALTISEN, S. DEVISMES, A. DURAND, F. PETIT.*Gradual Stabilization under τ-Dynamics*, in "22nd International European Conference on Parallel and Distributed Computing (Euro-Par 2016)", Grenoble, France, August 2016, http://hal.upmc.fr/hal-01348413.

[20] L. ARANTES, R. FRIEDMAN, O. MARIN, P. SENS.*Probabilistic Byzantine Tolerance Scheduling in Hybrid Cloud Environments*, in "18th International Conference on Distributed Computing and Networking (ICDCN 2017)", Hyderabad, India, January 2017 [*DOI :* 10.1145/1235], https://hal.inria.fr/hal-01399026.

[21] H. ATTIYA, S. BURCKHARDT, A. GOTSMAN, A. MORRISON, H. YANG, M. ZAWIRSKI.*Specification and Complexity of Collaborative Text Editing*, in "Int. Symp. on Principles of Distributed Computing (PODC) 2016", Chicago, IL, United States, ACM (editor), Int. Symp. on Principles of Distributed Computing (PODC) 2016, July 2016, vol. PODC 2016, 10 [*DOI :* 10.1145/2933057.2933090], https://hal.inria.fr/hal-01351512.

[22] M. BOURNAT, A. K. DATTA, S. DUBOIS.*Self-Stabilizing Robots in Highly Dynamic Environments*, in "SSS 2016 - 18th International Symposium Stabilization, Safety, and Security of Distributed Sys-

tems", Lyon, France, Lecture Notes in Computer Science, Springer, November 2016, vol. 10083, p. 54-69 [*DOI :* 10.1007/978-3-319-49259-9_5], https://hal.archives-ouvertes.fr/hal-01416308.

[23] L. BLIN, F. BOUBEKEUR, S. DUBOIS.*Algorithme auto-stabilisant efficace en mémoire pour la construction d'un arbre couvrant de diamètre minimum*, in "ALGOTEL 2016 - 18èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications", Bayonne, France, May 2016, https://hal.archives-ouvertes.fr/hal-01302779.

[24] A. BLIN, C. COURTAUD, J. SOPENA, J. LAWALL, G. MULLER.*Maximizing Parallelism without Exploding Deadlines in a Mixed Criticality Embedded System*, in "28th EUROMICRO Conference on Real-Time Systems (ECRTS'16)", Toulouse, France, July 2016, https://hal.inria.fr/hal-01346979.

[25] A. BLIN, C. COURTAUD, J. SOPENA, J. LAWALL, G. MULLER.*Understanding the Memory Consumption of the MiBench Embedded Benchmark*, in "Netys", Marakech, Morocco, May 2016, https://hal.inria.fr/hal-01349421.

[26] G. BOSILCA, A. BOUTEILLER, A. GUERMOUCHE, T. HÉRAULT, Y. ROBERT, P. SENS, J. DONGARRA.*Failure Detection and Propagation in HPC systems*, in "SC 2016 - The International Conference for High Performance Computing, Networking, Storage and Analysis", Salt Lake City, United States, November 2016, https://hal.inria.fr/hal-01352109.

[27] L. BRIOT, P. URSO, M. SHAPIRO.*High Responsiveness for Group Editing CRDTs*, in "ACM International Conference on Supporting Group Work", Sanibel Island, FL, United States, November 2016 [*DOI :* 10.1145/2957276.2957300], https://hal.inria.fr/hal-01343941.

[28] F. CORIAT, A. FLADENMULLER, L. ARANTES, O. MARIN.*Crowdsourcing-based architecture for post-disaster geolocation: A comparative performance evaluation*, in "The 15th IEEE International Symposium on Network Computing and Applications (NCA 2016)", Cambridge, MA, United States, October 2016, p. 1 - 9 [*DOI :* 10.1109/NCA.2016.7778583], http://hal.upmc.fr/hal-01416297.

[29] R. CORTÉS, X. BONNAIRE, O. MARIN, L. ARANTES, P. SENS.*GeoTrie: A Scalable Architecture for Location-Temporal Range Queries over Massive GeoTagged Data Sets*, in "The 15th IEEE International Symposium on Network Computing and Applications (NCA 2016)", Cambridge, MA, United States, 2016, https://hal.inria.fr/hal-01388949.

[30] A. GOTSMAN, H. YANG, C. FERREIRA, M. NAJAFZADEH, M. SHAPIRO.*'Cause I'm Strong Enough: Reasoning about Consistency Choices in Distributed Systems*, in "Symposium on Principles of Programming Languages", Saint Petersburg, FL, United States, January 2016, p. 371–384 [*DOI :* 10.1145/2837614.2837625], https://hal.inria.fr/hal-01243192.

[31] A. GRACIELA DE MORAES ROSSETTO, C. R. GEYER, L. ARANTES, P. SENS.*Implementing a flexible failure detector that expresses the confidence in the system*, in "LADC 2016 - 7th Latin-American Symposium on Dependable Computing", Cali, Colombia, October 2016, https://hal.inria.fr/hal-01352162.

[32] S. ISKOUNEN, T.-M.-T. NGUYEN, S. MONNET.*WiFi-Direct Simulation for INET in OMNeT++*, in "The 3rd OMNeT++ Community Summit", Brno, Czech Republic, September 2016, http://hal.upmc.fr/hal-01389581.

[33] D. JEANNEAU, L. A. RODRIGUES, E. DUARTE, L. ARANTES.*An Autonomic Hierarchical Reliable Broadcast Protocol for Asynchronous Distributed Systems with Failure Detector*, in "Latin-American Symposium on Dependable Computing (LADC)", Cali, Colombia, 2016, https://hal.inria.fr/hal-01355322.

[34] M. NAJAFZADEH, A. GOTSMAN, H. YANG, C. FERREIRA, M. SHAPIRO.*The CISE Tool: Proving Weakly-Consistent Applications Correct*, in "PaPoC 2016 - 2nd Workshop on the Principles and Practice of Consistency for Distributed Data", Londres, United Kingdom, ACM, April 2016 [*DOI :* 10.1145/2911151.2911160], https://hal.inria.fr/hal-01350636.

[35] L. A. RODRIGUES, L. ARANTES, E. DUARTE.*An Autonomic Majority Quorum System*, in "30th IEEE International Conference on Advanced Information Networking and Applications (AINA)", Crans-Montana, Switzerland, 2016 [*DOI :* 10.1016/J.JPDC.2015.10.007], http://hal.upmc.fr/hal-01349513.

[36] A. TOMSIC, T. CRAIN, M. SHAPIRO.*PhysiCS-NMSI: efficient consistent snapshots for scalable snapshot isolation*, in "Int. W. on Principles and Practice of Consistency for Distr. Data (PaPoC) 2016", London, United Kingdom, ACM (editor), Int. W. on Principles and Practice of Consistency for Distr. Data (PaPoC) 2016, April 2016, vol. PaPoC 2016, 4 [*DOI :* 10.1145/2911151.2911166], https://hal.inria.fr/hal-01350657.

[37] M. ZAWIRSKI, C. BAQUERO, A. BIENIUSA, N. PREGUIÇA, M. SHAPIRO.*Eventually Consistent Register Revisited*, in "Int. W. on Principles and Practice of Consistency for Distributed Data (PaPoC)", London, United Kingdom, ACM (editor), Int. W. on Principles and Practice of Consistency for Distributed Data (PaPoC), ACM Sigops / EuroSys, April 2016, vol. PaPoC 2016, 7 [*DOI :* 10.1145/2911151.2911157], https://hal.inria.fr/hal-01242700.

Scientific Books (or Scientific Book chapters)

[38] A. GRACIELA DE MORAES ROSSETTO, C. ROLIM, V. LEITHARDT, C. R. GEYER, L. ARANTES, P. SENS.*A failure detector based on processes' relevance and the confidence degree in the system for self-healing in ubiquitous environments*, in "Pervasive Computing", C. DOBRE, F. XHAFA (editors), Next Generation Platforms for Intelligent Data Collection, 2016, 23, https://hal.inria.fr/hal-01355320.

Research Reports

[39] D. D. AKKOORATH, A. TOMSIC, M. BRAVO, Z. LI, T. CRAIN, A. BIENIUSA, N. PREGUIÇA, M. SHAPIRO.*Cure: Strong semantics meets high availability and low latency*, Inria ; Paris 6, February 2016, n^{o} RR-8858, https://hal.inria.fr/hal-01270776.

[40] A. BLIN, C. COURTAUD, J. SOPENA, J. LAWALL, G. MULLER.*Maximizing Parallelism without Exploding Deadlines in a Mixed Criticality Embedded System*, Inria, February 2016, n^{o} RR-8838, https://hal.inria.fr/hal-01268078.

[41] M. BOURNAT, A. K. DATTA, S. DUBOIS.*Self-Stabilizing Robots in Highly Dynamic Environments* , LIP6 UMR 7606, Inria, UPMC Sorbonne Universités, France ; University of Nevada, Las Vegas, United States, September 2016, https://hal.inria.fr/hal-01368920.

[42] M. BOURNAT, S. DUBOIS, F. PETIT.*Computability of Perpetual Exploration in Highly Dynamic Rings*, Laboratoire d'informatique de Paris 6 [LIP6] ; UPMC Sorbonne Universités/CNRS/Inria - EPI REGAL, December 2016, https://hal.archives-ouvertes.fr/hal-01417565.

[43] M. NAJAFZADEH, A. GOTSMAN, H. YANG, C. FERREIRA, M. SHAPIRO.*The CISE Tool: Proving Weakly-Consistent Applications Correct*, Inria Paris Rocquencourt, February 2016, n^o RR-8870, https://hal.archives-ouvertes.fr/hal-01279495.

[44] A. G. M. ROSSETTO, L. ARANTES, P. SENS, C. R. GEYER.*Impact: an Unreliable Failure Detector Based on Processes' Relevance and the Confidence Degree in the System*, Université Pierre et Marie Curie ; Inria Paris-Rocquencourt - Regal ; Universidade Federal do Rio Grande do Sul, September 2016, https://hal.inria.fr/hal-01136595.

[45] M. SHAPIRO, M. SAEIDA ARDEKANI, G. PETRI.*Consistency in 3D*, Institut National de la Recherche en Informatique et Automatique (Inria), July 2016, n^o RR-8932, https://hal.archives-ouvertes.fr/hal-01343592.

Other Publications

[46] P. DARCHE.*Évolution des mémoires à semi-conducteurs à accès aléatoire*, February 2017, article E2491 des techniques de l'ingenieur, article de référence sur le domaine, http://hal.upmc.fr/hal-01341972.

References in notes

[47] V. BALEGAS, S. DUARTE, C. FERREIRA, R. RODRIGUES, N. PREGUIÇA, M. NAJAFZADEH, M. SHAPIRO.*Putting Consistency back into Eventual Consistency*, in "Euro. Conf. on Comp. Sys. (EuroSys)", Bordeaux, France, ACM, 2015, p. 6:1–6:16 [*DOI :* 10.1145/2741948.2741972], https://hal.inria.fr/hal-01248191.

[48] M. NAJAFZADEH, M. SHAPIRO.*Demo of the CISE tool*, November 2015, https://youtu.be/HJjWqNDh-GA.

Project-Team REO

Numerical simulation of biological flows

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

IN PARTNERSHIP WITH:
CNRS

Université Pierre et Marie Curie (Paris 6)

RESEARCH CENTER
Paris

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team REO

Creation of the Project-Team: 2005 April 01

Keywords:

Computer Science and Digital Science:

6.1.1. - Continuous Modeling (PDE, ODE)
6.1.4. - Multiscale modeling
6.1.5. - Multiphysics modeling
6.2.1. - Numerical analysis of PDE and ODE
6.3.1. - Inverse problems
6.3.2. - Data assimilation
6.3.4. - Model reduction

Other Research Topics and Application Domains:

2.2.1. - Cardiovascular and respiratory diseases
2.2.3. - Cancer
2.4.1. - Pharmaco kinetics and dynamics

1. Members

Research Scientists

Jean-Frédéric Gerbeau [Team leader, Inria, Senior Researcher, HDR]
Miguel Ángel Fernández Varela [Inria, Senior Researcher, HDR]
Céline Grandmont [Inria, Senior Researcher, HDR]
Damiano Lombardi [Inria, Researcher]
Marc Thiriet [CNRS, Researcher, HDR]
Marina Vidrascu [Inria, Senior Researcher]
Irene Vignon-Clementel [Inria, Researcher, HDR]
Sanjay Pant [Inria, Starting Research position, until Sep 2016]

Faculty Members

Laurent Boudin [UPMC, Associate Professor]
Muriel Boulakia [UPMC, Associate Professor]

Technical Staff

Axel Fourmont [Inria]
Fabien Raphel [Inria]

PhD Students

Matteo Aletti [Inria, granted by REVAMMAD project]
Chloé Audebert [Inria, granted by ANR iFLOW project]
Ludovic Boilevin-Kayl [Inria, from Feb 2016]
Chen-Yu Chiang [UPMC, from Feb 2016]
Nicolas Pozin [ALSI (Air Liquide Sante International), granted by CIFRE]
Alexandre This [Philips, granted by CIFRE]
Eliott Tixier [UPMC]
Mikel Landajuela Larma [Inria, until Mar 2016]

Post-Doctoral Fellows

Dena Kazerani [Inria, from Oct 2016]

Lara Trussardi [Inria, from Oct 2016]
Faisal Amlani [Inria, until Sep 2016]

Visiting Scientists

Stéphano Zonca [Politecnico di Milano, Italy, from Oct to Dec 2016]
Michele Annese [Universita degli Studi di Brescia, Italy, PhD Student, from Mar to Jul 2016]
Rodolfo Araya [University of Concepcion, Chile, Professor, from Apr 2016 to Jul 2016]

Administrative Assistant

Maryse Desnous [Inria]

Others

Valentin Amar [Inria, Internship (Paris 5), from Jul to Sep 2016]
Nathan Gold [Inria, Internship (York University), from May 2016 to Jul 2016]
Gaétan Gossard [Inria, Internship (UPMC), from May to Aug 2016]

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 (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.

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 [64] in the frame of combustion. It was later used to develop the Kiva code [54] at the Los Alamos National Laboratory, or the Hesione code [59], for example. It has a wide range of applications, besides the nuclear setting: diesel engines, rocket engines [57], 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 [60]. 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.

[0] http://www-sop.inria.fr/CardioSense3D/

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 [63], and Biot [55] 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 [56]), 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.

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.* [58], [62], [61] 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/M3DISIM 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

An important industrial partnership has been signed with the start-up companies Kephalios and Epygon, for the mathematical modeling of implantable cardiac devices.

6. New Software and Platforms

6.1. cardioXcomp

KEYWORDS: Cardiac Electrophysiology - Safety Pharmacology
FUNCTIONAL DESCRIPTION

cardioXcomp is a software dedicated to the safety pharmacology industry. It is developed in the framework of the joint laboratory (LabCom) "cardioXcomp" with the software company Notocord. Its purpose is to model the electrical potential of cardiomyocytes measured by a microelectrode array (MEA), and to model the effect of drugs on this signal. It was registered in November 2015 at the Agence pour la Protection des Programmes under the Inter Deposit Digital Number IDDNFR.001.480003.000.S.P.2015.000.31230.

- Participants: Jean-Frédéric Gerbeau, Fabien Raphel, Nejib Zemzemi
- Contact: Jean-Frédéric Gerbeau

6.2. FELiScE

Finite Elements for Life SCiences and Engineering problems
KEYWORDS: Finite element modeling - 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: Dominique Chapelle, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Philippe Moireau, Marina Vidrascu, Sebastien Gilles, Benoit Fabreges, Axel Fourmont, Mikel Landajuela Larma, Damiano Lombardi, Matteo Aletti, Irene Vignon-Clementel and Faisal Amlani
- Contact: Jean-Frédéric Gerbeau
- URL: http://felisce.gforge.inria.fr

6.3. MODULEF

FUNCTIONAL DESCRIPTION

MODULEF is a legacy finite element library developed at Inria since the 1980's. Here, we limit ourselves to recent developments done within this library.

A numerical method to approximate the constitutive laws for rubber elasticity derived from polymer physics are implemented in Modulef.

It is based on algorithms from stochastic geometry to generate suitable polymer networks, Delaunay tessellation algorithms to deal with steric effects (courtesy of the Inria project-team GAMMA2), the introduction of 1-dimensional finite elements for the polymer-chains in Modulef.

- Participants: Marina Vidrascu and Antoine Gloria
- Contact: Marina Vidrascu
- URL: https://www.rocq.inria.fr/modulef/

6.4. SHELDDON

SHELls and structural Dynamics with DOmain decomposition in Nonlinear analysis
FUNCTIONAL DESCRIPTION

SHELDDON is a finite element library based on the Modulef package which contains shell elements, nonlinear procedures and PVM subroutines used in domain decomposition or coupling methods, in particular fluid-structure interaction.

- Participants: Dominique Chapelle, Patrick Le Tallec and Marina Vidrascu
- Contact: Marina Vidrascu
- URL: https://gforge.inria.fr/projects/shelddon/

7. New Results

7.1. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Matteo Aletti, Faisal Amlani, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Mikel Landajuela Larma, Damiano Lombardi, Marina Vidrascu.

In [15] a simplified fluid-structure interaction method is proposed in order to deal with the simulation of fluids in elastic pipes. The motivation of this work is the modeling of the blood flow in arterioles. The structure is modeled by a non-linear Koiter shell, without bending. In addition, the presence of active elastic fibers is considered. The structure is lumped into the boundary condition of the fluid problem leading to a generalized Robin boundary condition. A finite elements discretization is proposed and several numerical test cases are presented to assess the properties of the method.

In [45] a reduced order modeling method is investigated to deal with multi-domain multi-physics problems. In particular we considered the case in which one problem of interest, described by a generic non-linear partial differential equation is coupled to one or several problems described by a set of linear partial differential equations. In order to speed up the resolution of the coupled system, a low-rank representation of the Poincaré-Steklov operator is built by a reduced-basis approach. A database for the secondary problems is built when the interface condition is set to be equal to a subset of the Laplace-Beltrami eigenfunctions on the surface. An online update is also introduced in order to guarantee stability and robustness. Several 3D fluid-fluid and fluid-structure couplings are presented as numerical experiments.

In [44] two new numerical methods for incompressible fluid/thin-walled structure interaction problems using unfitted meshes are proposed. The spatial discretization is based on different variants of Nitsche's method with cut elements. The degree of fluid-solid splitting (semi-implicit or explicit) is given by the order in which the space and time discretizations are performed. For the semi-implicit schemes, energy-based stability and a priori error estimates are derived and which guarantee the unconditional stability and optimal accuracy in the energy-norm of one the methods. Stability and a priori error estimates are also derived for one of the explicit schemes. Numerical experiments in a benchmark illustrate the performance of the different methods proposed.

7.2. Numerical methods for biological flows

Participants: Chloé Audebert, Jean-Frédéric Gerbeau, Céline Grandmont, Sanjay Pant, Marc Thiriet, Irene Vignon-Clementel.

In [16], we present a new approach for the outflow boundary conditions of Navier-Stokes equations in hemodynamics that consists in adding a 3D artificial part where the Navier-Stokes equations are modified to obtain an equivalent energy balance to a standard coupling with a 3-element Windkessel model. We investigate theoretically the stability of the system and compare it to previously introduced methods. We compare these coupling methods for numerical simulations of blood flow in three patient-specific models, which represent different flow regimes in the pulmonary and systemic circulations.

In [36], we highlight and present solutions to several challenges of the UKF method, a data-assimilation method, pertinent to reduced models of cardiovascular haemodynamics. These include methods to a) avoid ill-conditioning of covariance matrix; b) handle a variety of measurement types; c) include a variety of prior knowledge in the method; and d) incorporate measurements acquired at different heart-rates, a common situation in the clinic where patient-state differs between various clinical acquisitions.

In [18], we introduce a kinetic scheme to solve the 1D Euler equations of hemodynamics, which solution on several benchmark tests for both arterial and venous wall laws compares well with the literature. In particular, it is shown that it has a good behavior when the section area of a vessel is close to zero, which is an important property for collapsible or clamped vessels. The application to liver surgery shows that a closed-loop model of the global circulation, including 0D and 1D equations, is able to reproduce the change of waveforms observed after different levels of hepatectomy.

In [17], we explain with a 0D closed-loop lumped model the hemodynamics changes observed during partial hepatectomy in pigs [22]. The typical increase of portal pressure, increase of liver pressure loss, slight decrease of portal flow and major decrease in arterial flow are quantitatively captured by the model for a 75% hepatectomy. The different post-operative states, observed in experiments, are reproduced with the proposed model. Thus, an explanation for inter-subjects post-operative variability is proposed. This work needs to be translated to humans, in which liver flow modulation is a subject of surgery research [39].

In [24], we propose a computational approach for efficient design study of a reducer stent to be percutaneously implanted in enlarged right ventricular outflow tracts (RVOT) of repaired Tetralogy of Fallot. Hemodynamics of different designs are simulated in the stented RVOT via a reduce order model based on proper orthogonal decomposition on a reference device configuration. To validate the approach, forces exerted on the valve and on the reducer are monitored, varying with geometrical parameters, and compared with the results of full CFD simulations.

Peripheral pulmonary artery stenosis (PPS) is a congenital abnormality resulting in pulmonary blood flow disparity and right ventricular hypertension, for which optimal surgical strategies remain unclear. In [38], a proof of concept study, a constant shear stress hypothesis and structured pulmonary trees are used to derive adaptive outflow boundary conditions for 3D-0D postoperative blood flow simulations. This strategy provides better predictions of pulmonary flow distribution than the conventional strategy of maintaining outflow boundary conditions.

In [26] the effect of inserted needle on the subcutaneous interstitial flow is studied. The goal is to describe the physical stress affecting cells during acupuncture needling. The convective Brinkman equations are considered to describe the flow through a fibrous medium. Three-dimensional simulations are carried out by employing an ALE finite element model. Numerical studies illustrate the acute physical stress developed by the implantation of a needle.

In [32], a fully three-dimensional blood flow simulation through a complete rigid macrovascular circuit, namely the intracranial venous network, instead of a reduced order simulation and partial vascular network is presented. The biomechanical modeling step is carefully analyzed and leads to the description of the flow governed by the dimensionless Navier-Stokes equations for an incompressible viscous fluid. The equations are then numerically solved with a free finite element software using five meshes of a realistic geometry obtained from medical images to prove the feasibility of the pipeline. Some features of the intracranial venous circuit in the supine position such as asymmetric behavior in merging regions are discussed.

7.3. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Jean-Frédéric Gerbeau, Damiano Lombardi, Fabien Raphel, Eliott Tixier.

In [51] the variability of phenomena described by parametric partial differential equations is studied. In particular, given population statistics on a system observables, the probability density distribution of the parameters is sought such that the statistics of the model outputs match the observed ones. An uncertainty quantification step is solved once for all by using a non-instrusive approach, and then the inverse problem is solved by introducing an entropy regularisation. Several numerical experiments are considered to validate the approach and compare it to other existing techniques.

In [50] a reduced order modeling method is proposed in order to speed-up the solution of reaction diffusion equations. It is based on the Approximated Lax Pair method, the discretisation is carried out by adopting an empirical interpolation framework in order to deal with non-polynomial nonlinearities. Some numerical examples on the FKPP equations as well as the equations in electrophysiology are proposed.

We published in [25] a discussion about the Comprehensive in vitro Proarrhythmia Assay (CiPA), which is a nonclinical Safety Pharmacology paradigm for discovering electrophysiological mechanisms that are likely to confer proarrhythmic liability to drug candidates intended for human use. In particular, we presented the use of mathematical modeling in Safety Pharmacology to better understand the electric signals acquired by multielectrode arrays.

7.4. Lung and respiration modeling

Participants: Laurent Boudin, Muriel Boulakia, Céline Grandmont, Nicolas Pozin, Irene Vignon-Clementel.

In [46], we proved the existence of global weak solutions to the incompressible Navier-Stokes-Vlasov system in a three-dimensional time-dependent domain with absorption boundary conditions for the kinetic part. This model arises from the study of respiratory aerosol in the human airways. The proof is based on a regularization and approximation strategy designed for our time-dependent framework.

In [52] we develop a lung-ventilation model. The parenchyma is described as an elastic homogenized media, irrigated by the tracheo-bronchial tree, a nonlinear resistive pipe network. Both are strongly coupled, and an efficient algorithm that takes advantage of the tree dyadic structure is proposed. This framework is used with different types of boundary conditions, including a nonlinear Robin model of the surrounding lung structures, to exhibit global and local coupling effects, for various ventilations. The model is also compared to a more classical exit-compartment (0D) approach.

In [34], we present a new framework that is designed to simulate ventilation and particle fate throughout the respiration cycle, both difficult to dynamically image. The flow and the particle transport and deposition models in the main bronchi are coupled to 1D models that account for the distal lobar lung structures. This enables modeling of inspiration as well as expiration. This leads to differentiated particle deposition over time, and between lobes and generations. Strong agreement to previously collected regional rat experimental data is shown, as the 1D models account for lobe-dependent morphology.

7.5. Miscellaneous

Participants: Laurent Boudin, Jean-Frédéric Gerbeau, Damiano Lombardi, Sanjay Pant, Marina Vidrascu, Irene Vignon-Clementel.

In [47], we derive the Maxwell-Stefan formalism from the Boltzmann equation for mixtures with general cross-sections. The derivation uses the Hilbert asymptotic method for systems at low Knudsen and Mach numbers. We also formally prove that the Maxwell-Stefan coefficients can be linked to the direct linearized Boltzmann operator for mixtures. That allows to compute the values of the Maxwell-Stefan diffusion coefficients with explicit and simple formulae with respect to the cross-sections. We also justify the specific ansatz we use thanks to the so-called moment method.

In [19] we give a presentation of the mathematical and numerical treatment of plate dynamics problems including rotational inertia. The presence of rotational inertia in the equation of motion makes the study of such problems interesting. We employ HCT finite elements for space discretization and the Newmark method for time discretization in FreeFEM++, and test such methods in some significant cases: a circular plate clamped all over its lateral surface, a rectangular plate simply supported all over its lateral surface, and an L-shaped clamped plate.

In [31] we investigated a modified k-nearest neighbors method to assess the differential entropy of a probability density distribution given a set of samples. Instead of considering a classical Kozachenko-Leonenko approximation, an improved parametric gaussian representation is proposed. The method aims at improving the performances of the classical estimator when considering the probability density distribution of model observations, which are featured by a strong anisotropy or functional dependency.

In [49] a dynamical adaptive tensor method is proposed to build parsimonius discretisations for systems whose domain can be naturally decomposed as a product of sets. A modified Proper Generalised Decomposition step is introduced, that allows to project the equations residual on a tensorised space. Contrary to the majority of the methods proposed, the tensor rank is adapted to guarantee a chosen precision. The method is applied to the Vlasov-Poisson system of equations. In order to preserve the hamiltonian structure of the problem, a symplectic integrator is proposed. The convergence of the method is proved and several high-dimensional test-cases are presented in order to validate the approach.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Air Liquide Santé International

Participants: Céline Grandmont, Nicolas Pozin, Irene Vignon-Clementel.

CIFRE convention and contract with Air Liquide Santé International in the context of the ANRT on "Multiscale lung ventilation modeling in health and disease", for the PhD thesis of Nicolas Pozin (March 2014 - February 2017).

8.1.2. Philips Research

Participants: Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Alexandre This.

CIFRE convention and contract with Philips Research for the PhD thesis of Alexandre This (January 2016 - December 2018) on fusion data/simulation for the assessment of mitral regurgitation.

8.1.3. Kephalios & Epygon

Participants: Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Ludovic Boilevin-Kayl, Marina Vidrascu.

REO is an academic partner of the industrial project MIVANA, dedicated to the development of new technologies for mitral valve treatment. It is led by the start-up company Kephalios, with the participation of the start-up company Epygon, by the company MDB Texinov and the research institute IFTH. In this framework, REO has two bilateral contracts with Kephalios and Epygon on the modeling and simulation of two medical devices for mitral valve repair.

8.1.4. Instem/NOTOCORD

Participants: Muriel Boulakia, Damiano Lombardi, Jean-Frédéric Gerbeau, Fabien Raphel, Eliott Tixier.

REO partners with the software company NOTOCORD in the framework of the LabCom "cardioXcomp" (see ANR projects section). In 2016, the ANR funding came to an end, and NOTOCORD was acquired by the company Instem. Our collaboration with Instem/NOTOCORD will continue as a bilateral partnership with the purpose of developing the software cardioXcomp dedicated to the safety pharmacology industry.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

9.1.1.1. ANR Project "EXIFSI"

Participants: Faisal Amlani, Miguel Ángel Fernández Varela [Principal Investigator], Axel Fourmont, Mikel Landajuela Larma, Marina Vidrascu.

Period: 2012-2016
The aim of this project, coordinated by Miguel Ángel Fernández Varela, is to study mathematically and numerically new numerical methods for incompressible fluid-structure interaction.

9.1.1.2. ANR LabCom "CARDIOXCOMP"

Participants: Muriel Boulakia, Damiano Lombardi, Jean-Frédéric Gerbeau [Principal Investigator], Fabien Raphel, Eliott Tixier.

Period: 2013-2016.
This project, coordinated by Jean-Frédéric Gerbeau, is carried out in the framework of a joint laboratory ("LabCom" call of ANR) with the software company NOTOCORD. The focus is the mathematical modeling of a device measuring the electrical activity of cardiomyocytes. The overall objective of CardioXcomp is to enrich NOTOCORD's software with modeling and simulation solutions and provide to safety pharmacology research a completely new set incorporating state of the art signal processing and numerical simulation.

9.1.1.3. ANR Project "iFLOW"

Participants: Chloé Audebert, Jean-Frédéric Gerbeau, Irene Vignon-Clementel [co-Principal Investigator].

Period: 2013-2017.
This ANR-TecSan, co-managed by Eric Vibert (Paul Brousse Hospital) and Irene Vignon-Clementel, aims at developing an Intraoperative Fluorescent Liver Optimization Workflow to better understand the relationship between architecture, perfusion and function in hepatectomy.

Other partners: DHU Hepatinov - Hôpital Paul Brousse, Inria Mamba, Fluoptics, IfADo, MID.

9.1.1.4. 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.

9.1.1.5. Participation to other ANR projects

- Laurent Boudin is a member of the ANR Blanc project Kibord on kinetic models in biology and related domains
- Laurent Boudin is a member of the ANR TecSan Oxhelease
- Céline Grandmont is a member of the ANR TecSan Oxhelease
- Marina Vidrascu is a member of the ANR ARAMIS
- Irene Vignon-Clementel is a member of the project iLite (09/16-), RHU-santé grant, a large French hospital-medical research consortium that aims at developing innovations for liver and tissue engineering (Inria PI: Dirk Drasdo).

9.1.2. Inria initiatives

9.1.2.1. ADT Project "MENAMES"

Participants: Miguel Ángel Fernández Varela [Principal Investigator], Axel Fourmont, Marina Vidrascu.

Period: 2014-2016

The aim of this project, coordinated by Miguel Ángel Fernández Varela, is to implement in the FELiScE library the shell elements included in the shelddon and Modulef libraries.

9.1.2.2. ADT Project “PARASOL”

Participants: Miguel Ángel Fernández Varela [Principal Investigator], Axel Fourmont, Marina Vidrascu.

Period: 2016-2017
The aim of this project, coordinated by Miguel Ángel Fernández Varela, is to implement in the FELiScE library several balancing domain decomposition methods (BDD) for solid-mechanics.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. REVAMMAD

Title: "Retinal Vascular Modeling, Measurement and Diagnosis"

Programm: FP7

Duration: April 2013 - March 2017

Coordinator: University of Lincoln

Partners: See the web site http://revammad.blogs.lincoln.ac.uk/partners/

Inria contact: J.-F. Gerbeau

REVAMMAD is a European Union project aimed at combatting some of the EU’s most prevalent chronic medical conditions using retinal imaging. The project aims to train a new generation of interdisciplinary scientists for the academic, clinical and industrial sectors, and to trigger a new wave of biomedical interventions. The role of REO team within this consortium is to propose a mathematical model and a simulation tool for the retina hemodynamics. See http://revammad.blogs.lincoln.ac.uk for more details.

9.2.2. Collaborations in European Programs, Except FP7 & H2020

9.2.2.1. SimInhale COST

Participant: Irene Vignon-Clementel.

Action MP1404, a pan-European network of experts in the field of inhaled medicine

9.3. International Initiatives

9.3.1. Trans-Atlantic Network of Excellence for Cardiovascular Research

Participants: Jean-Frédéric Gerbeau, Sanjay Pant, Irene Vignon-Clementel [correspondant].

Period: 2010-2016

This network, funded by the Leducq foundation, is working on the multi-scale modeling of single ventricle hearts for clinical decision support.

Other partners: see http://modelingventricle.clemson.edu/home.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Visiting Professor: Rodolfo Araya, University of Concepcion (Chile), from Apr 2016 to Jul 2016
- Visiting PhD student: Michele Annese, Universita degli Studi di Brescia (Italy), from Mar to Jul 2016
- Visiting PhD student: Stefano Zonca, Politecnico di Milano (Italy), from Oct to Sep 2016

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

- Matteo Aletti
  - Co-organizer of the monthly Junior Seminar of Inria Paris.
- Laurent Boudin
  - Member of the organizing and scientific committees of the "Recent advances in kinetic equations and applications" workshop, June 2016, Paris
  - Member of the organizing committee of the 5th "Forum Emploi Maths", December 2016, Paris
- Jean-Frédéric Gerbeau
  - Local organizing Committee of the SIAM conference on Parallel Processing 2016. Paris, France.
- Sanjay Pant
  - Organizing committee member, 5th International Conference on Computational and Mathematical Biomedical Engineering (CMBE) 2017
- I. Vignon-Clementel
  - Organized a minisymposium at the COSINE conference, May 25th-26th, Bordeaux, France
  - Organized a minisymposium at the ECCOMAS congress, June 4th-9th, Crete, Greece
  - Programme committee member, Computational and Mathematical Biomedical Engineering Conference
  - Conference steering committee, International Conference on Engineering Frontiers in Pediatric and Congenital Heart Disease, 2015-present

10.1.2. Scientific Events Selection

10.1.2.1. Reviewer

- Jean-Frédéric Gerbeau
  - Member of the Scientific Program Committee of the Millennium Science Initiative, a program of the Ministry of Economy of Chile.
  - Expert for Horizon2020 FET OPEN RIA Call 2015/2.
- Irene Vignon-Clementel
  - Expert for “Appel à projets générique”, ANR 2016.
- Marina Vidrascu

- Expert for FONDECYT - Chile “Projects for Initiation in Research” 2016

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Jean-Frédéric Gerbeau
  - Editor-in-Chief of Mathematical Modelling and Numerical Analysis (M2AN), SMAI/EDP Sciences.
  - Series editor of “SEMA SIMAI Series”, Springer.
  - Member of the editorial board of Journal Advances in Computational Mathematics (ACOM), Springer
  - Member of the editorial board of International Journal for Numerical Methods in Biomedical Engineering (IJNMBE), Wiley.
  - Member of the editorial board of Communications in Applied and Industrial Mathematics, SIMAI/De Gruyter.
  - Member of the editorial board of Journal for Modeling in Ophthalmology, Kugler.
- Marc Thiriet
  - Member of the editorial board of Digital Medicine

10.1.4. Research Administration

- Laurent Boudin
  - Expert evaluator for ANVUR (VQR 2011-2014), Italy
  - Member of the Board of Mathematics Licence (EFU de Licence de mathématiques), UPMC
  - Member of the think-tank for third-year programs in Mathematics at UPMC.
  - Member of the IREM (Institutes for Research on Mathematics Teaching) Scientific Committee.
  - Member of the SMAI (French Society for applied and industrial mathematics) Teaching Committee.
- Muriel Boulakia
  - Supervisor of the teaching of mathematics at the engineer school Polytech Paris-UPMC
- Miguel Ángel Fernández Varela
  - Co-president of the Scientific Positions Commission, Inria Paris
- Jean-Frédéric Gerbeau
  - Service activity at Inria: Délégué Scientifique / Chairman of the project-teams’ committee of Inria Paris research center; Member of the Inria Evaluation Committee.
  - Service activity in other French institutions: member of the scientific committee of Labex NUMEV, Montpellier.
  - Service activity abroad: member of the Reference Committee of the PhD program Mathematical Models and Methods in Engineering (Politecnico di Milano, Italy).
- Céline Grandmont
  - Member of the Evaluation Committee Inria (2015–)
  - Head of the HCERES evaluation Jury of Imath lab. Toulon Univ.
- Marc Thiriet
  - Vice-President & Council Member of the International Society of Digital Medicine
- I. Vignon-Clementel

  - Organizing the monthly seminar at Inria Paris on “modeling and scientific computing”, now joint seminar "Rencontres Inria-LJLL en calcul scientifique" (until June 2016)
  - Committee member for PhD students at Inria "Commission consultative des doctorants", since July 2016.
  - Mediator between PhD students and their supervisors for Inria Paris-Rocquencourt

10.1.5. Conferences

- Matteo Aletti
  - Minisymposium talk, SIMAI2016, Sep 13-16, 2016 Milano, Italy
  - Minisymposium talk, ECCOMAS Congress 2016, Jun 5-10, 2016, Crete, Greece
  - Presentation at REVAMMAD (EU Marie Curie ITN) meeting, Jun 2016, Lincoln, UK
- Rodolfo Araya
  - Seminar, Laboratoire de Mathématiques de Besançon, Université de Franche-Comté, Besançon, May 26
  - Seminar, Groupe de Modélisation Mathématique, Mécanique et Numérique, Université de Caen Basse-Normandie, Jun 6
  - Minisymposium talk, The Mathematics of Finite Elements and Applications 2016 (MAFELAP 2016) conference, Jun 14-17, London, UK
- Chloé Audebert
  - Seminar, Journée interne du Laboratoire Jacques-Louis Lions, Nov 16, 2016, Paris, France.
  - Seminar, BioMécanique et BioIngénierie (BMBI), UTC, Nov 15, 2016, Compiègne, France.
  - Minisymposium talk,Word Congress on Computational Mechanics (WCCM), Jul 24-29, 2016, Seoul, Korea
  - Minisymposium talk, European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS), Jun 5-10, 2016, Crete Island, Greece
  - Open Brain in HPB Surgery, Club Innovation ACHBT, Jun 3-5, 2016, Carnac, France.
  - Congrès National d’Analyse Numérique (CANUM), May 9-13, 2016, Obernai, France
  - Talk, Saint-Antoine hospital, May 3, 2016, Paris, France
- Laurent Boudin
  - Seminar, Applied Mathematics, Department of Mathematics and Informatics, Univ. Novi Sad, Serbia, July 2016
- Muriel Boulakia
  - Workshop ANR IFSMACS, Toulouse, Nov 2016
  - Invited talk, Workshop Carleman estimates, unique continuation, University College of London, Nov 2016
  - Seminar LMAC, Compiègne, Oct 2016
  - Workshop Mathematics and Health, LJLL, UPMC, May 2016
  - Workshop ANR IFSMACS, Paris, Mar 2016
  - Seminar PDE, IECL, Nancy, Feb 2016
- Miguel Ángel Fernández Varela
  - Invited Speaker IWH Symposium on Simulation and Optimization of Extreme Fluids, Oct 2016, Heidelberg, Germany

  - Minisymposium talk, The Mathematics of Finite Elements and Applications 2016 (MAFELAP 2016) conference, Jun 14-17, London, UK
  - Invited Speaker, Workshop on geometrically unfitted finite element methods, Jan 6-8, 2016, London, UK
- Jean-Frédéric Gerbeau
  - Invited lecturer, CISM-ECCOMAS International Summer School (6 hours), June 2016, Udine, Italia.
  - Invited lecturer, “Numerical methods for PDEs”, Institut Henri Poincaré (9 hours), Oct 2016, Paris, France.
  - Invited speaker, Workshop: Mathematical Modeling in Cardiovascular Healthcare, Oct 2016, Emory University, USA
  - Invited speaker, Workshop “Boundary layer and Fluid-Structure Interaction”, Jan 2016, Bordeaux, France.
  - Invited speaker, 2d conference “Mathematical Modelling of Complex Systems”, Dec 2016, Châtenay Malabry, France.
  - Seminar at Collège de France, Pierre-Louis Lions chair, May 2016, Paris, France.
  - Minisymposium talk, European Congress of Mathematics (ECM), July 2016, Berlin, Germany.
  - Minisymposium talk, World Congress of Computational Mechanics (WCCM), July 2016, Seoul, Korea.
  - Minisymposium talk, SIMAI conference, Sep 2016, Milan, Italy.
- Céline Grandmont
  - Invited Speaker IWH Symposium on Simulation and Optimization of Extreme Fluids, Oct 2016, Heidelberg, Germany
  - Seminar, Ecole Centrale, Apr 2016
  - Invited Speaker, Journées Jeunes Edépistes, Mar 2016, Bordeaux
  - Invited Speaker, Boundary Layers and Fluid-Structure Interactions, Jan 2016, Bordeaux
- Mikel Landajuela
  - Seminar, Séminaire d’analyse numérique, Université de Genève, Mar 8, 2016, Geneva, Switzerland
- Damiano Lombardi
  - Invited talk, ALGORITMY 2016, Mar 13-18, 2016, Podbaske, Slovakia
  - Invited talk, SIMAI 2016, Sep 13-16, 2016, Milano, Italy
  - Contributed talk, Workshop on Reduced Order Modeling, Nov 7-10, 2016, Institut Henri Poincaré, Paris
- Sanjay Pant
  - Contributed talk, The 12th World Congress on Computational Mechanics (WCCM XII), Jul 2016, Seoul, Korea
  - Contributed talk, 5th International Conference on Engineering Frontiers in Pediatric and Congenital Heart Disease, Jun 2016 , Orlando, Florida, USA
  - Contributed talk, Computational modeling in healthcare: Making confident predictions in a world of error and uncertainty, Apr 2016, Glasgow, UK
- Nicolas Pozin
  - Minisymposium talk, European Congress on Computational Methods in Applied Sciences and Engineering - ECCOMAS 2016, Jun 5-10, 2016, Creta, Greece

- Marc Thiriet
  - Invited Speaker, 6th Annual Academic Congress of Chinese Society of Digital Medicine and1st International Conference on Digital Medicine & Medical 3D Printing, Jun 17-19, 2016, Nanjing, China
  - Minisymposium talk, 16th International Society for Therapeutic Ultrasound (ISTU), Mar 14-16, 2016, Tel-Aviv, Israel
- Alexandre This
  - Seminar, Inria Paris Junior Seminar, Oct 18, 2016, Paris
- Eliott Tixier
  - Minisymposium talk, SIAM Conference on Uncertainty Quantification, Apr 5-8, 2016, Lausanne, Switzerland
- Irene Vignon-Clementel
  - Seminar, Paul Brousse Hospital, Nov 18th, Villejuif, France
  - Seminar, DKFZ, Nov 15th, Heidelberg, Germany
  - Invited talk, SimInhale workshop, Oct 17th-19th, Prague, Czech Republic
  - Invited talk, GRIC Journées Françaises de Radiologie, Oct 13th, Paris, France
  - Minisymposium talk, CMBBE conference, September 20th-22nd, Tel Aviv, Israel
  - Seminar, Dassault Systems, July 20th, Velizy-Villacoublay, France
  - Minisymposium talk, SIAM Conference on the Life Sciences, July 11th-14th, Boston, USA
  - Presentation for the Chinese Academy of Science, June 29th, Paris, France
  - Invited talk, Inria National Scientific Days, June 20th-22th, Rennes, France
  - Invited talk, International Conference on Engineering Frontiers in Pediatric and Congenital Heart Disease, June 9th-10th, Orlando, USA
  - Minisymposim Keynote, ECCOMAS congress, June 4th-9th, Crete, Greece
  - Minisymposium talk, COSINE conference, May 25th-26th, Bordeaux, France
  - Presentation, Demi-journée Math-Industrie, LJLL-UPMC, May 10th 2016, Paris, France
  - Invited Keynote, Computational modelling in healthcare: Making confident predictions in a world of error and uncertainty (workshop), April 26th 2016, Glasgow, UK
  - Minisymposium talk, UQ SIAM conference, EPFL, April 5th-8th, 2016, Lausanne, Switzerland
  - Invited talk, workshop: towards a unified framework for benchmarking multicellular models and modelling/simulation software, Leipzig University, March 14th-16th, 2016, Leipzig, Germany
  - Podium talk, The 8th International Bio-Fluids Symposium, February 12-14, 2016, CaltechTech, Pasadena, USA
  - Seminar, Department of Mechanical Eng., UC at Berkeley, Feb. 10th, 2016, Berkeley, USA
  - Seminar, HeartFlow company, Feb. 9th, 2016, Mountain View, USA

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence :

- Ludovic Boilevin-Kayl

  - Calculus, 60h, L1, UPMC
- Laurent Boudin
  - Introduction to series for signal theory, 18h, L2, UPMC
  - Shared studies supervision in mathematics licence for approximately 500 students, 48h, L2-L3, UPMC
- Muriel Boulakia
  - Scilab, 35h, L2, UPMC
  - Nonlinear systems and optimization, 35h, L3, Polytech'Paris
  - Hilbertian analysis, 50h, L3, Polytech'Paris
  - Oral tests in numerical analysis, 20h, L3, UPMC
- Miguel Ángel Fernández Varela
  - Analysis and Scientific Computing, 30h, L3, ENPC
- Jean-Frédéric Gerbeau
  - Numerical Analysis and Optimization, 32h, L3, Ecole Polytechnique.
- Céline Grandmont
  - Ordinary differential equations, 24h, L3, UPMC
- Damiano Lombardi
  - Numerical Methods, 48h, L3, Polytech'Paris
- Eliott Tixier
  - Linear algebra, 60h, L2, Polytech'Paris
- Irene Vignon-Clementel
  - Mathematics for biology, 54h, L1, Université de Versailles Saint Quentin
  - Numerical simulations of blood flow, 1h30, as part of the undergraduate "continuum mechanics", AgroParisTech

Master :

- Laurent Boudin
  - Basics for numerical methods, 36h, M1, UPMC
- Muriel Boulakia
  - Preparatory course for teaching admission examination "Agrégation", 15h, M2, UPMC
- Miguel Ángel Fernández Varela
  - Numerical methods for bio-fluids simulation, 9h, M2, Universidade de Vigo, Spain
- Irene Vignon-Clementel
  - Modélisation hémodynamique & simulation numérique comme outil pour la chirurgie, 1h, M2, Université Paris Sud
- Jean-Frédéric Gerbeau
  - Numerical methods in hemodynamics (20h), M2, UPMC / Univ Paris-Sud / Ecole Polytechnique.
  - Seminar for M2 students of the master “Math SV” (1h), M2, Univ Paris-Sud, December, 2015
  - Seminar for M2 students at Ecole des Mines (3h), Paris, February, 2015

10.2.2. Supervision

HdR : Irene Vignon-Clementel, *Blood and air flow multi-scale simulations based on real data*, defended on March 31, 2016

PhD in progress: Chloé Audebert, *Modeling of liver hemodynamics*, since October 2013. Supervisors: J.-F. Gerbeau & I. Vignon-Clementel.

PhD : Francesco Bonaldi, *Modélisation Mathématique et Numérique de Multi-Structures avec couplage Magnéto-Electro-Thermo-Elastique*, defended on July 6, 2016. Supervisors: F. Krasucki & M. Vidrascu

PhD : Mikel Landajuela, *Coupling schemes and unfitted mesh methods for fluid-structure interaction*, defended in March 29, 2016. Supervisor: M.A. Fernández Varela.

PhD in progress: Matteo Aletti, *Multiscale retinal vascular modeling*, since January 2014. Supervisors: J.-F. Gerbeau & D. Lombardi.

PhD in progress: Eliott Tixier, *Stem cells electrophysiology*, since September 2014. Supervisors: J-F. Gerbeau & D. Lombardi.

PhD in progress: Nicolas Pozin, *Multiscale lung ventilation modeling in health and disease*, since March 2014. Supervisors: C. Grandmont & I. Vignon-Clementel.

PhD in progress: Andrea Bondesan, *Kinetic and fluid models, numerical and asymptotic analysis*, since October 2015. Supervisors: L. Boudin, B. Grec & S. Martin.

PhD in progress: Ludovic Boilevin-Kayl, *Modeling of cardiac implantable devices*, since February 2016. Supervisors: J.-F. Gerbeau & M.A. Fernández Varela

PhD in progress: Alexandre This, *Fusion data/simulation for the assessment of mitral regurgitation*, since January 2016. Supervisor: J.-F. Gerbeau

PhD in progress: Chen-Yu Chiang, *Transport on biological systems and some applications*, since February 2016. Supervisor: M. Thiriet

10.2.3. Juries

- Laurent Boudin
  - PhD committee: Alexandra de Cecco, Université Paul Sabatier (referee), Anthony Preux, Université Paris-Saclay
- Muriel Boulakia
  - PhD committee: Andjela Davidovic, Inria Bordeaux Sud-Ouest; Ibtissem Ben Aïcha, Université d'Aix-Marseille
- Miguel Ángel Fernández Varela
  - PhD committee: Moctar Ndiaye, Université Paul Sabatier (president), Davide Baroli, Politecnico di Milano, Simone Brugiapaglia, Politecnico di Milano; Rocco M. Lancellotti, Politecnico di Milano (referee); Paolo Pacciarini, Politecnico di Milano
- Jean-Frédéric Gerbeau
  - PhD committees: Julien Sigüenza, Univ Montpellier (referee). Anna Tagliabue, Politecnico di Milano (referee).
  - Hiring committee: Inria Bordeaux (CR2); Inria Paris (CR2).
- Céline Grandmont
  - Hiring committee: Rennes Univ. (Professor position), Marseille Univ. (Professor position)
  - PhD committee: M. Ndiaye, Université Paul Sabatier (president), B. Polizzi, Univ. de Nice (referee), B. Burtschell, Ecole Polytechnique (referee), P. Jounieaux, UPMC (president)
  - Member of the «Agrégation» Jury in mathematics
- Marc Thiriet
  - PhD committee: M. Haddadi, Université Paris Est–Créteil (referee)

- Marina Vidrascu
  - PhD committee: F. Bonaldi, Université de Montpellier; M Hédi, Tunis El-Manar & UPMC; F. Cheick, Tunis El-Manar & UPMC
- Irene Vignon-Clementel
  - PhD committee: Gabrielle Fournet, CEA & Université Paris-Saclay (referee)

10.3. Popularization

- Céline Grandmont
  - Conference : "Filles et Maths : une équation lumineuse", 60 students secondary school level, Feb 2016
  - Popularization paper with J.-F. Gerbeau : "Maths, médecine et entreprises : des collaborations gagnantes", brochure Maths Société Express, 2016
  - Conference "Métier": Master 1 Maths students, UPMC, Nov 2016
- Irene Vignon-Clementel
  - Telerama, Interview (Richard Senejoux), Mar 10, 2016
  - Presentation, Inauguration of Inria Paris research center in presence of the Minister of Research and presidents of Universities, ANR, EPST, media, etc. Mar 10, 2016, Paris
  - High school conference, Mar 14, 2016, Lycée St François d'Assise, Montigny le Bretonneux

11. 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^o 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^o 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^o 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^o 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^o 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^o 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^o 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^o 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] M. LANDAJUELA LARMA.*Coupling schemes and unfitted mesh methods for fluid-structure interaction*, Université Pierre et Marie Curie - Paris VI, March 2016, https://tel.archives-ouvertes.fr/tel-01366696.

[13] I. VIGNON-CLEMENTEL.*Blood and air flow multi-scale simulations based on real data*, UPMC - Paris 6 Sorbonne Universités, March 2016, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01418167.

Articles in International Peer-Reviewed Journal

[14] F. ALAUZET, B. FABRÈGES, M. A. FERNÁNDEZ, M. LANDAJUELA.*Nitsche-XFEM for the coupling of an incompressible fluid with immersed thin-walled structures*, in "Computer Methods in Applied Mechanics and Engineering", January 2016, vol. 301, p. 300-335 [*DOI :* 10.1016/J.CMA.2015.12.015], https://hal.inria.fr/hal-01149225.

[15] M. ALETTI, J.-F. GERBEAU, D. LOMBARDI.*A simplified fluid-structure model for arterial flow. Application to retinal hemodynamics*, in "Computational Methods in Applied Mechanics and Engineering", 2016, vol. 306, p. 77-94 [*DOI :* 10.1016/J.CMA.2016.03.044], https://hal.archives-ouvertes.fr/hal-01296940.

[16] G. ARBIA, I. VIGNON-CLEMENTEL, T.-Y. HSIA, J.-F. GERBEAU.*Modified Navier-Stokes equations for the outflow boundary conditions in hemodynamics*, in "European Journal of Mechanics - B/Fluids", June 2016 [*DOI :* 10.1016/J.EUROMECHFLU.2016.06.001], https://hal.archives-ouvertes.fr/hal-01328501.

[17] C. AUDEBERT, M. BEKHEIT, P. BUCUR, E. VIBERT, I. VIGNON-CLEMENTEL.*Partial hepatectomy hemodynamics changes: Experimental data explained by closed-loop lumped modeling*, in "Journal of Biomechanics", November 2016 [*DOI :* 10.1016/J.JBIOMECH.2016.11.037], https://hal.archives-ouvertes.fr/hal-01404771.

[18] C. AUDEBERT, P. BUCUR, M. BEKHEIT, E. VIBERT, I. VIGNON-CLEMENTEL, J.-F. GERBEAU.*Kinetic scheme for arterial and venous blood flow, and application to partial hepatectomy modeling*, in "Computer Methods in Applied Mechanics and Engineering", February 2017, vol. 314, p. 102-125 [*DOI :* 10.1016/J.CMA.2016.07.009], https://hal.archives-ouvertes.fr/hal-01347500.

[19] F. BONALDI, G. GEYMONAT, F. KRASUCKI, M. VIDRASCU.*Mathematical and numerical modeling of plate dynamics with rotational inertia*, in "AACE Journal", 2016, p. 1 - 20, https://hal.inria.fr/hal-01413037.

[20] L. BOUDIN, F. SALVARANI.*Opinion dynamics: kinetic modelling with mass media, application to the Scottish independence referendum*, in "Physica A", 2016, vol. 444, p. 448-457 [*DOI :* 10.1016/J.PHYSA.2015.10.014], https://hal.archives-ouvertes.fr/hal-01114091.

[21] 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.

[22] P. BUCUR, M. BEKHEIT, C. AUDEBERT, I. E. VIGNON-CLEMENTEL, E. VIBERT.*Simplified technique for 75% and 90% hepatic resection with hemodynamic monitoring in large white swine model*, in "Journal of Surgical Research", 2017 [*DOI :* 10.1016/J.JSS.2016.09.018], https://hal.archives-ouvertes.fr/hal-01405171.

[23] E. BURMAN, A. ERN, M. A. FERNÁNDEZ.*Fractional-step methods and finite elements with symmetric stabilization for the transient Oseen problem*, in "ESAIM: Mathematical Modelling and Numerical Analysis", April 2016 [*DOI :* 10.1051/M2AN/2016028], https://hal.inria.fr/hal-01218328.

[24] A. CAIAZZO, R. GUIBERT, I. E. VIGNON-CLEMENTEL.*A reduced-order modeling for efficient design study of artificial valve in enlarged ventricular outflow tracts*, in "Computer Methods in Biomechanics and Biomedical Engineering", 2016, vol. 19, n^{o} 12, p. 1314 - 1318 [*DOI :* 10.1080/10255842.2015.1133811], https://hal.inria.fr/hal-01414386.

[25] I. CAVERO, J.-M. GUILLON, V. BALLET, M. CLEMENTS, J.-F. GERBEAU, H. HOLZGREFE.*Comprehensive in vitro Proarrhythmia Assay (CiPA): Pending issues for successful validation and implementation*, in "Journal of Pharmacological and Toxicological Methods", 2016 [*DOI :* 10.1016/J.VASCN.2016.05.012], https://hal.inria.fr/hal-01328481.

[26] Y. DELEUZE, M. THIRIET, T. W. SHEU.*On three-dimensional ALE finite element model for simulating deformed interstitial medium in the presence of a moving needle*, in "Computers and Fluids", 2016 [*DOI :* 10.1016/J.COMPFLUID.2016.08.001], http://hal.upmc.fr/hal-01355051.

[27] M. A. FERNÁNDEZ, J. MULLAERT.*Convergence and error analysis for a class of splitting schemes in incompressible fluid-structure interaction*, in "IMA Journal of Numerical Analysis", October 2016, vol. 36, n^{o} 4, p. 1748-1782 [*DOI :* 10.1093/IMANUM/DRV055], https://hal.inria.fr/hal-01102975.

[28] 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.

[29] 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.

[30] 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", June 2016 [*DOI :* 10.1002/CNM.2813], https://hal.inria.fr/hal-01239931.

[31] 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.

[32] O. MIRAUCOURT, S. SALMON, M. SZOPOS, M. THIRIET.*Blood flow in the cerebral venous system: modeling and simulation*, in "Computer Methods in Biomechanics and Biomedical Engineering", 2016 [*DOI :* 10.1080/10255842.2016.1247833], http://hal.upmc.fr/hal-01384285.

[33] J. M. OAKES, P. HOFEMEIER, I. VIGNON-CLEMENTEL, J. SZNITMAN.*Aerosols in Healthy and Emphysematous In Silico Pulmonary Acinar Rat Models*, in "Journal of Biomechanics", 2016, vol. 49, n^o 11, p. 2213-2220 [*DOI :* 10.1016/J.JBIOMECH.2015.11.026], https://hal.inria.fr/hal-01244458.

[34] J. M. OAKES, S. C. SHADDEN, C. GRANDMONT, I. VIGNON-CLEMENTEL.*Aerosol Transport Throughout Inspiration and Expiration in the Pulmonary Airways*, in "International Journal for Numerical Methods in Biomedical Engineering", 2016, https://hal.inria.fr/hal-01413455.

[35] S. PANT, C. CORSINI, C. BAKER, T.-Y. HSIA, G. PENNATI, I. VIGNON-CLEMENTEL.*Data assimilation and modelling of patient-specific single-ventricle physiology with and without valve regurgitation*, in "Journal of Biomechanics", July 2016, vol. 49, n^o 11, p. 2162–2173 [*DOI :* 10.1016/J.JBIOMECH.2015.11.030], https://hal.inria.fr/hal-01240146.

[36] S. PANT, C. CORSINI, C. BAKER, T.-Y. HSIA, G. PENNATI, I. VIGNON-CLEMENTEL.*Inverse problems in reduced order models of cardiovascular haemodynamics: aspects of data-assimilation and heart-rate variability*, in "Journal of the Royal Society Interface", 2017, https://hal.inria.fr/hal-01413446.

[37] E. SCHENONE, A. COLLIN, J.-F. GERBEAU.*Numerical simulation of electrocardiograms for full cardiac cycles in healthy and pathological conditions*, in "International Journal for Numerical Methods in Biomedical Engineering", May 2016, vol. 32, n^o 5, In press (in International Journal for Numerical Methods in Biomedical Engineering) [*DOI :* 10.1002/CNM.2744], https://hal.inria.fr/hal-01184744.

[38] W. YANG, J. A. FEINSTEIN, I. E. VIGNON-CLEMENTEL.*Adaptive outflow boundary conditions improve post-operative predictions after repair of peripheral pulmonary artery stenosis*, in "Biomechanics and Modeling in Mechanobiology", 2016, vol. 15, n^o 5, p. 1345-1353 [*DOI :* 10.1007/S10237-016-0766-5], https://hal.inria.fr/hal-01414295.

Articles in National Peer-Reviewed Journal

[39] E. VIBERT, I. VIGNON-CLEMENTEL, D. SAMUEL.*Prévention de l'insuffisance hépatique post-opératoire par modulation pneumatique de la veine porte*, in "Hépato-Gastro & Oncologie Digestive", January 2016, vol. 23, n^{o} 1, p. 90-4 [*DOI :* 10.1684/HPG.2015.1250], https://hal.inria.fr/hal-01414577.

International Conferences with Proceedings

[40] Y. YIN, O. SEDLACZEK, J. LOTZ, J. OLESCH, K. BREUHAHN, D. DRASDO, I. VIGNON-CLEMENTEL.*Tumor Microvasculature in Lung Cancer and Diffusion-Weighted MRI: Preliminary Results*, in "IEEE Nuclear Science Symposuim & Medical Imaging Conference", Strasbourg, France, October 2016, https://hal.inria.fr/hal-01421152.

[41] Y. YIN, O. SEDLACZEK, A. WARTH, M. GONZÁLEZ-VALLINAS, K. BREUHAHN, I. VIGNON-CLEMENTEL, D. DRASDO.*Quantitative Estimation of Tumor Cellularity Based on Histology Data*, in "IEEE Nuclear Science Symposuim & Medical Imaging Conference, Oct 2016", Strasbourg, France, October 2016, https://hal.inria.fr/hal-01421163.

Scientific Books (or Scientific Book chapters)

[42] Y. DELEUZE, M. THIRIET, T. W. SHEU.*On Three-Dimensional ALE Finite Element Model For Simulating Interstitial Medium Deformation in the Presence of a Moving Needle*, in "Advances in Computational Fluid-Structure Interaction and Flow Simulation", Y. BAZILEVS, K. TAKIZAWA (editors), Birkhäuser Basel, October 2016 [*DOI :* 10.1007/978-3-319-40827-9_27], http://hal.upmc.fr/hal-01240292.

[43] M. THIRIET, W.-H. SHEU, A. GARON.*Biofluid Flow and Heat Transfer*, in "Handbook of Fluid Dynamics", R. W. JOHNSON (editor), CRC Press, 2016, 2nd Edition [*DOI :* 10.1201/B19031-35], http://hal.upmc.fr/hal-01393604.

Research Reports

[44] M. A. FERNÁNDEZ, M. LANDAJUELA.*Unfitted mesh formulations and splitting schemes for incompressible fluid/thin-walled structure interaction*, Inria, April 2016, n^{o} RR-8908, https://hal.inria.fr/hal-01309462.

Other Publications

[45] M. ALETTI, D. LOMBARDI.*A Reduced Order representation of the Poincaré-Steklov operator: an application to coupled multi-physics problems*, November 2016, working paper or preprint, https://hal-auf.archives-ouvertes.fr/hal-01396286.

[46] L. BOUDIN, C. GRANDMONT, A. MOUSSA.*Global existence of solutions to the incompressible Navier-Stokes-Vlasov equations in a time-dependent domain*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01312262.

[47] L. BOUDIN, B. GREC, V. PAVAN.*The Maxwell-Stefan diffusion limit for a kinetic model of mixtures with general cross sections*, April 2016, working paper or preprint, https://hal.inria.fr/hal-01303312.

[48] J. BOUYSSIER, M. BENDAHMANNE, Y. COUDIÈRE, J.-F. GERBEAU, J. PEDRON, P. ZITOUN, N. ZEMZEMI.*Parameters estimation approach for the MEA/hiPSC-CM assays*, September 2016, Journées scientifiques du LIRYC, Poster, https://hal.inria.fr/hal-01409683.

[49] V. EHRLACHER, D. LOMBARDI.*A dynamical adaptive tensor method for the Vlasov-Poisson system*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01335507.

[50] J.-F. GERBEAU, D. LOMBARDI, E. SCHENONE.*Approximated lax pairs and empirical interpolation for nonlinear parabolic partial differential equations* , February 2016, working paper or preprint, https://hal.inria.fr/hal-01278778.

[51] J.-F. GERBEAU, D. LOMBARDI, E. TIXIER.*A moment-matching method to study the variability of phenomena described by partial differential equations*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01391254.

[52] N. POZIN, S. MONTESANTOS, I. KATZ, M. PICHELIN, I. E. VIGNON-CLEMENTEL, C. GRANDMONT.*A tree-parenchyma coupled model for lung ventilation simulation*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01407055.

[53] 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*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01421398.

References in notes

[54] 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.

[55] 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.

[56] 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].

[57] 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.

[58] 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.

[59] R. MOTTE.*A numerical method for solving particle-fluid equations*, in "Trends in numerical and physical modeling for industrial multiphase flows", Cargèse, France, 2000.

[60] 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.

[61] F. SACHSE.*Computational Cardiology: Modeling of Anatomy, Electrophysiology, and Mechanics*, Springer-Verlag, 2004.

[62] J. SUNDNES, G. LINES, X. CAI, B. NIELSEN, K.-A. MARDAL, A. TVEITO.*Computing the electrical activity in the heart*, Springer-Verlag, 2006.

[63] K. TERZAGHI.*Theoretical Soil Mechanics*, John Wiley and Sons, New-York, 1943.

[64] 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:

5.1. - Human-Computer Interaction
5.3. - Image processing and analysis
5.4. - Computer vision
5.9. - Signal processing
5.10. - Robotics
5.11. - Smart spaces
6.1. - Mathematical Modeling
6.4. - Automatic control
8.2. - Machine learning
8.5. - Robotics

Other Research Topics and Application Domains:

5.6. - Robotic systems
6.5. - Information systems
6.6. - Embedded systems
7. - Transport and logistics

1. Members

Research Scientists

Fawzi Nashashibi [Team leader, Inria, Senior Researcher, HDR]
Guy Fayolle [Inria, Senior Researcher (emeritus)]
Jean-Marc Lasgouttes [Inria, Researcher]
Gérard Le Lann [Inria, Senior Researcher (emeritus)]
Vicente Milanés [Inria, Starting Research position, until Aug 2016]
Anne Verroust-Blondet [Inria, Researcher, HDR]

Technical Staff

Azary Abboud [Inria, Expert Engineer, from Feb 2016]
Raoul de Charette [Inria, Expert Engineer]
Pierre Merdrignac [Inria, Expert Engineer, until Sep 2016]
Oyunchimeg Shagdar [Inria, Expert Engineer, part time until Aug 2016]
Ahmed Soua [Inria, Expert Engineer, from Apr 2016]
Thomas Streubel [Inria, Expert Engineer, from Feb 2016]
Armand Yvet [Inria]

PhD Students

Mohammad Abualhoul [Inria, Mines ParisTech]
Zayed Alsayed [VEDECOM, Télécom ParisTech]
Pierre de Beaucorps [Inria granted by Valéo Etudes Electroniques, UPMC]
Carlos Flores [Inria, Mines ParisTech]
Fernando Garrido Carpio [VEDECOM, Mines ParisTech]
David González Bautista [Inria, Mines ParisTech]

Mohamed Maddouri [Inria, Univ. Rouen, until Jan 2016]
Francisco Navas [Inria, Mines ParisTech]
Dinh-Van Nguyen [grant from the Ministry of National Education of Vietnam, Mines ParisTech]
Danut-Ovidiu Pop [Inria, Univ. Rouen, from Jun 2016]
Guillaume Trehard [Inria granted by Valéo Etudes Electroniques, Mines ParisTech, until Feb 2016]

Visiting Scientists

Aidos Ibrayev [Kazakh National University, PhD Student, until Jun 2016]
Sakriani Watiasri Sakti [NAIST Japan, Assistant professor, until Mar 2016]

Administrative Assistant

Chantal Chazelas [Inria]

Others

Benjamin Boualam Murano [Min. de l'Education Nationale, student, until Apr 2016]
Rafael Colmenares Prieto [Inria, Engineering student, from Sep 2016]
Thierry Ernst [Yogoko, Research Engineer]
Aitor Gomez Torres [Inria, internship, from Sep 2016]
Juan Jose Larez Urdaneta [Inria, internship, from Aug 2016]
Pablo Marin Plaza [Inria, internship, from Sep 2016]
Kenneth Martinez Torres [Inria, internship, from Jun 2016 until Aug 2016]
Alexis Meyer [Inria, internship, from Mar 2016 until Sep 2016]
Daniel Sanchez Aranguren [Inria, internship, from Aug 2016]
Jose Emilio Traver Becerra [Inria, internship, until Jan 2016]
Myriam Vaca Recalde [Inria, internship, until Sep 2016]
Alfredo Valle Barrio [Inria, internship, from Sep 2016]
Itheri Yahiaoui [Univ. Reims, Assistant Professor]

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.

[0] LaRA is a Joint Research Unit (JRU) associating three French research teams: Inria's project-team RITS, Mines ParisTech's CAOR and LIVIC.

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.

3. Research Program

3.1. Vehicle guidance and autonomous navigation

Participants: Zayed Alsayed, Pierre de Beaucorps, Raoul de Charette, Rafael Colmenares Prieto, Aitor Gomez Torres, Fernando Garrido Carpio, David González Bautista, Pierre Merdrignac, Alexis Meyer, Vicente Milanés, Francisco Navas, Fawzi Nashashibi, Carlos Flores, Dinh-Van Nguyen, Danut-Ovidiu Pop, Oyunchimeg Shagdar, Thomas Streubel, Guillaume Trehard, Anne Verroust-Blondet, 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: Zayed Alsayed, Raoul de Charette, Rafael Colmenares Prieto, Aitor Gomez Torres, Pierre Merdrignac, Alexis Meyer, Fawzi Nashashibi, Dinh-Van Nguyen, Danut-Ovidiu Pop, Guillaume Trehard, 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.

3.1.2. Cooperative Multi-sensor data fusion

Participants: Pierre Merdrignac, Fawzi Nashashibi, Oyunchimeg Shagdar.

Since data are noisy, inaccurate and can also be unreliable or unsynchronized, the use of data fusion techniques is required in order to provide the most accurate situation assessment as possible to perform the perception task. RITS team worked a lot on this problem in the past, but is now focusing on collaborative perception approach. Indeed, the use of vehicle-to-vehicle or vehicle-to-infrastructure communications allows an improved on-board reasoning since the decision is made based on an extended perception.

As a direct consequence of the electronics broadly used for vehicular applications, communication technologies are now being adopted as well. In order to limit injuries and to share safety information, research in driving assistance system is now orientating toward the cooperative domain. Advanced Driver Assistance System (ADAS) and Cybercars applications are moving towards vehicle-infrastructure cooperation. In such scenario, information from vehicle based sensors, roadside based sensors and a priori knowledge is generally combined thanks to wireless communications to build a probabilistic spatio-temporal model of the environment. Depending on the accuracy of such model, very useful applications from driver warning to fully autonomous driving can be performed.

The Collaborative Perception Framework (CPF) is a combined hardware/software approach that permits to see remote information as its own information. Using this approach, a communicant entity can see another remote entity software objects as if it was local, and a sensor object, can see sensor data of others entities as its own sensor data. Last year we developed the basic hardware modules that ensure the well functioning of the embedded architecture including perception sensors, communication devices and processing tools.

Finally, since vehicle localization (ground vehicles) is an important task for intelligent vehicle systems, vehicle cooperation may bring benefits for this task. A new cooperative multi-vehicle localization method using split covariance intersection filter was developed during the year 2012, as well as a cooperative GPS data sharing method.

In the first method, each vehicle estimates its own position using a SLAM (Simultaneous Localization And Mapping) approach. In parallel, it estimates a decomposed group state, which is shared with neighboring vehicles; the estimate of the decomposed group state is updated with both the sensor data of the ego-vehicle and the estimates sent from other vehicles; the covariance intersection filter which yields consistent estimates even facing unknown degree of inter-estimate correlation has been used for data fusion.

In the second GPS data sharing method, a new collaborative localization method is proposed. On the assumption that the distance between two communicative vehicles can be calculated with a good precision, cooperative vehicle are considered as additional satellites into the user position calculation by using iterative methods. In order to limit divergence, some filtering process is proposed: Interacting Multiple Model (IMM) is used to guarantee a greater robustness in the user position estimation.

Accidents between vehicles and pedestrians (including cyclists) often result in fatality or at least serious injury for pedestrians, showing the need of technology to protect vulnerable road users. Vehicles are now equipped with many sensors in order to model their environment, to localize themselves, detect and classify obstacles, etc. They are also equipped with communication devices in order to share the information with other road users and the environment. The goal of this work is to develop a cooperative perception and communication system, which merges information coming from the communications device and obstacle detection module to improve the pedestrian detection, tracking, and hazard alarming.

Pedestrian detection is performed by using a perception architecture made of two sensors: a laser scanner and a CCD camera. The laser scanner provides a first hypothesis on the presence of a pedestrian-like obstacle while the camera performs the real classification of the obstacle in order to identify the pedestrian(s). This is a learning-based technique exploiting adaptive boosting (AdaBoost). Several classifiers were tested and learned in order to determine the best compromise between the nature and the number of classifiers and the accuracy of the classification.

3.1.3. Planning and executing vehicle actions

Participants: Fernando Garrido Carpio, David González Bautista, Vicente Milanés, Fawzi Nashashibi, Francisco Navas, Carlos Flores.

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). Today, RITS is also investigating the opportunity of fuzzy-based control for specific maneuvers. First results have been recently obtained for reference trajectories following in roundabouts and normal straight roads.

3.2. V2V and V2I Communications for ITS

Participants: Thierry Ernst, Oyunchimeg Shagdar, Gérard Le Lann, Pierre Merdrignac, Mohammad Abualhoul, Fawzi Nashashibi.

Wireless communications are expected to play an important role for road safety, road efficiency, and comfort of road users. Road safety applications often require highly responsive and reliable information exchange between neighboring vehicles in any road density condition. Because the performance of the existing radio communications technology largely degrades with the increase of the node density, the challenge of designing wireless communications for safety applications is enabling reliable communications in highly dense scenarios. Targeting this issue, RITS has been working on medium access control design and visible light communications, especially for highly dense scenarios. The works have been carried out considering the vehicle behavior such as vehicle merging and vehicle platooning.

Unlike many of the road safety applications, the applications regarding road efficiency and comfort of road users, on the other hand, often require connectivity to the Internet. Based on our expertise in both Internet-based communications in the mobility context and in ITS, we are now investigating the use of IPv6 (Internet Protocol version 6 which is going to replace the current version, IPv4, in a few years from now) for vehicular communications, in a combined architecture allowing both V2V and V2I.

The wireless channel and the topology dynamics need to be studied when understanding the dynamics and designing efficient communications mechanisms. Targeting this issue, we 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 geocast message delivery, and access point selection.

Below follows a more detailed description of the related research issues.

3.2.1. *Geographic multicast addressing and routing*

Participants: Oyunchimeg Shagdar, Thierry Ernst.

Many ITS applications such as fleet management require multicast data delivery. Existing work on this subject tackles mainly the problems of IP multicasting inside the Internet or geocasting in the VANETs. To enable Internet-based multicast services for VANETs, we introduced a framework that:

i) defines a distributed and efficient geographic multicast auto-addressing mechanism to ensure vehicular multicast group reachability through the infrastructure network,
ii) introduces a simplified approach that locally manages the group membership and distributes the packets among them to allow simple and efficient data delivery.

3.2.2. Platooning control using visible light communications

Participants: Mohammad Abualhoul, Oyunchimeg Shagdar, Fawzi Nashashibi.

The main purpose of our research is to propose and test new successful supportive communication technology, which can provide stable and reliable communication between vehicles, especially for the platooning scenario. Although VLC technology has a short history in comparison with other communication technologies, the infrastructure availability and the presence of the congestion in wireless communication channels lead to propose VLC technology as a reliable and supportive technology which can takeoff some loads of the wireless radio communication. The first objective of this work is to develop an analytical model of VLC to understand its characteristics and limitations. The second objective is to design vehicle platooning control using VLC. In platooning control, a cooperation between control and communication is strongly required in order to guarantee the platoon's stability (e.g. string stability problem). For this purpose we work on VLC model platooning scenario, to permit for each vehicle the trajectory tracking of the vehicle ahead, altogether with a prescribed inter-vehicle distance and considering all the VLC channel model limitations. The integrated channel model of the main Simulink platooning model will be responsible for deciding the availability of the Line-of-Sight for different trajectory's curvatures, which means the capability of using light communication between each couple of vehicles in the platooning queue. At the same time the model will compute all the required parameters acquired from each vehicle controller.

3.2.3. V2X radio communications for road safety applications

Participants: Mohammad Abualhoul, Pierre Merdrignac, Oyunchimeg Shagdar, Fawzi Nashashibi.

While 5.9 GHz radio frequency band is dedicated to ITS applications, the channel and network behaviors in mobile scenarios are not very well known. In this work we theoretically and experimentally study the radio channel characteristics in vehicular networks, especially the radio quality and bandwidth availability. Based on our study, we develop mechanisms for efficient and reliable V2X communications, channel allocation, congestion control, and access point selection, which are especially dedicated to road safety and autonomous driving applications.

3.2.4. Safety-critical communications in intelligent vehicular networks

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 [48]. 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 [46], 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 [50];
- build macroscopic variables that measure the overall state of the underlying graph, in order to improve the local propagation of information [47];
- make the underlying model as sparse as possible, in order to improve BP convergence and quality [49].

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 [43] 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 [45]):

$$Q(x,y)\pi(x,y) = q(x,y)\pi(x) + \widetilde{q}(x,y)\widetilde{\pi}(y) + \pi_0(x,y). \tag{2}$$

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 [44], 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

  $$F(x,y,z) = \sum_{i,j,k \geq 0} f(i,j,k) x^i y^j z^k \tag{3}$$

  satisfies the functional equation

  $$K(x,y)F(x,y,z) = c(x)F(x,0,z) + \widetilde{c}(y)F(0,y,z) + c_0(x,y), \tag{4}$$

  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. Discrete-event simulation for urban mobility

We have developed two simulation tools to study and evaluate the performance of different transportation modes covering an entire urban area.

- one for collective taxis, a public transportation system with a service quality provided will be comparable with that of conventional taxis (system operating with or without reservations, door-to-door services, well adapted itineraries following the current demand, controlling detours and waits, etc.), and with fares set at rates affordable by almost everyone, simply by utilizing previously wasted vehicle capacity;
- the second for 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.

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 (i.e. in La Rochelle). 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. Today, in order to rely on a professional and maintained solution, we have chosen to migrate to the RTMaps SDK development platform. Today, all our developments and demonstrations are using this efficient prototyping platform. Thanks to RTMaps we have been able to do all the demonstrations on our cybercars: cycabs, Yamaha AGV and new Cybus platforms. 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. DOLAR

FUNCTIONAL DESCRIPTION

This software performs real-time obstacle detection and tracking using laser data scanned with one or several laser sensors with different geometric configurations. Obstacle detection is based on laser data segmentation while obstacle tracking uses PHD-based filtering techniques.

- Authors: Raoul de Charette, Fawzi Nashashibi and Evangeline Pollard
- Contact: Fawzi Nashashibi

5.2. FEMOT

Fuzzy Embedded MOTor
FUNCTIONAL DESCRIPTION

FEMOT is an experimental motor for implementing fuzzy logic controllers, including all the fuzzy stages (fuzzification, inference, and defuzzification). This library has been compiled in Microsoft Visual (MVS) Studio and RTMaps. The proposed library is modular and adaptable to different situations and scenarios, especially for autonomous driving applications. FEMOT allows the development of the fuzzy rules to be written as sentences in an almost natural language. It allows the user to define variables and their fuzzy rules and to join them with other variables in rules to yield crisp signals for the controllers.

This software is used for the arbitration and control for fully automated functions. The behaviour of a human driver can be emulated with this technique. First simulations are showing promising results, and the library allows an easy adaptation in decision marking situations.

- Participants: Joshue Perez Rastelli and Vicente Milanés
- Contact: Fawzi Nashashibi

5.3. MELOSYM

FUNCTIONAL DESCRIPTION

MELOSYM is the latest laser based Hierarchical ML-SLAM algorithm developed by RITS. It contains all the functions needed to perform the vehicle localization and the mapping of the environment. Windows compatible, it was initially developed under the RTMAPS platform but the version includes a standalone version.

- Participants: Fawzi Nashashibi, Benjamin Lefaudeux, Jianping Xie and Paulo Lopes Resende
- Contact: Fawzi Nashashibi

5.4. PML-SLAM

- Participants: Zayed Alsayed and Fawzi Nashashibi
- Contact: Fawzi Nashashibi

5.5. Platools

KEYWORD: Telecommunications

- Participant: Marios Makassikis
- Contact: Thierry Ernst

5.6. SODA

SOftwares for Driving Automation
KEYWORD: Environment perception
FUNCTIONAL DESCRIPTION

This software has been developed in the context of the French ABV (Automatisation Basse Vitesse) project. This package contains the functions that are necessary to automate the vehicle navigation in its secured lane.

- Participants: Paulo Lopes Resende and Fawzi Nashashibi
- Contact: Fawzi Nashashibi

5.7. STEREOLOC-3D

FUNCTIONAL DESCRIPTION

STEREOLOC is the package performing stereovision based localization and mapping. It performs semi-dense mapping of outdoor large environments and provides real-time estimates of the vehicle position.

- Participants: Benjamin Lefaudeux and Fawzi Nashashibi
- Contact: Fawzi Nashashibi

5.8. Taxi-col

KEYWORD: Mobile Computing, Transportation

- Participant: Eugenie Lioris
- Contact: Fawzi Nashashibi

5.9. 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

5.10. SimConVA

Participants: Pierre Merdrignac, Oyunchimeg Shagdar, Jean-Marc Lasgouttes.

This software was developed during the SINETIC FUI project. It aims at interfacing the network simulator ns-3 and the prototyping software RTMaps.

The source code of is software is a library to generate an RTMaps component that triggers and controls ns-3. The component handle emission and reception of data packets between RTMaps and ns-3 for every vehicle. It can also deal with the mobility of vehicles in ns-3 based on the localization known in ns-3.

This software was used with the communication stack developed in RITS. It has been shown during the SINETIC project how this it can be used for simulating and emulating cooperative driver assistance systems. Particularly, the software has been tested on cooperative platoons. The tests were conducted on both simulation and real platforms to prove how the such software can be used during the development phase and that it is fully compatible with the architecture already present in the experimental vehicles.

- Contact: Fawzi Nashashibi

6. New Results

6.1. Low Speed Vehicle Localization using WiFi-FingerPrinting

Participants: Dinh-Van Nguyen, Myriam Vaca Recalde, Fawzi Nashashibi.

Recently, the problem of fully autonomous navigation of vehicle has gained major interest from research institutes and private companies. In general, these researches rely on GPS in fusion with other sensors to track vehicle in outdoor environment. However, as indoor environment such as car park is also an important scenario for vehicle navigation, the lack of GPS poses a serious problem. In [39] we present an approach to use WiFi Fingerprinting as a replacement for GPS information in order to allow seamlessly transition of localization architecture from outdoor to indoor environment. Often, movement speed of vehicle in indoor environment is low (10-12km/h) in comparison to outdoor scene but still surpasses human walking speed (3-5km/h, which is usually maximum movement speed for effective WiFi localization). We propose an ensemble classification method together with a motion model in order to deal with the above issue. Experiments show that proposed method is capable of imitating GPS behavior on vehicle tracking.

6.2. Free navigation space estimation

Participants: Raoul de Charette, Rafael Colmenares Prieto, Alexis Meyer, Fawzi Nashashibi.

Autonomous vehicles need to know where they can physically drive. In the past, lane detection was used to bound the driving area of the vehicle but road markings do not exist in many urban scenario thus perception needs to estimate the free navigation space with other means.

To contrast with the state of the art two approaches were developed and will be published soon. The first approach is using a monocular setup and use an absurd logic to identify the flow of the scene and extract the ego motion. The second method still under research is to develop a hybrid approach to segment the navigation space using energy minimization to label the scene assuming learning on the go.

6.3. Pedestrian Recognition using Convolutional Neural Networks

Participants: Danut-Ovidiu Pop, Fawzi Nashashibi.

Pedestrian detection is of highly importance for a large number of applications, especially in the elds of automotive safety, robotics and surveillance. In spite of the widely varying methods developed in recent years, pedestrian detection is still an open challenge whose accuracy and robustness has to be improved. This year we focused on the improvement of the classification component in the pedestrian detection task by adopting two approaches: 1) by combining three image modalities (intensity, depth and ow) to feed a unique convolutional neural network (CNN) and 2) by fusing the results of three independent CNNs. The evaluations have been performed on the Daimler stereo vision data set.

6.4. Reliability estimation and information redundancy for accurate localization

Participants: Zayed Alsayed, Anne Verroust-Blondet, Fawzi Nashashibi.

Our goal is to improve localization systems performances in order to be able to navigate in urban and peri-urban environments. For this purpose, we choose to study the reliability of a SLAM method that incrementally builds a map of the surrounding environment from an information given by a set of 2D laser points.
This year, we focused on SLAM failure and non-failure scenarios.

- Experimental data acquired on the VEDECOM demonstrator in the context of ITS Bordeaux demonstrations in 2015 were analyzed. This evaluation showed in [30] that the SLAM concept seems better suited to urban scenarios, while algorithms such as lane marking detection could offer a good alternative in peri-urban environments.
- In parallel, we worked on designing a reliability measure associated to the pose given by our SLAM considering the geometrical configurations of the 2D laser points describing the environment and the computations done in the maximum likelihood matching process.

6.5. Feature Selection for road obstacles classification

Participants: Itheri Yahiaoui, Pierre Merdrignac, Anne Verroust-Blondet.

In order to ensure the ability of an automated vehicle to be autonomous in a real environment, we must equip it with tools (hardware and software) to meet the requirement of such an application as safety, real-time processing, understanding and intelligence, etc. To contribute to these objectives a perception system is of vital importance. The one on road obstacles detection and classification is of particular interest for us. In this work a large number of geometric features have been proposed to describe different class objects like vehicles, pedestrians, cyclists and static obstacles from 2D laser points. A binary classification was performed with an Adaboost algorithm. In order to improve this work and enhance the classification rate, we have constructed new binary and multiclass classifiers, using SVM and logistic regression, with optimal choices of kernel parameters and models. We have defined several decision strategies by tracking objects in the video sequences, which lead to obtain the most probable target object. On the other hand, we have studied different dependence measures between the proposed features and the classes, leading the selection of the best set of features. As measures of dependence, we have used nonparametric estimate of mutual information, Fisher information and Pearson correlation. We have used also the Akaïke criterion in order to select the best models (the best subset of features) in logistic regression.

6.6. Motion planning techniques

Participants: David González Bautista, Fernando Garrido Carpio, Vicente Milanés, Fawzi Nashashibi, Myriam Vaca Recalde, Jose Emilio Traver Becerra.

The latest developments in the Intelligent Transportation Systems (ITS) field allow emerging technologies to show promising results at increasing passengers comfort and safety, while decreasing energy consumption, emissions and travel time. Despite of great efforts, fully automated driving still remains unsolved, where research challenges such as navigation in urban dynamic environments with obstacle avoidance capabilities–i.e. Vulnerable Road Users (VRU) and vehicles–and cooperative maneuvers among automated and semi-automated vehicles, still need further efforts for a real environment implementation. A deep state-of-the-art review has been conducted to find the gaps in this important topic into the autonomous vehicle field, with special attention to overtaking and obstacle avoidance maneuvers [21].

Having this in mind, a novel local path planning algorithm combining both off-line and real-time generation has been proposed in [32], providing a significant reduction on the computational time with respect to prior implementations from RITS team. This new local planning architecture for urban environments benefits from *a priori* knowledge of the geometry of the road layout, vehicle's kinematics and dynamics, among others, to produce local smooth path for the vehicle to navigate. The planner relies on several databases containing optimized trajectories for a G^2 continuous path generation. Four different type of databases have been generated to provide our system with a naturalistic driving style, allowing the car to maintain smooth trajectories according to the characteristics of the road [33].

Based on the accuracy of current digital maps, it is possible to know before-hand the way-points that define the route by which the vehicle will pass to reach a predefined destination. Furthermore, the original route can be generated in real-time and modified on-demand according to the user needs through the use of Automatic Global Planners (AGP) [42]. That way, since urban scenarios can present several consecutive curves in a short period of time, a smoother and more comfortable path generation can be done by extending the planning horizon up to two curves. There, a set of paths are analyzed by considering the angles of the curves and the distances to them in order to find the best joint point for the consecutive curves.

In this sense, a speed planning algorithm has also been designed to increase passenger comfort and set continuous speed profiles [35]. The approach permits to improve the comfort in automated vehicles by integrating the speed profile with the previously computed path, constraining the global acceleration in the whole ride (longitudinal and lateral accelerations according to ISO 2631-1). It also minimizes distance error problems by associating the speed profile w.r.t. distance in the path instead of the time. The planner has been tested against other techniques in the state-of-the-art, providing better results.

The proposed architecture has been validated both on simulation (with Pro-Sivic and RTMaps) and on the Inria Rocquencourt terrain. The results showed a smoother tracking of the curves, reduction on the execution times and reduced global accelerations increasing comfort. Future works will improve the capacity to deal with dynamic obstacles, conducting avoidance maneuvers if possible, or returning to the original lane if not. The maneuver will be decided by building an occupancy grid with the information given by the perception system. It will provide the best point near the obstacle to carry out the avoidance trajectory by loading the pre-computed curves.

6.7. Plug&Play control for highly non-linear systems: Stability analysis of autonomous vehicles

Participants: Francisco Navas, Vicente Milanés, Fawzi Nashashibi.

The final stage for automating a vehicle relies on the control algorithms. They are in charge of providing the proper behavior and performance to the vehicle, leading to provide fully automated capabilities. Controllability and stability of dynamic complex systems are the key aspects when it comes to design intelligent control algorithms for vehicles.

Nowadays, the problem is that control systems are "monolithic". That means that a minor change in the system could require the entire redesign of the control system. It addresses a major challenge, a system able to adapt the control structure automatically when a change occurred.

An autonomous vehicle is built by combining a set-of-sensors and actuators together with sophisticated algorithms. Since sensors and actuators are prone to intermittent faults, the use of different sensors is better and more cost effective than duplicating the same sensor type. The problem is to deal with the different availability of each sensor/actuator and how the vehicle should react to these changes. A methodology that improves the security of autonomous driving systems by providing a framework managing different sensor/actuator setups should be carried out. New trends are proposing intelligent algorithms able to handle any unexpected circumstances as unpredicted uncertainties or even fully outages from sensors. This is the case of Plug&Play control, which is able to provide stability responses for autonomous vehicles under uncontrolled circumstances, including modifications on the input/output sensors.

In order to meet with the idea of automatically handling those changes into the system, different research lines should be followed:

- Reconfiguration of existing controllers whenever changes are introduced in the system being controlled. In that line, the already commercially available Adaptive Cruise Controller (ACC) system, and its evolution by adding vehicle-to-vehicle communication (CACC) are examined. Plug&Play control is used for providing stable transitions between both controllers when the vehicle-to-vehicle communication link is changing from available to available or vice versa. More detail can be found in [38]. Gain scheduling approaches can be achieved by using the same structure. An Advanced-CACC is developed by using it. Hybrid behaviors between controllers with different head times are carried out depending on the traffic situation.
- Online closed loop identification of the vehicle and its components. Plug&Play control also provides a way for doing online closed loop identification of any system as open loop like systems. Here, the obtained models for the vehicle will be compared with the physical lateral model (Bicycle and 2GDL) and the longitudinal model together with the tire models (Pacejka, Dugoff and Buckhardt). It is also possible to identify new sensors or actuators connected to the system.
- Automatic control reconfiguration to achieve optimal performance together with identification of the new situation. Once a new situation has been identified in the system, the controller should be reconfigured to achieve the optimal performance of the autonomous vehicle.

6.8. Using Fractional Calculus for Cooperative Car Following Control

Participants: Carlos Flores, Vicente Milanés, Fawzi Nashashibi.

In the field of Advanced Driver Assistance Systems (ADAS), there are two main types of systems: passive and active ones. Specifically the active ADAS, they are capable of taking partial or complete control of the vehicle. Among these techniques, Car-Following has arisen as one important solution to traffic jams, driver comfort and safety.

Scoping on the evolution of the control involved in Car Following, it can be remarked the improved version of the cruise control system, Adaptive Cruise Control (ACC). This system allows the vehicle to maintain a desired distance gap measured by raging sensors (LiDAR, radars, etc), by controlling longitudinally the vehicle through the throttle and brake.

Afterward, the addition of Vehicle to Vehicle (V2V) communication links allowed the vehicles to maintain even shorter distances between each of the string members, by performing a Cooperative ACC (CACC). Focusing on CACC formations, a control structure must be conceived to guarantee stability and string stability as well. As a core of the control structure, the controller must be able to maintain the vehicle in the desired spacing in a stable, robust and comfortable.

Towards achieving this goals, it is proposed to use fractional order calculus to gain a more flexible frequency response and at the same time satisfy more demanding design requirements. This mathematical has been used for years for different applications providing good results and outperforming classical techniques in the industrial control field, due to its capability of describing systems more accurately than integer order calculus. Several research lines are stated to achieve these objectives:

- An exhaustive identification process of the experimental platforms dynamics. Allowing further comparison between the empirical identified dynamics of the real vehicle and a theoretical mathematical dynamic model. Such permits to design much more effective and stable control algorithms for both the lateral and longitudinal command of the vehicle.
- Conception of a Car-Following gap regulation controller using fractional order calculus, which has been proven that yields a more accurate description of real processes. The controller should satisfy more demanding design requirements [31], allowing to extend the scope of Car Following controllers' design. This controller should be framed into an appropriate control structure both for ACC and CACC
- Further investigation on the effects of communication delays and latency in the V2V links, as well as study different control structures that react not with the preceding vehicle's behavior but also other string members.

6.9. Decision making for automated vehicles in urban environments

Participants: Pierre de Beaucorps, Thomas Streubel, Anne Verroust-Blondet, Fawzi Nashashibi.

The development of automated vehicles in urban environments requires a robust sensing system followed by an adaptive situation assessment. This is the basis for smart decision making in the driving process without collisions or taking high risks. We address this aspect of automated driving in a project with the sensor developer VALEO. The focus is on complex urban traffic scenarios, e.g. intersections and roundabouts, including multiple road users.

In a first step, we developed a new multi-agent driving simulation as a tool to explore human behavior in relevant traffic scenarios. We conducted a study with 10 test persons driving in a scene with one dummy car to acquire data and understand the human decision process in risky situations. This data was used to retrieving speed profiles for the trajectory planning. The path planning was established with Bezier curves. Further, a robust decision making algorithm utilizes the trajectory planning coupled with a risk assessment. The latter is estimating the post-encroachment time (PET), which is the time between one vehicle leaving a collision zone in an intersection area and the other car entering this same zone. Based on this estimation a risk is assigned to every predetermined speed profile and the one with lowest acceptable risk is chosen to be send to the controller of the automated vehicle. The results showed better performance than the drivers in our study. The so equipped automated vehicle is integrated in our simulation environment and was presented to our project partners in several intersection and roundabout scenarios with a real driver in the same scene.

6.10. Transposition of autonomous vehicle architecture

Participants: Raoul de Charette, Pablo Marin Plaza, Fawzi Nashashibi.

With the development of autonomous vehicles, many software and hardware architectures exist in the world to handle perception, control, decision, planning. Studies were conducted to see how an alien software architecture could be transposed to our Cycabs platforms. Lightweight Communications and Marshalling has been implemented on our platforms to communicate fully with the Carlos 3 architecture, allowing the alien software pipeline to control fully our vehicle. Results and studies include stability of the communication, impact on the control quality, and planning comparison.

6.11. Fusion of Perception and V2P Communication Systems for Safety of Vulnerable Road Users

Participants: Pierre Merdrignac, Oyunchimeg Shagdar, Fawzi Nashashibi.

With cooperative intelligent transportation systems (C-ITS), vulnerable road users (VRU) safety can be enhanced by multiple means.

On one hand, perception systems are based on embedded sensors to protect VRUs. However, such systems may fail due to the sensors' visibility conditions and imprecision. On the other hand, Vehicle-to-Pedestrian (V2P) communication can contribute to the VRU safety by allowing vehicles and pedestrians to exchange information. This solution is, however, largely affected by the reliability of the exchanged information, which most generally is the GPS data. Since perception and communication have complementary features, we can expect that a fusion between these two approaches can be a solution to the VRU safety.

In this work, we proposed a cooperative system that combines the outputs of communication and perception. After introducing theoretical models of both individual approaches, we developed a probabilistic association between perception and V2P communication information by means of multi-hypothesis tracking (MHT).

Experimental studies were conducted to demonstrate the applicability of this approach in real-world environments. Our results showed that the cooperative VRU protection system can benefit of the redundancy coming from the perception and communication technologies both in line-of-sight (LOS) and non-LOS (NLOS) conditions. We established that the performances of this system are influenced by the classification performances of the perception system and by the accuracy of the GPS positioning transmitted by the communication system.

More detail can be fund in [24]

6.12. Study and Evaluation of Laser-based Perception and Light Communication for a Platoon of Autonomous Vehicles

Participants: Mohammad Abualhoul, Pierre Merdrignac, Oyunchimeg Shagdar, Fawzi Nashashibi.

Visible Light Communication (VLC) is a new emerging technology that is being proposed as a reliable and supportive choice for short range communications in ITS.

On the same context, Laser Range Finders (LRF) sensors are used for the vehicular environment perception. Compared to VLC, LRF can provide more coverage range and extended viewing angle.

To take the full advantages of both technologies features, we have studied and demonstrated the proposal of using VLC for information exchange among the platoon members and LRF for inter-vehicle distance estimation. A handover algorithm was proposed to manage the switching process for any failure occurrence by assessing LRF and VLC performance using three different metrics: LRF confidence value, vehicles angular orientation, and the VLC link latency.

The evaluation of the proposed system is verified using VLC prototype and Pro-SiVIC Simulator driving platoon of two autonomous vehicles over different curvature scenarios. Our results showed that the proposed combination are extending the VLC limitations and satisfying the platooning requirement. However, in the very sharp curvature, LRF was capable of driving the platoon except for the 90° curve scenario, the system experienced non-stable behavior due to the LRF area of interest limitation.

More detail can be fund in [27].

6.13. Solutions for Safety-Critical Communications in IVNs

Participant: Gérard Le Lann.

In 2016, we have followed a divide-and-conquer approach. Rather than considering medium-range omnidirectional communications, we have split the problem space in two sub-domains, longitudinal short-range SC communications and lateral short-range communications. Our research has been directed at MAC protocols, string-wide message dissemination based on longitudinal communications, and distributed agreement algorithms based on longitudinal and lateral communications. New results are:

- A rigorous characterization of what is meant by SC communications in IVNs: the space-time bounds acceptability (STBA) requirements, as follows:
  - $STBA_1$: a MAC protocol is acceptable if and only if the distance traveled in λ time units by any vehicle involved in a SC scenario is an order of magnitude smaller than average vehicle size.
  - $STBA_2$: a string-wide message dissemination algorithm, or a string-wide distributed agreement algorithm, is acceptable if and only if the distance traveled in Δ time units by any vehicle involved in a SC scenario is smaller than average vehicle size.
- Specification of SWIFT (Synchronous Wireless Interference-Free Transmissions), a collision-free MAC protocol that solves the BCAD and the TBMA problems introduced in [48] (no solutions given in this publication), and that also achieves fast string-wide acknowledged message dissemination,
- Analytical formulae of worst-case upper bounds λ and Δ achieved with SWIFT [36],
- Specification of Fast Distributed Agreement (FastDA), a problem that arises in IVNs in the presence of conflicting concurrent SC events (e.g., lane changes and brutal braking), under two instances, single-lane (longitudinal) agreement and multilane (lateral and longitudinal) agreement [37],
- Specifications of solutions to FastDA: the Eligo algorithm for the single-lane string-wide agreement (SLA), and the LHandshake protocol for the multilane agreement (MLA),
- Analytical formulae of worst-case upper bounds Δ_{SLA} and Δ_{MLA} achieved with Eligo and LHandshake, respectively [37],
- Verification that SWIFT, Eligo and LHandshake meet the STBA requirements.

It turns out that SWIFT, Eligo, and LHandshake outperform existing stochastic solutions.

6.14. Large scale simulation interfacing

Participants: Ahmed Soua, Jean-Marc Lasgouttes, Oyunchimeg Shagdar.

In order to efficiently design and validate a cooperative intelligent transportation system, a complete simulation environment handling both mobility and communication is required. We are interested here in a so-called system-level view, focusing on simulating all the components of the system (vehicle, infrastructure, management center, etc.) and its realities (roads, traffic conditions, risk of accidents, etc.). The objective is to validate the reference scenarios that take place on a geographic area where a large number of vehicles exchange messages using 802.11p protocol. This simulation tool is be done by coupling the SUMO microscopic simulator and the ns-3 network simulator thanks to the simulation platform iTETRIS.

We have focused in this part of the project on how to reduce the execution time of large scale simulations. To this end, we designed a new simulation technique called Restricted Simulation Zone which consists on defining a set of vehicles responsible of sending the message and an area of interest around them in which the vehicles receive the packets. In fact, the messages emitted by the vehicles located outside the interference zone are not useful for the simulation of the ego-vehicle, and therefore limiting the transmission area to a useful one reduces obviously the number of nodes involved in the transmission operation and thus reduces the processing time of messages. To corroborate the efficiency of our proposal, we compare it with an already existing simulation tool called COLOMBO. The simulation results have shown that our technique outperforms COLOMBO in terms of simulation execution time in the case of large scale simulations (when the number of vehicles exceeds 2400 nodes).

6.15. Belief propagation inference for traffic prediction

Participant: Jean-Marc Lasgouttes.

This work [50], in collaboration with Cyril Furtlehner (TAO, Inria), deals with real-time prediction of traffic conditions in a setting where the only available information is floating car data (FCD) sent by probe vehicles. The main focus is on finding a good way to encode some coarse information (typically whether traffic on a segment is fluid or congested), and to decode it in the form of real-time traffic reconstruction and prediction. Our approach relies in particular on the belief propagation algorithm.

The work about the theoretical aspects of encoding real valued variables into a binary Ising model has now been published [23].

Moreover, following an agreement signed with the city of Vienna (Austria) and the company SISTeMA ITS (Italy), we obtained access to large amounts of data. We are now working on assessing the performance of our techniques in real-world city networks.

6.16. Random Walks in Orthants

Participant: Guy Fayolle.

The Second Edition of the Book [45] *Random walks in the Quarter Plane*, prepared in collaboration with R. Iasnogorodski (St-Petersburg, Russia) and V. Malyshev (MGU, Moscow), is complete and now in the Springer Production Department. It will be published in the collection *Probability Theory and Stochastic Processes*. **Part II** of this second edition borrows specific case-studies from queuing theory, and enumerative combinatorics. Five chapters have been added, including examples and applications of the general theory to enumerative combinatorics. Among them:

- Explicit criterion for the finiteness of the group, both in the genus 0 and genus 1 cases.
- Chapter *Coupled-Queues* shows the first example of a queuing system analyzed by reduction to a BVP in the complex plane.
- Chapter *Joining the shorter-queue* analyzes a famous model, where maximal homogeneity conditions do not hold, hence leading to a system of functional equations.
- Chapter *Counting Lattice Walks* concerns the so-called *enumerative combinatorics*. When counting random walks with small steps, the nature (rational, algebraic or holonomic) of the generating functions can be found and a precise classification is given for the basic (up to symmetries) 79 possible walks.

6.17. Facing ADAS validation complexity with usage oriented testing

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 maybe 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. The COVADEC project (type FUI/FEDER 15) aims to provide methods and techniques to deal with these problems. The test cases automatic generation relies on a *Model Based Testing (MBT)* approach. The tool used for MBT is the software MaTeLo (Markov Test Logic), developed by the company All4Tec. MaTeLo is an MBT tool, which makes it possible to build a model of the expected behaviour of the system under test and then to generate, from this model, a set of test cases suitable for particular needs. MaTeLo is based on Markov chains, and, for non-deterministic generation of test cases, uses the Monte Carlo methods. To cope with the inherent combinatorial explosion, we couple the graph generated by MaTeLo to an ad hoc *random scan Gibbs sampler (RSGS)*, which converges at geometric speed to the target distribution. Thanks to these test acceleration techniques, MaTeLo also makes it possible to obtain a maximal coverage of system validation by using a minimum number of test cases. As a consequence, the number of driving kilometers needed to validate an ADAS is reduced, see [40], [41].

6.18. Broadcast Transmission Networks with Buffering

Participant: Guy Fayolle.

In collaboration with P. Muhlethaler, we analyzed the so-called back-off technique of the IEEE 802.11 protocol in broadcast mode with waiting queues. In contrast to existing models, packets arriving when a station (or node) is in back-off state are not discarded, but are stored in a buffer of infinite capacity. As in previous studies, the key point of our analysis hinges on the assumption that the time on the channel is viewed as a random succession of transmission slots (whose duration corresponds to the length of a packet) and mini-slots during which the back-off of the station is decremented. These events occur independently, with given probabilities. The state of a node is represented by a two-dimensional Markov chain in discrete-time, formed by the back-off counter and the number of packets at the station. Two models are proposed both of which are shown to cope reasonably well with the physical principles of the protocol. Stability (ergodicity) conditions are obtained and interpreted in terms of maximum throughput. Several approximations related to these models are also discussed in [20].

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 and the post-doc of Thomas Streubel under the framework of VALEO project “Daring”
- *SMART* project: on the *Design and development of multisensor fusion system for road vehicles detection and tracking*. This project funds the internship of Alfredo Valle Bario.
- A CIFRE like PhD thesis is ongoing between VALEO and Inria (M. 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 CAMPUS project (PIA french project) 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.

TATA Motors European Technical Centre (TMETC): a new partnership was born in 2016 with the aim of developing a highly automated vehicle. Technology transfer from Inria to TATA Motors includes localization and mapping as well as control-command codes.

Renault Group: Collaboration between Renault and RITS re-started in 2016 and is expected to know a major growth in 2017. Three different research teams in Renault are now working separately with RITS on different topics.

A first concrete action was the beginning of a CIFRE PhD thesis funded by Renault and the French ANRT. The thesis deals with the accurate localization of an autonomous vehicle on a highway using mainly on-board low-cost perception sensors.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

8.1.1.1. COCOVEA

Title: Coopération Conducteur-Véhicule Automatisé

Instrument: ANR

Duration: November 2013 - April 2017

Coordinator: Jean-Christophe Popieul (LAMIH - University of Valenciennes)

Partners: LAMIH, IFSTTAR, Inria, University of Caen, COMETE, PSA, CONTINENTAL, VALEO, AKKA Technologies, SPIROPS

Inria contact: Fawzi Nashashibi

Abstract: CoCoVeA project aims at demonstrating the need to integrate from the design of the system, the problem of interaction with the driver in resolving the problems of sharing the driving process and the degree of freedom, authority, level of automation, prioritizing information and managing the operation of the various systems. This approach requires the ability to know at any moment the state of the driver, the driving situation in which he finds himself, the operating limits of the various assistance systems and from these data, a decision regarding activation or not the arbitration system and the level of response.

8.1.1.2. VALET

Title: Redistribution automatique d'une flotte de véhicules en partage et valet de parking

Instrument: ANR

Duration: January 2016 - December 2018

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.

8.1.2. FUI

8.1.2.1. Sinetic

Title: Système Intégré Numérique pour les Transports Intelligents Coopératifs

Instrument: FUI

Duration: December 2014 - May 2017

Coordinator: Thomas Nguyen (Oktal)

Partners: Oktal, ALL4TEC, CIVITEC, Dynalogic, Inria, EURECOM, Renault, Armines, IFSTTAR, VEDECOM

Inria contact: Jean-Marc Lasgouttes

Abstract: The purpose of the project SINETIC is to create a complete simulation environment for designing cooperative intelligent transport systems with two levels of granularity: the system level, integrating all the components of the system (vehicles, infrastructure management centers, etc.) and its realities (terrain, traffic, etc.) and the component-level, modeling the characteristics and behavior of the individual components (vehicles, sensors, communications and positioning systems, etc.) on limited geographical areas, but described in detail.

8.1.2.2. PAC V2X

Title: Perception augmentée par coopération véhicule avec l'infrastructure routière

Instrument: FUI

Duration: September 2016 - August 2019

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). VEDECOM is financing the PhD theses of Mr. Fernando Garrido Carpio and Mr. Zayed Alsayed.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. CityMobil2

Type: COOPERATION (TRANSPORTS)

Instrument: Large-scale integrating project

Objectif: NC

Duration: September 2012 - August 2016

Coordinator: University of Rome La Sapienza, CTL (Italy)

Partners: Inria (France), DLR (Germany), GEA Chanard (Switzerland), POLIS (Belgium), ERT (Belgium), EPFL (Switzerland),...(45 partners!)

Inria contact: Fawzi Nashashibi

Abstract: The CityMobil2 goal is to address and to remove three barriers to the deployment of automated road vehicles: the implementation framework, the legal framework and the unknown wider economic effect. CityMobil2 features 12 cities which will revise their mobility plans and adopt wherever they will prove effective automated transport systems. Then CityMobil2 will select the best 5 cases (among the 12 cities) to organize demonstrators. The project will procure two sets of automated vehicles and deliver them to the five most motivated cities for a 6 to 8 months demonstration in each city. CityMobil2 will establish a workgroup that will deliver a proposal for a European Directive to set a common legal framework to certify automated transport systems.

See also: http://www.citymobil2.eu/en/

8.2.1.2. AutoNet2030

Title: Co-operative Systems in Support of Networked Automated Driving by 2030

Objectif: NC

Duration: November 2013 - October 2016

Coordinator: Andras KOVACS – BROADBIT (Hungary)

Partners: BROADBIT (Hungary), BASELABS (Germany), CRF (Italy), Armines (France), VOLVO (Sueden), HITACHI EUROPE (France), EPFL (Switzerland), ICCS (Greece), TECHNISCHE UNIVERSITAET DRESDEN (Germany) (9 partners)

Inria contact: Fawzi Nashashibi

AutoNet2030 shall develop and test a co-operative automated driving technology, based on a decentralized decision-making strategy which is enabled by mutual information sharing among nearby vehicles. The project is aiming for a 2020-2030 deployment time horizon, taking into account the expected preceding introduction of co-operative communication systems and sensor based lane-keeping/cruise-control technologies. By taking this approach, a strategy can be worked out for the gradual introduction of fully automated driving systems, which makes the best use of the widespread existence of co-operative systems in the near-term and makes the deployment of fully automated driving systems beneficial for all drivers already from its initial stages.

See also: http://www.autonet2030.eu

8.2.1.3. 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 - December 2018

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

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** 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/

RITS is member of the Working Group on Automation: **iMobility**. This group has been created and is animated by ERTICO ITS Europe. The Automation Working Group was formed under the iMobility Forum, with the initial high level aims of exploring and promoting the potential of highly automated vehicles and applications and working towards the development of a roadmap for the deployment of automated systems.

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. International Academics Partners

NAIST – Japan: RITS has a close cooperation with NAIST (Nara institute of Science and Technology), Japan since 2009. Based on this collaboration NAIST and Inria established the MoU agreement to accelerate and strengthen future research collaborations and the exchange of researchers and students. During the period February 2015 - March 2016 RITS hosted Sakriani Watiasri Sakti, assistant professor at NAIST.

Seoul National University - S. Korea: An International Cooperation Agreement has been signed between RITS team of Inria and the Vehicle Dynamics and Control Laboratory (VDCL) of Seoul National University (SNU). RITS and VDCL recognize the value of educational, cultural, and scientific exchanges between international research laboratories, and have determined that sufficient interest exists to establish an academic and research partnership for collaborative research and education in the area of future intelligent vehicle systems for sustainable safety and environment.

International Chaire "Drive4U": Inria-RITS, Mines ParisTech, EPFL, Univ. of Berkeley (PATH Program) and Shanghai Jiao Tong Univ. (SJTU) are the academic partners of the international Chaire GAT, funded and supported by: VALEO Group, SAFRAN Group and MPSA Group (Peugeot-Citroën). A recent NDA has been signed recently. This Chaire will promote and fund academic activities related to Ground Automated Transportation and autonomous driving.

8.3.2. Participation in Other International Programs

8.3.2.1. ICT-Asia

SIM-Cities

Title: "Sustainable and Intelligent Mobility for Smart Cities"

International Partner (Institution - Laboratory - Researcher):

- Nanyang Technical University (NTU), School of Electrical and Electronic Engineering – Singapore. Prof. Dan Wei Wang

- National University of Singapore (NUS), Department of Mechanical Engineering – Singapore. Dr. Marcelo Ang

- Kumamotoo University - Japan. Intelligent Transportation Systems Lab, Graduate School of Science and Technology, Prof. James Hu / Prof. Ogata

- Shanghai Jiao-Tong University (SJTU), Department of Automation – China. Prof. Ming Yang

- Hanoi University of Science and Technology, International Center MICA Institute – Vietnam. Prof. Eric Castelli

- Inria, RITS Project-Team – France. Dr. Fawzi Nashashibi

- Inria, e-Motion/CHROMA Project-Team – France. Dr. Christian Laugier

- Ecole Centrale de Nantes, IRCCyN – France. Prof. Philippe Martinet

Duration: Jan. 2015 - May 2017

Start year: 2015

This project aims at conducting common research and development activities in the field of sustainable transportation and advanced mobility of people and goods in order to move in the direction of smart, clean and sustainable cities.

RITS and MICA lab have obtained from the Vietnamese Program 911 the financing of the joint PhD thesis of Dinh-Van Nguyen (co-directed by Eric Castelli from MICA lab and Fawzi Nashashibi).

8.3.2.2. ECOS Nord – Venezuela

ECOS Nord

Title: "Les Techniques de l'Information et de la Communication pour la Conception de Systèmes Avancés de Mobilité durable en Milieu Urbain."

International Partner (Institution - Laboratory - Researcher):

- Simon Bolivar University, Department of Mecatronics – Venezuela. Dr. Gerardo Fernandez

- Inria, RITS Project-Team – France. Dr. Fawzi Nashashibi

Duration: Jan. 2014 - Dec. 2017

Start year: 2014

The main objective of this project is to contribute scientifically and technically to the design of advanced sustainable mobility systems in urban areas, particularly in dense cities where mobility, comfort and safety needs are more important than in other types of cities. In this project, we will focus on the contribution of advanced systems of perception, communication and control for the realization of intelligent transport systems capable of gradually integrating into the urban landscape. These systems require the development of advanced dedicated urban infrastructures as well as the development and integration of on-board intelligence in individual vehicles or mass transport.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

Sakriani Watiasri Sakti, assistant professor at NAIST, from February 2015 until March 2016. A part of the work done during her stay has been published in [51].

Aidos Ibrayev, PhD student, from Kazakhstan.

Pablo Marin Plaza, PhD student, from Universidad Carlos III de Madrid, Spain.

8.4.1.1. Internships

Rafael Colmenares Prieto, Juan Jose Larez Urdaneta, Daniel Sanchez Aranguren from Simon Bolivar University, Venezuela.

Aitor Gomez Torres, Alfredo Valle Barrio and Myriam Vaca Recalde from Universidad Politécnica de Madrid, Spain.

Jose Emilio Traver Becerra from Universidad de Extremadura, Spain.

Kenneth Martinez Torres from Universidad del Turabo, Porto Rico.

Alexis Meyer from Télécom SudParis.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

Fawzi Nashashibi: member of the steering committee of VENITS 2016 (2nd International Workshop on Vehicular Networking and Intelligent Transportation Systems), Sep. 2016, Valencia, Spain

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

Guy Fayolle: scientific program committee member of *The IEEE Computer Society's MASCOTS 2016 (Modelling, Analysis and Simulation of Computer and Telecommunication Systems)*, September 21-23, Imperial College, London, UK.

Vicente Milanés: international program committee member of IEEE International Conference on Intelligent Transportation Systems (ITSC 2016).

Fawzi Nashashibi: international program committee member of IEEE Intelligent Vehicles Symposium (IV 2016).

Fawzi Nashashibi: IPC for the 2016 IEEE Intelligent Transportation Systems Conference (ITSC'2016), 8th Workshop on Planning, Perception and Navigation for Intelligent Vehicles, Nov. 2016, Rio de Janeiro.

Fawzi Nashashibi: member of the Technical Program Committee of the 14th Int. Conference on Control, Automation, Robotics and Vision, Nov. 2016, Phuket.

Fawzi Nashashibi: member of the Technical Program Committee of the 2016 International Conference on Computer Science and Artificial Intelligence, Aug. 2016, Guilin.

Oyunchimeg Shagdar: program committee member of the IEEE International Conference on Advanced Information Networking and Applications (AINA 2016).

Ahmed Soua: technical program committee member of IEEE International Conference on Communications and of International Workshop on Vehicular Adhoc Networks for Smart Cities (IWVSC 2016).

Anne Verroust-Blondet: technical program committee member of MMEDIA 2016.

9.1.2.2. Reviewer

Raoul de Charette: IEEE International Conference on Intelligent Transportation Systems (ITSC 2016).

Guy Fayolle: IEEE MASCOTS 2016 (Modelling, Analysis and Simulation of Computer and Telecommunication Systems) conference.

Jean-Marc Lasgouttes: IEEE MASCOTS 2016 (Modelling, Analysis and Simulation of Computer and Telecommunication Systems) conference.

Fawzi Nashashibi: International conference on Multidisciplinary in IT and Communication Science and Technologies 2016 (AIC-MITC), Transportation Research Board (TRB 2016), IEEE International Conference on Robotics and Automation (ICRA 2016), IEEE Intelligent Transportation Systems Conference (ITSC 2016), IEEE International Conference on Intelligent Robots and Systems (IROS 2016), International Conference on Advanced Robotics and Computer Vision (ICARCV 2016), Reconnaissance des Formes et Intelligence Artificielle (RFIA 2016).

Anne Verroust-Blondet: IEEE International Conference on Intelligent Transportation Systems (ITSC 2016).

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Guy Fayolle: associate editor of the journal *Markov Processes and Related Fields*

Vicente Milanés: editor-in-chief of Journal of Computer and Communications, associate Editor of Journal of Intelligent Transportation and Urban Planning.

Fawzi Nashashibi: associate editor of IEEE Transactions on Intelligent Vehicles, associate editor of the IEEE Transactions on Intelligent Transportation Systems, Editor of the Journal of Internet of Things (IoT) ISTE Editions (London).

Oyunchimeg Shagdar: associate editor of Wiley International Journal of Communication Systems.

9.1.3.2. Reviewer - Reviewing Activities

Raoul de Charette: *Journal of Computer and Communications*.

Guy Fayolle: *AAP, MPRF, PTRF, QUESTA, European Journal of Combinatorics, JSP, Physica A*.

Jean-Marc Lasgouttes: *Journal of Applied Probability, IEEE Transactions on Knowledge and Data Engineering*.

Pierre Merdrignac: *IEEE Trans. on Intelligent Transportation Systems*.

Fawzi Nashashibi: *IEEE Transactions on Intelligent Vehicles, IEEE Transactions on Intelligent Transportation Systems, IEEE Intelligent Transportation Systems Magazine*.

Ahmed Soua: *IEEE Communications Letters*.

Thomas Streubel: *IEEE Transactions on Intelligent Vehicles*.

Anne Verroust-Blondet: *Transactions on Pattern Analysis and Machine Intelligence*.

9.1.4. Invited Talks

Raoul de Charette: "Perception research for Autonomous Vehicles", Sino-European Workshop on Information and Technology, Beijin, October 10-11 2016.

Guy Fayolle: "Random walks in the quarter-plane: explicit criterions for the finiteness of the associated group in the genus 1 case", Int. Conf. on Random Walks in Cones with Big Jumps, Tours, December 14-16 2016.

Fawzi Nashashibi:

1. Plenary keynote at the B.I.G. Forum (Big data Initiative of Gyeonggi). Invitation by the Geyonggi Province and Seoul National University, Oct. 19-20, Pangyo, Korea.
2. Keynote speaker at the Plenary session of the International Conference on Advanced Robotics and Computer Vision (ICARCV'2016), Nov. 15, Phuket, Thailand.
3. Keynote at the French ASPROM Workshop "‘De la connectée à la voiture autonome : technologies, enjeux et applications"’, Feb. 10-11, Paris, France.
4. Keynote at the "‘Algorithms for HRI Workshop 2016"’, July 21, Paris, France.

Anne Verroust-Blondet: "Sketch-based 3D model retrieval", Whitehead Lectures in Cognition, Computation and Culture, Goldsmiths, University of London, 24 May 2016.

9.1.5. Scientific Expertise

Guy Fayolle is scientific advisor at the Robotics Laboratory of Mines ParisTech.

Jean-Marc Lasgouttes is member of the Conseil Académique of Université Paris-Saclay.

9.1.6. Research Administration

Jean-Marc Lasgouttes is a member of the Comité Technique Inria.

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.

Anne Verroust-Blondet is the scientific correspondent of the European affairs and of the International relations of Inria Paris. She is member of the COST-GTRI committee at Inria.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence: Fawzi Nashashibi, “Programmation avancée”, 84h, L1, Université Paris-8 Saint-Denis, France.

Master: Jean-Marc Lasgouttes, “Analyse de données”, 54h, second year of Magistère de Finance (M1), University Paris 1 Panthéon Sorbonne, France.

Master: Fawzi Nashashibi, “Image synthesis and 3D Infographics”, 12h, M2, INT Télécom Sud-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.

Master: Oyunchimeg Shagdar, “Nouvelles technologies de communication”, 32h, M2, University Versailles Saint-Quentin, France.

Master : Pierre Merdrignac, Introduction au traitement d’image, 12h (M1), Université Versailles Saint-Quentin en Yvelines, France.

Master: David González Bautista and Anne Verroust-Blondet, "Le véhicule autonome. Présentation des recherches de l’équipe-projet RITS",1.5 h, 2nd year, Ecole des Ponts ParisTech, France, Sept. 2016.

Doctorat: Jean-Marc Lasgouttes, “Analyse de données fonctionnelles”, 31.5h, Mastère Spécialisé “Expert en sciences des données”, INSA-Rouen, France

9.2.2. Supervision

PhD: Mohammad Abualhoul, “Visible light and radio communication for cooperative autonomous driving: applied to vehicle convoy”, Mines ParisTech, December 2016, supervisor: Fawzi Nashashibi, co-supervisor: Oyunchimeg Shagdar. (cf. [17])

PhD: Guillaume Trehard, “Gestion des intersections pour la conduite urbaine autonome”, Mines ParisTech, February 2016, supervisor: Fawzi Nashashibi.(cf. [18]).

PhD in progress : Zayed Al-Sayed, “Système de localisation redondant en environnement extérieur ouvert pour véhicule urbain automatique”, Télécom ParisTech, October 2014, supervisor: Anne Verroust-Blondet, co-supervisor: Guillaume Bresson.

PhD in progress : David González Bautista, “Contrôle coordonné et planification dynamique des trajectoires pour un système de transport cybernétique en milieu urbain instrumenté”, Mines ParisTech, January 2014, supervisor: Fawzi Nashashibi.

PhD in progress : Pierre de Beaucorps, “Autonomous vehicle: behavior prediction and Interaction with road users”, UPMC Paris, January 2016, supervisor: Anne Verroust-Blondet, co-supervisor: Fawzi Nashashibi.

PhD in progress : Fernando Garrido Carpio, “Optimal trajectory generation for autonomous vehicles in urban environments” , Mines ParisTech, November 2014, supervisor: Fawzi Nashashibi, co-supervisors: Vicente Milanés, Joshué Pérez.

PhD in progress : Carlos Flores, “Analysis and design of cooperative systems for trains of green cars”, Mines ParisTech, December 2015, supervisor: Fawzi Nashashibi, co-supervisor: Vicente Milanés.

PhD in progress : Francisco Navas, “Plug&Play control for highly non-linear systems: Stability analysis of autonomous vehicles”, Mines ParisTech, October 2015, supervisor: Fawzi Nashashibi, co-supervisor: Vicente Milanés.

PhD in progress : Dinh-Van Nguyen, "Wireless sensor networks for indoor mapping and accurate localization for low speed navigation in smart cities", Mines ParisTech, December 2015, supervisor: Fawzi Nashashibi, co-supervisor: Eric Castelli.

Starting PhD: Farouk Ghallabi, “Environment modeling and simultaneous localization of a mobile vehicle on motorways: a multi-sensor approach”, Mines ParisTech, October 2016, supervisor: Fawzi Nashashibi.

9.2.3. Juries

Guy Fayolle was a jury member of the PhD thesis defense of Mohamed Hadded "Design and Optimization of Access Control Protocols in Vehicular Ad Hoc Networks (VANETs)", Télécom SudParis, 30 November 2016. Fawzi Nashashibi was a reviewer and examiner of the following PhD theses:

- Rémy Sohier – « Sensibilité artistique et création vidéoludique : Critique du jeu par le jeu et modèle ludographique ». Université Paris VIII Saint-Denis, 28 November 2016 – Reviewer

- Yiyuan Huang – « Méthode de création numérique et interactivité inconsciente entre philosophie de l'oeuvre et psychologie du participant. Ré-exploration de la peinture traditionnelle chinoise de manière numérique en utilisant l'installation artistique interactive ». Université de Paris VIII Saint-Denis, département Arts et Technologies de l'Image. 14 June 2016 – Reviewer

- Adela Sirbu – « Dynamic Machine Learning for Supervised and Unsupervised Classification". Thèse en co-tutelle de l'Université BABES-BOLYAI, CLUJ-NAPOCA et de l'Université de Normandie délivrée par INSA Rouen. 6 June 2016 – Reviewer

- Carlos Fernandez-Lopez – «Road Scene Interpretation For Autonomous Navigation Fusing Stereo Vision and Digital Maps». Université de Alcalà (Spain). 23 September 2016 – Examiner

Anne Verroust-Blondet was:

- the external examiner of the PhD Examining Committee of Prashant Aparajeya's PhD "Medialness-based shape invariant feature transformation", Goldsmidth, University of London (UK), 23 May 2016.

- a jury member of the PdD thesis defense of Arthur Truong "Analyse du contenu expressif des gestes corporels", Télécom SudParis, 21 September 2016.

9.3. Popularization

An article about YoGoKo (spinoff RITS) was published by bpifrance. *"YoGoKo joue la connectivité routière"*. Date: 09/02/2016.

"La Tribune" published an article about the Link&Go prototype developed by AKKA Technologies and Inria-Rits. *"Voiture autonome : la LINK & GO d'AKKA veut sortir des frontières"*, pp.72-73. Date: 29/06/2016. Journalist: Marie-Annick Depagneux.

Fawzi Nashashibi was interviewed in the "L'Auto-Journal" journal. Title: *"Les voitures autonomes sont là."*. Date : 21/01-03/02 16. OJD : 105052, pp.30-32. Journalist: Brice Perrin.

Fawzi Nashashibi was interviewed in the "Usine Nouvelle" journal. Title: *"À la recherche du véhicule vraiment autonome"*. Date: 14/04/2016. Also: *"Le véhicule autonome fait avancer la simulation"*, pp. 48-49. Journalist: Frédéric Parisot.

Fawzi Nashashibi was interviewed in the "webcar center" (http://www.webcarcenter.com). Title: *"Conduite autonome : les premières voitures sont là"*. Date: 10/02/2016. Journalist: Brice Perrin.

Fawzi Nashashibi was interviewed by "L'Automobile & L'Entreprise" (http://www.automobile-entreprise.com). Journalist: Séverine Fontaine.

Fawzi Nashashibi was interviewed by "Télérama" magazine. Title: *"IA : c'est vraiment demain?"*. Date : 23/04-29/04 16. OJD : 578680. Journalists: Olivier Tesquet / Richard Sénéjoux.

Fawzi Nashashibi was interviewed by "Innovation Review". Title: *"Voitures autonomes : quel est le rôle du « deep learning » ?"*. Journalist: Thibault Lescuyer.

Fawzi Nashashibi was interviewed by "Les Echos" newspaper. Title: *"Google Car, le robot qui a remis la voiture autonome dans la course"*. Date: 02-03 Sept. 2016, pp.10-11. Journalist: Jacques Henno.

Fawzi Nashashibi was interviewed by "Science & Vie Special Edition - High Tech -". Title: *"Voiture connectée : la révolution cachée"*, pp. 78-83. Date: N 43/2016.

Fawzi Nashashibi was interviewed by "Industrie & Technologies". Title: *"Pourquoi et comment les véhicules autonomes vont devoir coopérer"*. Date: 11/10/2016. Journalist: Juliette Raynal. (https://www.industrie-techno.com/pourquoi-et-comment-les-vehicules-autonomes-vont-devoir-cooperer.46065).

Fawzi Nashashibi participated to a radio program on France Culture, La méthode scientifique, *"Jusqu'où ira l'autonomie des voitures ?"*, September 28, 2016.

Fawzi Nashashibi participated to a round table on *"Transports et énergies du futur"* for the event *Science en direct*, organized at the Cité des sciences et de l'industrie, Paris, October 8, 2016.

RITS team members participated to the making of a short movie dedicated to Teletoon Channel. Title/Serie: *"Culture Décode"*. Date: 09/12/2016. Journalist: Benjamin Brun (Tralalère: http://www.tralalere.com/).

10. Bibliography

Major publications by the team in recent years

[1] M. ABUALHOUL, O. SHAGDAR, F. NASHASHIBI.*Visible Light Inter-Vehicle Communication for Platooning of Autonomous Vehicles*, in "2016 IEEE Intelligent Vehicles Symposium IV2016", Gothenburg, Sweden, June 2016, https://hal.inria.fr/hal-01308430.

[2] G. FAYOLLE, R. IASNOGORODSKI, V. A. MALYSHEV.*Random walks in the Quarter Plane*, Applications of Mathematics, Springer-Verlag, 1999, n^{o} 40.

[3] C. FLORES, V. MILANÉS, F. NASHASHIBI.*Using Fractional Calculus for Cooperative Car-Following Control*, in "Intelligent Transportation Systems Conference 2016", Rio de Janeiro, Brazil, IEEE, November 2016, https://hal.inria.fr/hal-01382821.

[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] D. GONZÁLEZ BAUTISTA, J. PÉREZ RASTELLI, R. LATTARULO, V. MILANÉS, F. NASHASHIBI.*Continuous curvature planning with obstacle avoidance capabilities in urban scenarios*, in "2014 IEEE 17th International Conference on Intelligent Transportation Systems (ITSC)", Qingdao, China, October 2014, https://hal.inria.fr/hal-01086888.

[6] G. LE LANN.*Cohorts and groups for safe and efficient autonomous driving on highways*, in "Vehicular Networking Conference (VNC)", IEEE, 2011, p. 1-8.

[7] H. LI, F. NASHASHIBI.*Multi-vehicle cooperative localization using indirect vehicle-to-vehicle relative pose estimation*, in "ICVES 2012 - IEEE International Conference on Vehicular Electronics and Safety", Istanbul, Turkey, IEEE, July 2012, p. 267 - 272 [*DOI :* 10.1109/ICVES.2012.6294256], https://hal.inria.fr/hal-00763825.

[8] H. LI, F. NASHASHIBI.*Cooperative Multi-Vehicle Localization Using Split Covariance Intersection Filter*, in "IEEE Intelligent Transportation Systems Magazine", April 2013, vol. 5, n^{o} 2, p. 33-44, https://hal.inria.fr/hal-00833707.

[9] M. MAROUF, E. POLLARD, F. NASHASHIBI.*Automatic parking and platooning for electric vehicles redistribution in a car-sharing application*, in "IOSR Journal of Electrical and Electronics Engineering", 2015, vol. 10, n^{o} 1, 9, https://hal.inria.fr/hal-01254336.

[10] 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.

[11] 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^{o} 1, p. 123-154 [*DOI :* 10.1007/s10472-015-9470-x], https://hal.inria.fr/hal-01186220.

[12] P. MERDRIGNAC, O. SHAGDAR, F. NASHASHIBI.*Fusion of Perception and V2P Communication Systems for Safety of Vulnerable Road Users*, in "IEEE Transactions on Intelligent Transportation Systems", 2016, https://hal.inria.fr/hal-01399150.

[13] P. MORIGNOT, J. PÉREZ RASTELLI, F. NASHASHBI.*Arbitration for balancing control between the driver and ADAS systems in an automated vehicle: Survey and approach*, in "2014 IEEE Intelligent Vehicles Symposium (IV)", Dearborn, United States, June 2014, p. 575 - 580 [*DOI :* 10.1109/IVS.2014.6856577], https://hal.inria.fr/hal-01081302.

[14] F. NAVAS, V. MILANÉS, F. NASHASHIBI.*Using Plug&Play Control for stable ACC-CACC system transitions*, in "Intelligent Vehicles Symposium 2016", Gothemburg, Sweden, June 2016 [*DOI :* 10.1109/IVS.2016.7535464], https://hal.inria.fr/hal-01304542.

[15] P. PETROV, F. NASHASHIBI.*Modeling and Nonlinear Adaptive Control for Autonomous Vehicle Overtaking*, in "IEEE Transactions Intelligent Transportation Systems", August 2014, vol. 15, n^{o} 4, p. 1643–1656.

[16] G. TREHARD, E. POLLARD, B. BRADAI, F. NASHASHIBI.*On line Mapping and Global Positioning for autonomous driving in urban environment based on Evidential SLAM*, in "Intelligent Vehicles Symposium - IV 2015", Seoul, South Korea, June 2015, https://hal.inria.fr/hal-01149504.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[17] M. ABUALHOUL.*Visible Light and Radio Communication for Cooperative Autonomous Driving: applied to vehicle convoy*, MINES ParisTech, December 2016, https://hal.inria.fr/tel-01447124.

[18] G. TREHARD.*Evidence theory applications for localization and mapping in an urban context*, Ecole Nationale Supérieure des Mines de Paris, February 2016.

Articles in International Peer-Reviewed Journal

[19] D. CHRISTIE, A. KOYMANS, T. CHANARD, J.-M. LASGOUTTES, V. KAUFMANN.*Pioneering driverless electric vehicles in Europe: the City Automated Transport System (CATS)*, in "Transportation Research Procedia", 2016, vol. 13, p. 30-39 [*DOI :* 10.1016/J.TRPRO.2016.05.004], https://hal.inria.fr/hal-01357309.

[20] G. FAYOLLE, P. MUHLETHALER.*A Markovian Analysis of IEEE 802.11 Broadcast Transmission Networks with Buffering*, in "Probability in the Engineering and Informational Sciences", June 2016, vol. 30, n^{o} 3, 19 [*DOI :* 10.1017/S0269964816000036], https://hal.inria.fr/hal-01166082.

[21] 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.

[22] Y. HAN, F. MOUTARDE.*Analysis of Large-Scale Traffic Dynamics in an Urban Transportation Network Using Non-Negative Tensor Factorization*, in "International Journal of Intelligent Transportation Systems Research",

January 2016, vol. 14, n^{o} 1, p. 36-49 [*DOI :* 10.1007/s13177-014-0099-7], https://hal-mines-paristech.archives-ouvertes.fr/hal-01085971.

[23] 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^{o} 1, p. 123-154 [*DOI :* 10.1007/s10472-015-9470-x], https://hal.inria.fr/hal-01186220.

[24] P. Merdrignac, O. Shagdar, F. Nashashibi.*Fusion of Perception and V2P Communication Systems for Safety of Vulnerable Road Users*, in "IEEE Transactions on Intelligent Transportation Systems", 2016, https://hal.inria.fr/hal-01399150.

[25] Z. Yasseen, A. Verroust-Blondet, A. Nasri.*Shape Matching By Part Alignment Using Extended Chordal Axis Transform*, in "Pattern Recognition", September 2016, vol. 57 [*DOI :* 10.1016/j.patcog.2016.03.022], https://hal.inria.fr/hal-01356174.

[26] Z. Yasseen, A. Verroust-Blondet, A. Nasri.*View selection for sketch-based 3D model retrieval using visual part shape description*, in "Visual Computer", 2016, 19 [*DOI :* 10.1007/s00371-016-1328-7], https://hal.inria.fr/hal-01396333.

International Conferences with Proceedings

[27] M. Abualhoul, P. Merdrignac, O. Shagdar, F. Nashashibi.*Study and Evaluation of Laser-based Perception and Light Communication for a Platoon of Autonomous Vehicles*, in "2016 IEEE 19th Intelligent Transportation Systems Conference (ITSC 2016)", Rio de Janeiro, Brazil, November 2016, https://hal.inria.fr/hal-01366147.

[28] M. Abualhoul, O. Shagdar, F. Nashashibi.*Visible Light Inter-Vehicle Communication for Platooning of Autonomous Vehicles*, in "2016 IEEE Intelligent Vehicles Symposium IV2016", Gothenburg, Sweden, June 2016, https://hal.inria.fr/hal-01308430.

[29] N. Achir, Y. Bouchaala, P. Muhlethaler, O. Shagdar.*Comparison of Spatial Aloha and CSMA using Simple Stochastic Geometry Models for 1D and 2D Networks*, in "ICT 2016 - 23rd International Conference on Telecommunications, 2016", Thessalonique, Greece, May 2016 [*DOI :* 10.1109/ICT.2016.7500470], https://hal.inria.fr/hal-01368875.

[30] G. Bresson, M.-C. Rahal, D. Gruyer, M. Revilloud, Z. Alsayed.*A Cooperative Fusion Architecture for Robust Localization: Application to Autonomous Driving*, in "IEEE Intelligent Transportation Systems Conference 2016", Rio de Janeiro, Brazil, November 2016, https://hal.archives-ouvertes.fr/hal-01379322.

[31] C. Flores, V. Milanés, F. Nashashibi.*Using Fractional Calculus for Cooperative Car-Following Control*, in "Intelligent Transportation Systems Conference 2016", Rio de Janeiro, Brazil, IEEE, November 2016, https://hal.inria.fr/hal-01382821.

[32] F. Garrido, D. Gonzalez Bautista, V. Milanés, J. Pérez, F. Nashashibi.*Optimized trajectory planning for Cybernetic Transportation Systems*, in "9th IFAC Symposium on Intelligent Autonomous Vehicles IAV 2016", Leipzig, Germany, June 2016, vol. 49, n^{o} 15, p. 1-6, https://hal.inria.fr/hal-01356691.

[33] F. GARRIDO, D. GONZALEZ BAUTISTA, V. MILANÉS, J. PÉREZ, F. NASHASHIBI.*Real-time planning for adjacent consecutive intersections*, in "19th International IEEE Conference on Intelligent Transportation Systems - ITSC 2016", Rio de Janeiro, Brazil, November 2016, https://hal.inria.fr/hal-01356706.

[34] L. GONZALEZ, J. PÉREZ, D. GONZALEZ BAUTISTA, F. NASHASHIBI.*Perception and Control Strategies for Autonomous Docking for Electric Freight Vehicles* , in "Transport Research Arena TRA 2016", Warsaw, Poland, April 2016 [*DOI :* 10.1016/J.TRPRO.2016.05.116], https://hal.inria.fr/hal-01397934.

[35] D. GONZALEZ BAUTISTA, V. MILANÉS, J. PÉREZ RASTELLI, F. NASHASHIBI.*Speed Profile Generation based on Quintic Bezier Curves for Enhanced Passenger Comfort*, in "2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC 2016)", Rio de Janeiro, Brazil, November 2016, https://hal.inria.fr/hal-01388396.

[36] G. LE LANN.*A Collision-Free MAC Protocol for Fast Message Dissemination in Vehicular Strings*, in "2016 IEEE Conference on Standards for Communications & Networking", Berlin, Germany, IEEE ComSoc, October 2016, 7, https://hal.inria.fr/hal-01402119.

[37] G. LE LANN.*Fast Distributed Agreements and Safety-Critical Scenarios in VANETs*, in "2017 IEEE International Conference on Computing, Networking and Communications", Santa Clara, CA, United States, 2017 IEEE International Conference on Computing, Networking and Communications, IEEE ComSoc, January 2017, 7, https://hal.inria.fr/hal-01402159.

[38] F. NAVAS, V. MILANÉS, F. NASHASHIBI.*Using Plug&Play Control for stable ACC-CACC system transitions*, in "Intelligent Vehicles Symposium 2016", Gothemburg, Sweden, June 2016 [*DOI :* 10.1109/IVS.2016.7535464], https://hal.inria.fr/hal-01304542.

[39] D.-V. NGUYEN, M. VACA RECALDE, F. NASHASHIBI.*Low Speed Vehicle Localization using WiFi Finger-Printing*, in "International Conference on Control, Automation, Robotics and Vision, ICARCV 2016", Phuket, Thailand, November 2016, https://hal.inria.fr/hal-01395973.

[40] L. RAFFAELLI, G. FAYOLLE, F. VALLÉE.*ADAS Reliability and Safety*, in "20ème Congrès de maîtrise des risques et de sûreté de fonctionnement", Saint-Malo, France, E. LARDEUX, A. BRACQUEMOND (editors), Congrès Lambda MU 20, Institut pour la Maîtrise des Risques (IMdR), October 2016, 10, https://hal.inria.fr/hal-01398428.

[41] L. RAFFAELLI, F. VALLÉE, G. FAYOLLE, P. DE SOUZA, X. ROUAH, M. PFEIFFER, S. GÉRONIMI, F. PÉTROT, S. AHIAD.*Facing ADAS validation complexity with usage oriented testing*, in "ERTS 2016", Toulouse, France, January 2016, 13, https://hal.inria.fr/hal-01277494.

[42] M. VACA RECALDE, J. E. TRAVER, V. MILANÉS, J. PÉREZ, D. GONZALEZ BAUTISTA, F. NASHASHIBI.*Automated Global Planner for Cybernetic Transportation Systems*, in "14th international conference on control, automation, robotics and vision (ICARCV)", Phuket, Thailand, November 2016, https://hal.inria.fr/hal-01401380.

References in notes

[43] 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.

[44] 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^{o} 4, Accepted for publication in the journal MPRF (Markov Processes and Related Fields), https://hal.inria.fr/hal-01086684.

[45] G. FAYOLLE, R. IASNOGORODSKI, V. A. MALYSHEV.*Random walks in the Quarter Plane*, Applications of Mathematics, Springer-Verlag, 1999, n^{o} 40.

[46] 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^{o} 2, p. 317-348.

[47] 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.

[48] 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^{o} 9143, 14, Mobile Ad Hoc Networks [*DOI :* 10.1007/978-3-319-19662-6_24], https://hal.inria.fr/hal-01172595.

[49] 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.

[50] V. MARTIN.*Modélisation probabiliste et inférence par l'algorithme Belief Propagation*, Ecole Nationale Supérieure des Mines de Paris, May 2013, http://hal.inria.fr/tel-00867693.

[51] S. SAKTI, O. SHAGDAR, F. NASHASHIBI, S. NAKAMURA.*Context Awareness and Priority Control for ITS based on Automatic Speech Recognition*, in "International conference on ITS Telecommunications", Copenhagen, Denmark, December 2015, https://hal.inria.fr/hal-01225312.

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

Keywords:

Computer Science and Digital Science:

4. - Security and privacy
4.2. - Correcting codes
4.3. - Cryptography
4.3.1. - Public key cryptography
4.3.2. - Secret key cryptography
7.2. - Discrete mathematics, combinatorics
7.8. - Information theory
7.13. - Quantum algorithms

Other Research Topics and Application Domains:

6.4. - Internet of things
6.5. - Information systems
9.8. - Privacy

1. Members

Research Scientists

Anne Canteaut [Team leader, Inria, Senior Researcher, HDR]
André Chailloux [Inria, Researcher]
Pascale Charpin [Inria, Senior Researcher, Emeritus, HDR]
Gaëtan Leurent [Inria, Starting Research position]
Anthony Leverrier [Inria, Researcher on leave from Corps des Mines]
María Naya Plasencia [Inria, Researcher]
Nicolas Sendrier [Inria, Senior Researcher, HDR]
Jean-Pierre Tillich [Inria, Senior Researcher, HDR]

PhD Students

Xavier Bonnetain [Univ. Paris VI, from Sept 2016]
Rodolfo Canto Torres [Inria]
Kevin Carrier [Min. de la Défense, from Oct 2016]
Kaushik Chakraborty [Inria]
Julia Chaulet [Thales, granted by CIFRE]
Thomas Debris [Univ. Paris VI, from Aug 2016]
Sébastien Duval [Univ. Paris VI]
Antoine Grospellier [ENS Lyon, from Sep 2016]
Adrien Hauteville [Univ. Limoges]
Virginie Lallemand [Inria, until Oct 2016]
Vivien Londe [Univ. Bordeaux, from Sep 2016]
Yann Rotella [Inria]

Post-Doctoral Fellows

Irene Márquez Corbella [Inria, until Apr 2016]
Nicky Mouha [FWO grant (Belgium), until Jun 2016]

Visiting Scientist

Thomas Peyrin [NTU Singapore, from Feb 2016 until Mar 2016, and June 2016]

Administrative Assistant

Christelle Guiziou [Inria]

Others

Xavier Bonnetain [Inria, internship, from Mar 2016 until Aug 2016]
Rémi Bricout [ENS Paris, internship, from Mar 2016 until Aug 2016]
Vivien Londe [Inria, internship, from Apr 2016 until July 2016]
Thomas Debris [ENS Cachan, internship, from Mar 2016 until Aug 2016]
Ghazal Kachigar [Inria, internship, from Mar 2016 until Sep 2016]

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 very recently, with a submission deadline in November 2017.

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:

[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/

(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 unconditional 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

5.1.1. Post-quantum symmetric cryptanalysis

We have been considering the problem of symmetric cryptography in the future environment that will see the arrival of quantum computers. Indeed, this environment will pose a real problem for the majority of asymmetric primitives, but little is known about the implications for the security of symmetric primitives. Confidence in our symmetric primitives is entirely based on our knowledge within the field of cryptanalysis, but in reality, we do not know much about the symmetric post-quantum attacks. If we want post-quantum systems to be reliable and efficient, we need to understand how adversaries might exploit this new computing power. This year, two preliminary results have been obtained within the team and published at CRYPTO 2016 [51] and in the *IACR Transactions on Symmetric Cryptology* [23]. They include surprising results demonstrating that, in some scenarios, some symmetric systems can also become vulnerable to the quantum computer. Recently María Naya-Plasencia has been awarded an ERC starting grant, QUASYModo, to work on this subject. This grant will enable us to continue this work in more depth.

5.1.2. Real-word impact of some theoretical cryptanalytic works

Weak cryptography can be used long after weaknesses have been found by the academic community. For instance, Rogaway warned that the predictable IV used in TLS was a problem in 2002, but it took a public demonstration with a practical exploit in 2011 (the BEAST attack) for servers and clients to implement countermeasures. The same happened with the use of compression (CRIME), unsecure version fallback (POODLE), and known biases in RC4 (RC4NOMORE), to name a few examples. In joint works at NDSS and ACM CCS, K. Bhargavan from the PROSECCO project-team and G. Leurent showed two almost practical attacks against deprecated cryptographic primitives that are still used in real-world applications. The SLOTH attack targeted the use of MD5 in TLS for in-protocol signatures, and the Sweet32 attack targeted the use of 64-bit block ciphers: Blowfish in OpenVPN, and 3DES in TLS. Moreover, the SLOTH attack received a distinguished paper award at NDSS.

5.1.3. Symmetric ciphers for homomorphic encryption schemes

In order to avoid the (extremely) high expansion rate of homomorphic encryption, a solution consists in transmitting to the server the ciphertext c obtained by encrypting m with a symmetric scheme (the corresponding secret key encrypted by the homomorphic cipher is also transmitted). The server then needs to compute m encrypted with the homomorphic scheme from c, i.e. the server needs to homomorphically evaluate the decryption circuit of the symmetric cipher. Hybrid encryption schemes dedicated to this application then require the use of symmetric ciphers with very specific features. Our team has two important contributions on this topic: the design of new appropriate solutions based on stream ciphers [44], and the attack of a cipher proposed by Méaux et al. in this context [48], [32].

5.1.4. Awards

BEST PAPERS AWARDS :
[58] **Post-Quantum Cryptography - 7th International Workshop, PQCrypto 2016**. A. PHESSO, J.-P. TILLICH.
[41] **Network and Distributed System Security Symposium – NDSS 2016**. K. BHARGAVAN, G. LEURENT.

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: Nicolas Sendrier and Gregory Landais
- 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: Nicolas Sendrier and Gregory Landais
- 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 varaint of the algorithm due to May, Meurer and Thomae.

- Participants: Nicolas Sendrier and Gregory Landais
- Contact: Anne Canteaut
- URL: https://gforge.inria.fr/projects/collision-dec/

7. New Results

7.1. Symmetric cryptology

Participants: Xavier Bonnetain, Anne Canteaut, Pascale Charpin, Sébastien Duval, Virginie Lallemand, Gaëtan Leurent, Nicky Mouha, María Naya Plasencia, Yann Rotella.

7.1.1. Block ciphers

Our recent results mainly concern either the analysis and design of lightweight block ciphers.
Recent results:

- Design and study of a new construction for low-latency block ciphers, named *reflection ciphers*, which generalizes the so-called α-reflection property exploited in PRINCE. This construction aims at reducing the implementation overhead of decryption on top of encryption [13].
- Design of a new permutation for wide-block block ciphers: N. Mouha and S. Gueron have proposed a family of cryptographic permutations, named Simpira, that supports inputs of $128b$ bits, where b is a positive integer [50]. This wide-block permutation is mainly based on the AES round-function. It then achieves a very high throughput on virtually all modern 64-bit processors that have native instructions for AES.
- Analysis of the division property against block ciphers [42], [26]: A. Canteaut, together with C. Boura, gave a new approach to the division property, which has been recently introduced as a distinguishing property on block ciphers. This work provides a simpler and more general view of the division property which allows the attacker to take into account the characteristics of the building-blocks of the cipher. As an illustration, this new approach provides low-data distinguishers against reduced-round Present, which reach a much higher number of rounds than previously known distinguishers of the same type.
- Modes of operation for full disk encryption [52]: L. Khati, N. Mouha and D. Vergnaud have classified various FDE modes of operation according to their security in a setting where there is no space to store additional data, like an IV or a MAC value. They also introduce the notion of a diversifier, which does not require additional storage, but allows the plaintext of a particular sector to be encrypted into different ciphertexts.

7.1.2. Authenticated encryption and MACs

A limitation of all classical block ciphers is that they aim at protecting confidentiality only, while most applications need both encryption and authentication. These two functionalities are provided by using a block cipher like the AES together with an appropriate mode of operation. However, it appears that the most widely-used mode of operation for authenticated encryption, AES-GCM, is not very efficient for high-speed networks. Also, the security of the GCM mode completely collapses when an IV is reused. These severe drawbacks have then motivated an international competition named CAESAR, partly supported by the NIST, which has been recently launched in order to define some new authenticated encryption schemes [0]. The project-team is involved in a national cryptanalytic effort in this area led by the BRUTUS project funded by the ANR.

Recent results:

- Attack against π-Cipher : G. Leurent and his coauthors have presented a guess-and-determine attack against some variants of the π-Cipher family, which is a second-round candidate to the Caesar competition. More precisely, they showed a key recovery attack with time complexity little higher than $2^{4\omega}$, and low data complexity, against variants of the cipher with ω-bit words, when the internal permutation is reduced to 2.5 rounds out of 3.
- Improved generic attacks against hash-based MAC [20]
- Cryptanalysis of 7 (out of 8) rounds of the Chaskey MAC [54]. This work has led the designers of Chaskey to increase the number of rounds.

7.1.3. Stream ciphers

Stream ciphers provide an alternative to block-cipher-based encryption schemes. They are especially well-suited in applications which require either extremely fast encryption or a very low-cost hardware implementation.

Recent results:

- Design of encryption schemes for efficient homomorphic-ciphertext compression (see Section 5.1.3): A. Canteaut, M. Naya-Plasencia together with their coauthors have investigated the constraints on the symmetric cipher imposed by this application and they have proposed some solutions based on additive IV-based stream ciphers [44], [30].
- Cryptanalysis of the FLIP family of stream ciphers: S. Duval, V. Lallemand and Y. Rotella have exhibited an attack against a new family of stream ciphers intended for use in Fully Homomorphic Encryption systems, and proposed by Méaux et al. at Eurocrypt 2016 [48], [32]. More precisely, their attack applies to the early version of FLIP. It exploits the structure of the filter function and the constant internal state of the cipher. The proposed algorithm then recovers the secret key for the two instantiations originally proposed by Méaux et al.
- New types of correlation attacks against filter generators: A. Canteaut and Y. Rotella presented a new family of attacks against filter generators, which exploit a change of the primitive root defining the LFSR [45]. Most notably, an attack can often be mounted by considering non-bijective monomial mappings. In this setting, a divide-and-conquer strategy applies, based on a search within a multiplicative subgroup of $\mathbb{F}_{2^n}$ where n is the LFSR length. If the LFSR length is not a prime, a fast correlation involving a shorter LFSR can then be performed.

7.1.4. 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

[0] http://competitions.cr.yp.to/caesar.html

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:

- Cryptographic properties of involutions: P. Charpin, together with S. Mesnager and S. Sarkar, has provided a rigorous study of involutions over the finite field of order 2^n which are relevant primitives for cryptographic designs [19]. Most notably, they have focused on the class of involutions defined by Dickson polynomials [61].
- Construction of a new family of permutations over binary fields of dimension $(4k+2)$ with good cryptographic properties. An interesting property is that this family includes as a specific case the only known APN permutation of an even number of variables [64].
- Construction of cryptographic permutations over finite fields with a sparse representation: P. Charpin, together with N. Cepak and E. Pasalic, exhibited permutations which are derived from sparse functions via linear translators [14].
- New methods for determining the differential spectrum of an Sbox: P. Charpin and G. Kyureghyan have proved that the whole differential spectrum of an Sbox can be determined without examining all derivatives of the mapping, but only the derivatives with respect to an element within a hyperplane [18]. Also, they have proved that, for mappings of a special shape, it is enough to consider the derivatives with respect to all elements within a suitable multiplicative subgroup of $\mathbb{F}_{2^n}$.

7.1.5. Side-channel attacks

Physical attacks must be taken into account in the evaluation of the security of lightweight primitives. Indeed, these primitives are often dedicated to IoT devices in pervasive environments, where an attacker has an easy access to the devices where the primitive is implemented.

Recent results:

- Differential fault attack against the block cipher PRIDE [53]: the efficiency of this attack mainly originate from the design of the linear layer of the cipher which relies on the interleaved construction.
- Study of the criteria to quantify the resistance offered by an Sbox to differential power analysis [17]. This work by K. Chakraborty and his coauthors shows that the classical criterion, called transparency order, has many limitations; an alternative definition is then proposed.

7.1.6. Security of Internet protocols

Cryptographic primitives are used to in key-exchange protocols such as TLS, IKE and SSH, to verify the integrity of the exchange. The recent works by K. Bhargavan and G. Leurent show the real-word impact of some recent theoretical cryptanalytic works.

Recent results:

- Impact of hash function collisions on the security of TLS: most practitioners believe that the hash function only need to resist preimage attacks for this use. However, K. Bhargavan and G. Leurent have shown that collisions in the hash function are sufficient to break the integrity of these protocols, and to impersonate some of the parties [41], [34]. Since many protocols still allow the use of MD5 or SHA-1 (for which collision attacks are known), this results in some practical attacks, and extends the real-world impact of the collision attacks against MD5 and SHA-1. This work has already influenced the latest TLS 1.3 draft, and the main TLS libraries are removing support of MD5 signatures.
- Use of block ciphers operating on small blocks: It is well-known that most modes of operation, like CBC, are not secure if the same key is used for encrypting $2^{n/2}$ blocks of plaintext, where n is the block size. But this threat has traditionally been dismissed as impractical, even for 64-bit blocks, since it requires some prior knowledge of the plaintext and even then, it only leaks a few secret bits per gigabyte. In this context, K. Bhargavan and G. Leurent demonstrated two concrete attacks that exploit such short block ciphers [40]. First, they presented an attack on the use of 3DES in HTTPS that can be used to recover a secret session cookie. Second, they showed how a similar attack on Blowfish can be used to recover HTTP BasicAuth credentials sent over OpenVPN connections.

7.2. Code-based cryptography

Participants: Rodolfo Canto Torres, Julia Chaulet, Thomas Debris, Adrien Hauteville, Ghazal Kachigar, Irene Márquez Corbella, Nicolas Sendrier, Jean-Pierre Tillich.

The first cryptosystem based on error-correcting codes was a public-key encryption scheme proposed by McEliece in 1978; a dual variant was proposed in 1986 by Niederreiter. We proposed the first (and only) digital signature scheme in 2001. Those systems enjoy very interesting features (fast encryption/decryption, short signature, good security reduction) but also have their drawbacks (large public key, encryption overhead, expensive signature generation). Some of the main issues in this field are

- security analysis, including against a quantum adversary, implementation and practicality of existing solutions,
- reducing the key size, *e.g.*, by using rank metric instead of Hamming metric, or by using particular families of codes,
- addressing new functionalities, like hashing or symmetric encryption.

Recent results:

- J. Chaulet and N. Sendrier are working on the analysis Gallager's bit flipping algorithm for the decoding of QC-MDPC codes. A first outcome is an improved decoder with an adaptative threshold [47]. The ultimate goal of this work is to avoid side-channel attacks on QC-MDPC-McEliece by designing a failure-free constant-time decoder.
- We have started to explore whether generalized Reed-Solomon codes, and more generally MDS codes, can be used in a McEliece cryptosystem. We have first started by a fundamental work about MDS codes by first characterizing which MDS codes can be efficiently decoded with the rather general technique using error correcting pairs [25] We have also studied whether it is possible, if we know only a random generator matrix of a code admitting an error correcting pair, to recover the pair itself [55]. The latter problem is precisely the problem that an attacker wants to solve when he wants to perform a key attack on a McEliece system based on MDS codes admitting an error correcting pair. Finally, we have come up with what we believe to be a viable McEliece scheme based on Reed-Solomon codes by combining them with a generalized $U|U+V$ construction which hides at the same time the algebraic structure and even improves the decoding capacity of the code [57].
- Design of a new code-based stream cipher, named RankSynd, variant of Synd for the rank metric [49] and of the first Identity based Encryption Scheme relying on error correcting codes (paper currently under submission which is joint work of P. Gaborit, A. Hauteville, H. Phan and J.P. Tillich).
- Structural attacks against some variants of the McEliece cryptosystem based on subclasses of alternant/Goppa codes which admit a very compact public matrix, typically quasi-cyclic, quasi-dyadic, or quasi-monoidic matrices [22]. This result is obtained thanks to a new operation on codes called folding that exploits the knowledge of the automorphism group of the code [21].
- Cryptanalysis of a variant of McEliece cryptosystem based on polar codes [38].
- The previous work has been extended by exploring some structural properties of polar codes in [39]. In particular, we have been able to show that these codes have a very large automorphism group and have found an efficient way of counting the number of codewords of low weight.
- Cryptanalysis of all McEliece cryptosystems relying on algebraic geometry codes [73].
- Cryptanalysis of a code-based signature scheme proposed at PQCrypto 2013 by Baldi at al. [58]. This paper has received the best paper award of PQCrypto 2016.
- R. Canto Torres and N. Sendrier have investigated the information-set decoding algorithms applied to the case where the number of errors is sub-linear in the code length [46]. This situation appears in the analysis of the McEliece scheme based on quasi-cyclic Moderate Density Parity Check (MDPC) codes.
- We have also investigated other decoding techniques such as statistical decoding [74] or quantum algorithms [75]. The last work has led to the best known quantum algorithms for decoding a linear code.

7.3. Quantum Information

Participants: Xavier Bonnetain, Rémi Bricout, Kaushik Chakraborty, André Chailloux, Antoine Grospellier, Gaëtan Leurent, Anthony Leverrier, Vivien Londe, María Naya Plasencia, Jean-Pierre Tillich.

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 started in September 2016 on this topic. First, Antoine Grospellier, co-advised by A. Leverrier and O. Fawzi (Ens Lyon), will study efficient decoding algorithms for quantum LDPC codes. 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 is co-advised by A. Leverrier and G. Zémor (IMB) and his thesis is devoted to the design of better quantum LDPC codes: the main idea is to generalize the celebrated toric code of Kitaev by considering cellulations of manifolds in higher dimensions. A recent 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.

Recent results:

- Introduction of a new class of quantum LDPC codes, “Quantum expander codes”, featuring a simple and very efficient decoding algorithm which can correct arbitrary patterns of errors of size scaling as the square-root of the length of the code. These are the first codes with constant rate for which such an efficient decoding algorithm is known [36], [59].

7.3.2. Quantum cryptography

A recent approach to cryptography takes into account that all interactions occur in a physical world described by the laws of quantum physics. These laws put severe constraints on what an adversary can achieve, and allow for instance to design provably secure key distribution protocols. We study such protocols as well as more general cryptographic primitives such as coin flipping with security properties based on quantum theory.

Recent results:

- A. Chailloux, together with colleagues from IRIF and Jerusalem, established the existence of quantum weak coin flipping with arbitrarily small bias [12].
- A. Chailloux and international collaborators performed an experimental verification of multipartite entanglement in quantum networks [24].
- A. Chailloux and collaborators established the optimal bounds for quantum weak oblivious transfer [15].
- Security analysis of quantum key distribution with continuous variables [35].

7.3.3. Relativistic cryptography

Two-party cryptographic tasks are well-known to be impossible without complexity assumptions, either in the classical or the quantum world. Remarkably, such no-go theorems become invalid when adding the physical assumption that no information can travel faster than the speed of light. This additional assumption gives rise to the emerging field of relativistic cryptography. We recently started investigating such questions through the task of bit commitment. In a paper in *Physical Review Letters* in 2015, K. Chakraborty, A. Chailloux and A. Leverrier developed a security proof for a simple and easily implementable protocol that can achieve arbitrarily long commitment times, thereby establishing that relativistic cryptography is a very practical solution.

André Chailloux was awarded an ANR "Jeune chercheur" to develop the field of relativistic cryptography [31].

Recent results:

- R. Bricout and A. Chailloux [70] considered explicit attacks against the relativistic protocol for bit commitment mentioned above and proved that the security analysis published in *Physical Review Letters* 2015 is essentially tight.
- A drawback of the relativistic bit commitment protocol is that it requires that all communications remain perfectly synchronized during the entire commitment time, and a single network failure leads to aborting the protocol. K. Chakraborty, A. Chailloux and A. Leverrier proposed a more robust version of the protocol allowing to deal with such network failures, a required feature in order to implement the protocol in realistic conditions [16], [71].

7.3.4. Quantum cryptanalysis of symmetric primitives

Symmetric cryptography seems at first sight much less affected in the post-quantum world than asymmetric cryptography: its main known threat is Grover's algorithm, which allows for an exhaustive key search in the square root of the normal complexity. For this reason, it is 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. M. Naya Plasencia has recently been awarded an ERC Starting grant for her project named QUASYModo on this topic.

Recent results:

- Differential and linear attacks in the quantum setting: G. Leurent, A. Leverrier and M. Naya Plasencia, in collaboration with M. Kaplan, have obtained some results on quantum versions of differential and linear cryptanalysis [23]. They show that it is usually possible to use quantum computations to obtain a quadratic speed-up for these attacks, but not for all variants. Therefore, the best attack in the classical world does not necessarily lead to the best quantum one.
- Application of Simon's algorithm to symmetric cryptanalysis [51], [33]: Leurent et al. also proved that several attacks can be dramatically sped up using a quantum procedure known as Simon's algorithm for finding the period of a function. As a first application, the most widely used modes of operation for authentication and authenticated encryption (e.g. CBC-MAC, PMAC, GMAC, GCM, and OCB) are completely broken in this security model. These quantum attacks are also applicable to many CAESAR candidates: CLOC, AEZ, COPA, OTR, POET, OMD, and Minalpher. Second, Simon's algorithm can also be applied to slide attacks, leading to an exponential speed-up of a classical symmetric cryptanalysis technique in the quantum model.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

- **Thales** ($02/14 \to 01/17$)
  Funding for the supervision of Julia Chaulet's PhD.
  30 kEuros.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

- **ANR BLOC** $(10/11 \rightarrow 03/16)$
  Design and Analysis of block ciphers dedicated to constrained environments
  ANR program: Ingénierie numérique et sécurité
  Partners: INSA Lyon, Inria (project-team SECRET), University of Limoges (XLIM), CryptoExperts
  446 kEuros
  http://bloc.project.citi-lab.fr
  The BLOC project aims at providing strong theoretical and practical results in the domain of cryptanalysis and design of block ciphers.
- **ANR KISS** $(12/11 \rightarrow 02/16)$
  Keep your personal Information Safe and Secure
  ANR program: Ingénierie numérique et sécurité
  Partners: Inria (project-teams SMIS and SECRET), LIRIS, Gemalto, University of Versailles-St Quentin, Conseil Général des Yvelines
  64 kEuros
  The KISS project builds upon the emergence of new portable and secure devices known as Secure Portable Tokens (e.g., mass storage SIM cards, secure USB sticks, smart sensors) combining the security of smart cards and the storage capacity of NAND Flash chips. The idea promoted in KISS is to embed, in such devices, software components capable of acquiring, storing and managing securely personal data.
- **ANR BRUTUS** $(10/14 \rightarrow 09/18)$
  Authenticated Ciphers and Resistance against Side-Channel Attacks
  ANR program: Défi Société de l'information et de la communication
  Partners: ANSSI, Inria (project-team SECRET and project-team MARELLE), Orange, University of Lille, University of Rennes, University Versailles-Saint Quentin
  160 kEuros
  The Brutus project aims at investigating the security of authenticated encryption systems. We plan to evaluate carefully the security of the most promising candidates to the Caesar competition, by trying to attack the underlying primitives or to build security proofs of modes of operation. We target the traditional black-box setting, but also more "hostile" environments, including the hardware platforms where some side-channel information is available.
- **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.

9.1.2. Others

- **DGA-MI** $(09/15 \rightarrow 09/16)$
  Analysis of binary streams: reconstructing LDPC codes.
  28.6 kEuros.
  The objective of this contract was to examine the code reconstruction problem (from noisy observation) for LDPC codes.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. PQCRYPTO

Title: Post-quantum cryptography for long-term security

Programm: H2020

Duration: March 2015 - March 2018

Coordinator: Technische Universiteit Eindhoven (NL)

Partners:

Academia Sinica (Taiwan)

Bundesdruckerei (Germany)

Danmarks Tekniske Universitet (Denmark)

Katholieke Universiteit Leuven (Belgium)

Nxp Semiconductors Belgium Nv (Belgium)

Ruhr-Universität Bochum (Germany)

Stichting Katholieke Universiteit (Netherlands)

Technische Universiteit Eindhoven (Netherlands)

Technische Universitaet Darmstadt (Germany)

University of Haifa (Israel)

Inria contact: Nicolas Sendrier

Online banking, e-commerce, telemedicine, mobile communication, and cloud computing depend fundamentally on the security of the underlying cryptographic algorithms. Public-key algorithms are particularly crucial since they provide digital signatures and establish secure communication without requiring in-person meetings. Essentially all applications today are based on RSA or on the discrete-logarithm problem in finite fields or on elliptic curves. Cryptographers optimize parameter choices and implementation details for these systems and build protocols on top of these systems; cryptanalysts fine-tune attacks and establish exact security levels for these systems. Alternative systems are far less visible in research and unheard of in practice. It might seem that having three systems offers enough variation, but these systems are all broken as soon as large quantum computers are built. The EU and governments around the world are investing heavily in building quantum computers; society needs to be prepared for the consequences, including cryptanalytic attacks accelerated by these computers. Long-term confidential documents such as patient health-care records and state secrets have to guarantee security for many years, but information encrypted today using RSA or elliptic curves and stored until quantum computers are available will then be as easy to decipher as Enigma-encrypted messages are today. PQCRYPTO will allow users to switch to post-quantum cryptography: cryptographic systems that are not merely secure for today but that will also remain secure long-term against attacks by quantum computers. PQCRYPTO will design a portfolio of high-security post-quantum public-key systems, and will improve the speed of these systems, adapting to the different performance challenges of mobile devices, the cloud, and the Internet of Things. PQCRYPTO will provide efficient implementations of high-security post-quantum cryptography for a broad spectrum of real-world applications.

9.2.1.2. 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.

9.2.2. Collaborations in European Programs, Except FP7 & H2020

Program: COST

Project acronym: ICT COST Action IC1306

Project title: Cryptography for Secure Digital Interaction

Duration: January 2014 - November 2017

Coordinator: Claudio Orlandi, Aarhus University, Denmark

Other partners: see http://www.cost.eu/domains_actions/ict/Actions/IC1306

Abstract: The aim of this COST action is to stimulate interaction between the different national efforts in order to develop new cryptographic solutions and to evaluate the security of deployed algorithms with applications to the secure digital interactions between citizens, companies and governments.

Anne Canteaut is co-leader of the working group on cryptographic primitives. She co-organized a 2-day workshop for PhD students and early-career researchers in symmetric cryptography, DISC 2016 (Bochum, Germany, March 23-24 2016).

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. 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 - 2018

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.

9.3.1.2. Informal International Partners

- Otto-von-Guericke Universität Magdeburg, Institut für Algebra und Geometrie (Germany): Study of Boolean functions for cryptographic applications
- Nanyang Technological University (Singapore): cryptanalysis of symmetric primitives.
- Ruhr-Universität Bochum (Germany): design and cryptanalysis of symmetric primitives.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Leo Perrin, University of Luxemburg, visiting PhD student, June 2016.
- Thomas Peyrin, NTU Singapore, visiting scientist, Feb.-March 2016 and June 2016.

9.4.1.1. Internships

- Xavier Bonnetain, MPRI and Telecom ParisTech, March-Aug. 2016
- Rémi Bricout, MPRI and ENS Paris, March-Aug. 2016
- Thomas Debris, MPRI and ENS Cachan, March-Aug. 2016
- Ghazal Kachigar, Master cryptographie et mathématiques de l'information, Univ. Rennes, March-Sept. 2016
- Vivien Londe, Master de mathématiques, UPMC, April-July 2016

9.4.2. Visits to International Teams

9.4.2.1. Short Research Stays Abroad

- Ruhr-Universität Bochum, Bochum, Germany, January 18-22, work with Gregor Leander (G. Leurent)
- Instituto Superior Tecnico, Lisbon, Portugal, May 18-20, 2016, invitation to visit the group of quantum computation of Paulo Mateus (A. Leverrier)
- University of Oxford Mathematical Institute, Oxford, UK, May 25-26, invitation to the cryptography seminar (G. Leurent)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- DISC 2016, Workshop for early-career symmetric cryptographers funded by the COST Action IC1306, Bochum, Germany, March 23-24 2016. https://disc2016.compute.dtu.dk/, co-organizer: A. Canteaut.
- Research retreat (H2020 PQCRYPTO), September 21-22, 2016, Inria de Paris, organizer: N. Sendrier

10.1.1.2. Member of the Organizing Committees

- EuroS&P 2017: April 26-28, 2015, Paris (France): G. Leurent (poster chair)

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- FSE 2017: March 5-8, 2017, Tokyo, Japan: M. Naya-Plasencia (co-chair).

10.1.2.2. Member of the Conference Program Committees

- PQCrypto 2016: February 24-26, 2016, Fukuoka, Japan (N. Sendrier, J.P. Tillich)
- CT-RSA 2016: Feb. 29- March 4, 2016, San Francisco, USA (M. Naya Plasencia)
- FSE 2016: March 20-23, 2016, Bochum, Germany (A. Canteaut, G. Leurent)
- Eurocrypt 2016: May 8-12, 2016, Vienna, Austria (M. Naya Plasencia)
- Crypto 2016: August 14-18, 2016, Santa Barbara, USA (A. Canteaut)

- ACISP 2016: July 4-6, 2016, Melbourne, Australia (G. Leurent)
- Waifi 2016: July 13-15, 2016, Ghent, Belgium (A. Canteaut)
- YACC 2016: June 6-10, 2016, Porquerolles Island (A. Canteaut)
- SAC 2016: August 10-12, 2016, St. John's, NL, Canada (G. Leurent, M. Naya-Plasencia)
- Lightsec 2016: September 21-22, 2016, Cappadocia, Turkey (M. Naya-Plasencia)
- Redundancy 2016: September 26-29, 2016, St. Petersburg, Russia (P. Charpin)
- TQC 2016: September 27-29, 2016, Berlin, Germany (A. Chailloux);
- SETA 2016 (International Conference on SequEnces and Their Applications): October 9-14, 2016, Chengdu, China (P. Charpin).
- Asiacrypt 2016: December 4-8, 2016, Hanoi, Vietnam (A. Canteaut)
- Indocrypt 2016: December 11-14, 2016, Kolkata, India (G. Leurent)
- QIP 2017: January 16-20, 2017, Seattle, USA (A. Chailloux, A. Leverrier)
- Financial Crypto 2017: April 3-7, 2017, Sliema, Malta (G. Leurent)
- Fq13: June 4-9, 2017, Geata, Italy (A. Canteaut)
- Crypto 2017: August 20-24, 2017, Santa Barbara, CA, USA (G. Leurent)

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- *Designs, Codes and Cryptography*, associate editor: P. Charpin.
- *Finite Fields and Their Applications*, associate editors: A. Canteaut, P. Charpin.
- *Annals of telecommunications*, associate editor : J.-P. Tillich.
- *Applicable Algebra in Engineering, Communication and Computing*, associate editor: A. Canteaut.
- *IACR Transactions on Symmetric Cryptology*, associate editors: A. Canteaut and G. Leurent, co-editor-in-chief: M. Naya-Plasencia.

P. Charpin serves as a reviewer for *Mathematical Reviews*.

10.1.3.2. Editor for books or special issues

- Special issue in Coding and Cryptography, *Designs, Codes and Cryptography*, to appear, editors: P. Charpin, N. Sendrier and J-P. Tillich.
- *Contemporary Developments in Finite Fields and Applications*, 2016, World Scientific Publishing [62], co-editor: A. Canteaut.

10.1.4. Invited Talks

- G. Leurent *Breaking Symmetric Cryptosystems Using Quantum Period Finding*, TCCM-CACR 2016, Yinchuan, China, August 2016
- A. Leverrier, *Quantum Expander Codes*, Beyond i.i.d. in Information Theory, Barcelone, Spain, 18-22 June 2016

The members of the project-team have also been invited to give talks to some workshops or international seminars, including:

- A. Canteaut, *Another view of the division property* Dagstuhl seminar on symmetric cryptology, Dagstuhl, Germany, Jan. 10-14, 2016.
- A. Canteaut, *Stream Ciphers: A Practical Solution for Efficient Homomorphic-Ciphertext Compression*, CryptoAction Symposium 2016, Budapest, Hungary, April 6-8, 2016.
- A. Canteaut, *Algebraic Distinguishers against Symmetric Primitives*, Paris Crypto Day, France, June 30, 2016.

- A. Canteaut, *Comment concevoir un algorithme de chiffrement sûr et efficace : l'héritage de Shannon*, Shannon 100, workshop organized at the occasion of Shannon's 100th birthday, Institut Henri Poincaré, Paris October 26, 2016. The talk is available online at https://www.youtube.com/watch?v=BYlOO4MkVgU.
- A. Chailloux, *Cryptographie relativiste*, Coding, Cryptography and Algorithms (CCA), Paris, July 1, 2016.
- A. Chailloux, *Quantum Information Processing*, Journées Scientifiques Inria 2016, Rennes, France, June 2016.
- V. Lallemand, *Cryptanalysis of the FLIP Family of Stream Ciphers*, Paris Crypto Day, Sept. 6, 2016.
- G. Leurent, *Transcript Collision Attacks*, Dagstuhl seminar on symmetric cryptology, Dagstuhl, Germany, Jan. 10-14, 2016.
- A. Leverrier, *Distributing Secret Keys with Quantum Continuous Variables*, Recent Advances in Continuous-variable Quantum Information Theory, Barcelone, Spain, 16-8 April 2016
- M. Naya-Plasencia: *Pourquoi essaie-t-on de casser les fonctions cryptographiques ?*. Colloquium organised by the pre-GDR Sécurité Informatique: Colloque Sécurité informatique CNRS http://colloque-cybersecu.cnrs.fr/. Paris, France, Dec. 9, 2016.
- J.P. Tillich, *Attaining the capacity with Reed-Solomon codes through the $(U|U+V)$ construction and Koetter-Vardy soft decoding*, Journée Claude Shannon, Paris, July 1, 2016.

10.1.5. Leadership within the Scientific Community

- A. Canteaut serves as a chair of the steering committee of *Fast Software Encryption (FSE)*.
- N. Sendrier serves on the steering committee of *Post-quantum cryptography (PQCrypto)*.
- M. Naya Plasencia serves on the steering committee of the *Coding and Cryptography* group of GDR-IM https://crypto.di.ens.fr/c2:main;
- N. Sendrier is a member of the "Comité de pilotage" of the ANR (défi 9);
- Since 2014, JP. Tillich organizes a working group on code-based cryptography which meets on a monthly/bimonthly basis. It gathers people from the project-team, from the GRACE project-team (Inria Saclay), from the University of Limoges, from the University of Rennes and from the University of Rouen who all work on this topic.

10.1.6. Research Administration

- N. Sendrier has been a vice-chair of the "Commission d'Evaluation" at Inria until October 2016;
- A. Canteaut is a member of the "Comité de pilotage" of the Fondation Sciences Mathématiques de Paris;
- M. Naya-Plasencia is a member of *Inria Paris CES Committee* (Comité de suivi doctoral).
- M. Naya-Plasencia is a member of *Inria Paris Scientific Hiring Committee* (Assignement of PhD, post-doctoral and delegation Inria fundings).
- N. Sendrier served on the jury of PEDR CNRS INSII 2016.
- J.-P. Tillich is in charge of "Formation par la recherche" for the Paris Inria center;
- **Committees for the selection of professors, assistant professors and researchers**: Inria Paris Chargés de recherche (A. Canteaut), University Paris 8 assistant professor (A. Canteaut, M. Naya-Plasencia, JP Tillich), Inria Directeurs de recherche (N. Sendrier)

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: A. Canteaut, *Introduction to Symmetric Cryptography*, 7 hours, M1, Telecom ParisTech, France;

Master: A. Canteaut, *Error-correcting codes and applications to cryptology*, 12 hours, M2, University Paris-Diderot (MPRI), France;

Master: A. Chailloux, *Quantum computing*, 6 hours, M2, University Paris-Diderot (MPRI), France;

Master: N. Sendrier, *Code-based cryptography*, 4.5 hours, M2, University Paris-Diderot (MPRI), France;

Master: N. Sendrier, Information theory, 32 hours, M1, University of Versailles-St Quentin (MINT), France;

Master: J.-P. Tillich, *Introduction to Information Theory*, 32 hours, M2, Ecole Polytechnique, France.

The members of the project-team also gave advanced lectures to summer schools for PhD students:

- *UbiCrypt Spring School on Symmetric Cryptography*, Bochum, Germany, March 2016: A. Canteaut (9 hours). Some of the lectures are available online.

E-learning

Mooc: I. Marquez-Corbella and N. Sendrier, *Code-based cryptography*, 5 weeks, FUN, Inria, undergraduate and Master's degree students in mathematics or computer science.

Pedagogical resources: https://www.fun-mooc.fr/courses/inria/41006S02/session02/about

10.2.2. Supervision

PhD: Virginie Lallemand, *Cryptanalysis for symmetric cryptography*, University Pierre-et-Marie Curie, October 5, 2016, supervisors: M. Naya-Plasencia and A. Canteaut

PhD in progress: Julia Chaulet, *Study of public-key cryptosystems based on MDPC quasi-cyclic codes*, since February 2014, CIFRE convention with Thales, supervisor: N. Sendrier

PhD in progress: Kaushik Chakraborty, *Position-based Quantum Cryptography*, since October 2014, supervisors: A. Leverrier, J.P. Tillich

PhD in progress: Adrien Hauteville, *Rank-metric-based Cryptosystems*, since October 2014, supervisors: P. Gaborit (Univ. Limoges) and J.-P. Tillich

PhD in progress: Rodolfo Canto Torres, *Analysis of generic decoding algorithms for the Hamming metric and study of cryptosystems based on the rank metric*, since September 2015, supervisor: N. Sendrier

PhD in progress: Sébastien Duval, *Constructions for lightweight cryptography*, since October 2015, supervisor: A. Canteaut and G. Leurent

PhD in progress: Yann Rotella, *Finite fields and symmetric cryptography*, since October 2015, supervisor: A. Canteaut

PhD in progress: Xavier Bonnetain, *Cryptanalysis of symmetric primitives in the post-quantum world*, since September 2016, supervisor: M. Naya Plasencia and A. Canteaut

PhD in progress: Thomas Debris, *Quantum algorithms for decoding linear codes*, since September 2016, supervisor: J.-P. Tillich

PhD in progress: Antoine Grospellier, *LDPC codes: constructions and decoding*, since October 2016, supervisor: J.-P. Tillich

PhD in progress: Vivien Londe, *Study of quantum LDPC codes*, since September 2016, supervisors: G. Zémor and A. Leverrier

PhD in progress: Kevin Carrier, *Reconstruction of error-correcting codes*, since October 2016, supervisor: N. Sendrier

10.2.3. Juries

- Mohamed Nidhal Mejri, *Securing Vehicular Networks against Denial of Service attacks*, University Paris 13, May 19, 2016, committee: A. Canteaut;

- Tung Chou *Accelerating Pre- and Post-quantum Cryptography*, TU Eindhoven, The Netherlands, June 26, 2016, committee: N. Sendrier;
- Jean-Marie Le Bars, *Some studies about randomness in Computer Science*, HdR, University of Caen, June 29, 2016, committee: J.P. Tillich (reviewer);
- Tom Douce, *Realistic quantum information processing: from devices to computational models*, Université Paris Diderot, September 9, 2016, committee: A. Leverrier;
- Virginie Lallemand, *Cryptanalysis for symmetric cryptography*, University Pierre-et-Marie Curie, October 5, 2016, committee: M. Naya-Plasencia and A. Canteaut (supervisors)
- Brice Minaud, *Analysis of recent cryptographic primitives*,University of Rennes 1, October 7, 2016, committee: A. Canteaut;
- Pierre Karpmann, *Analysis of symmetric primitives*, University Paris-Saclay, October 18, 2016, committee: A. Canteaut (reviewer);
- Jean-Christophe De Neuville, *Contributions to post-quantum cryptography*, University of Limoges, December 1, 2016, committee: J.P. Tillich (reviewer).
- Zoé Amblard , *Quantum cryptography and applications to spatial communications*, University of Limoges, December 5, 2016, committee: J.P. Tillich (reviewer).
- Qian Guo, *Using coding techniques for attacking post-quantum cryptographic assumptions and systems*, Lund University, Sweden, December 13, 2016, committee: J.P. Tillich.

10.3. Popularization

- Nicolas Sendrier and Jean-Pierre Tillich, *Code-Based Cryptography: New Security Solutions Against a Quantum Adversary*, ERCIM News [67].
- Anne Canteaut gave a talk at the *dotSecurity 2016* conference for developers, at Théâtre des Variétés, Paris, April 2016 http://www.thedotpost.com/2016/05/anne-canteaut-the-struggle-for-secure-cryptography.
- Anne Canteaut gave a talk at *Séminaire général du département d'informatique de l'ENS* for Master students in computer science at ENS Paris, April 13, 2016 http://savoirs.ens.fr/expose.php?id=2516.
- André Chailloux gave a talk entitled *L'ordinateur quantique*, at Journées Art, Cerveau, Futur; Mouans-Sartoux, France, September 2016;
- Anne Canteaut gave a talk on cryptography at lycée Rodin, Paris, February 2, 2016.
- Sébastien Duval gave a talk on cryptography at lycée des 7 Mares, Maurepas, December 2, 2016
- Anne Canteaut has been involved in the AlKindi competition, which is a national competition on cryptanalysis for students in "Seconde" http://www.concours-alkindi.fr/.

  The best teams from Paris have been visiting the SECRET project-team in June 2016 https://www.youtube.com/watch?v=EVLHEOWAORc.
- Julia Chaulet participated to a general-public mediation about the use of mathematics in industry at "Salon Culture & Jeux Mathématiques", Paris, May 28, 2016.
- Yann Rotella hold a stand to explain cryptography at Futur en Seine, Carreau du Temple, Paris, June 12, 2016.

11. Bibliography

Major publications by the team in recent years

[1] A. CANTEAUT, B. CHEVALLIER-MAMES, A. GOUGET, P. PAILLIER, T. PORNIN, E. BRESSON, C. CLAVIER, T. FUHR, T. ICART, J.-F. MISARSKY, M. NAYA-PLASENCIA, J.-R. REINHARD, C. THUILLET, M. VIDEAU.*Shabal, a Submission to NIST's Cryptographic Hash Algorithm Competition*, October 2008, Submission to NIST.

[2] A. CANTEAUT, M. NAYA-PLASENCIA, B. VAYSSIÈRE.*Sieve-in-the-Middle: Improved MITM Attacks*, in "Advances in Cryptology - CRYPTO 2013, Part I", Lecture Notes in Computer Science, Springer, 2013, vol. 8042, p. 222–240.

[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] 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.

[5] P. CHARPIN, G. 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.

[6] 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^{o} 2248, p. 157–174.

[7] I. DINUR, G. LEURENT.*Improved Generic Attacks Against Hash-based MACs and HAIFA*, in "Advances in Cryptology - CRYPTO 2014", Santa Barbara, CA, United States, LNCS, Springer, August 2014, vol. 8616 [*DOI :* 10.1007/978-3-662-44371-2_9], https://hal.archives-ouvertes.fr/hal-01086177.

[8] J.-C. FAUGÈRE, A. OTMANI, L. PERRET, J.-P. TILLICH.*Algebraic Cryptanalysis of McEliece Variants with Compact Keys*, in "Advances in Cryptology - EUROCRYPT 2010", LNCS, Springer, 2010, n^{o} 6110, p. 279-298, http://dx.doi.org/10.1007/978-3-642-13190-5_14.

[9] P. JOUGUET, S. KUNZ-JACQUES, A. LEVERRIER, P. GRANGIER, E. DIAMANTI.*Experimental demonstration of long-distance continuous-variable quantum key distribution*, in "Nature Photonics", 2013, vol. 7, p. 378-381 [*DOI :* 10.1038/NPHOTON.2013.63], https://hal.archives-ouvertes.fr/hal-00798855.

[10] R. MISOCZKI, J.-P. TILLICH, N. SENDRIER, P. S. 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] V. LALLEMAND.*Cryptanalysis of symmetric ciphers*, Université Pierre et Marie Curie - Paris VI, October 2016, https://hal.inria.fr/tel-01405436.

Articles in International Peer-Reviewed Journal

[12] D. AHARONOV, A. CHAILLOUX, M. GANZ, I. KERENIDIS, L. MAGNIN.*A simpler proof of existence of quantum weak coin flipping with arbitrarily small bias*, in "SIAM Journal on Computing", May 2016, 48 [*DOI :* 10.1137/14096387X], https://hal.inria.fr/hal-01094114.

[13] C. BOURA, A. CANTEAUT, L. R. KNUDSEN, G. LEANDER.*Reflection ciphers*, in "Designs, Codes and Cryptography", January 2016, p. 1-23 [*DOI :* 10.1007/S10623-015-0143-X], https://hal.inria.fr/hal-01237135.

[14] N. CEPAK, P. CHARPIN, E. PASALIC.*Permutations via linear translators*, in "Finite Fields and Their Applications", 2017, https://hal.inria.fr/hal-01412487.

[15] A. CHAILLOUX, G. GUTOSKI, J. SIKORA.*Optimal bounds for quantum weak oblivious transfer*, in "Chicago Journal of Theoretical Computer Science", September 2016 [*DOI :* 10.4086/CJTCS.2016.013], https://hal.archives-ouvertes.fr/hal-00927537.

[16] K. CHAKRABORTY, A. CHAILLOUX, A. LEVERRIER.*Robust Relativistic Bit Commitment*, in "Physical Review A", December 2016 [*DOI :* 10.1103/PHYSREVA.94.062314], https://hal.inria.fr/hal-01409562.

[17] K. CHAKRABORTY, S. SARKAR, S. MAITRA, B. MAZUMDAR, D. MUKHOPADHYAY, E. PROUFF.*Redefining the transparency order*, in "Designs, Codes and Cryptography", 2016 [*DOI :* 10.1007/S10623-016-0250-3], https://hal.archives-ouvertes.fr/hal-01399584.

[18] P. CHARPIN, G. M. KYUREGHYAN.*On sets determining the differential spectrum of mappings*, in "International journal of information and Coding Theory", 2017, Special issue on the honor of Gerard Cohen, https://hal.inria.fr/hal-01406589.

[19] P. CHARPIN, S. MESNAGER, S. SARKAR.*Involutions over the Galois field F2n*, in "IEEE Transactions on Information Theory", 2016, vol. 62, n^{o} 4 [*DOI :* 10.1109/TIT.2016.2526022], https://hal.inria.fr/hal-01272943.

[20] I. DINUR, G. LEURENT.*Improved Generic Attacks Against Hash-Based MACs and HAIFA*, in "Algorithmica", November 2016 [*DOI :* 10.1007/S00453-016-0236-6], https://hal.inria.fr/hal-01407953.

[21] J.-C. FAUGÈRE, A. OTMANI, L. PERRET, F. DE PORTZAMPARC, J.-P. TILLICH.*Folding Alternant and Goppa Codes with Non-Trivial Automorphism Groups*, in "IEEE Transactions on Information Theory", 2016, vol. 62, n^{o} 1, p. 184 - 198 [*DOI :* 10.1109/TIT.2015.2493539], https://hal.inria.fr/hal-01244609.

[22] J.-C. FAUGÈRE, A. OTMANI, L. PERRET, F. DE PORTZAMPARC, J.-P. TILLICH.*Structural Cryptanalysis of McEliece Schemes with Compact Keys*, in "Designs, Codes and Cryptography", April 2016, vol. 79, n^{o} 1, p. 87-112 [*DOI :* 10.1007/S10623-015-0036-Z], https://hal.inria.fr/hal-00964265.

[23] M. KAPLAN, G. LEURENT, A. LEVERRIER, M. NAYA-PLASENCIA.*Quantum Differential and Linear Cryptanalysis*, in "IACR Transactions on Symmetric Cryptology", 2016, vol. 2016, n^{o} 1, https://hal.inria.fr/hal-01237242.

[24] W. MC CUTCHEON, A. PAPPA, B. A. BELL, A. MCMILLAN, A. CHAILLOUX, T. LAWSON, M. S. MAFU, D. MARKHAM, E. DIAMANTI, I. KERENIDIS, J. RARITY, M. TAME.*Experimental verification of multipartite entanglement in quantum networks*, in "Nature Communications", November 2016, vol. 7, 8 [*DOI :* 10.1038/NCOMMS13251], https://hal.inria.fr/hal-01409559.

[25] I. MÁRQUEZ-CORBELLA, R. PELLIKAAN.*A characterization of MDS codes that have an error correcting pair*, in "Finite Fields and Their Applications", 2016, vol. 40, p. 224 - 245 [*DOI :* 10.1016/J.FFA.2016.04.004], https://hal.inria.fr/hal-01408412.

Invited Conferences

[26] C. BOURA, A. CANTEAUT.*Another view of the division property*, in "Symmetric Cryptography (Dagstuhl Seminar 16021)", Dagstuhl, Germany, January 2016, https://hal.inria.fr/hal-01401320.

[27] A. CANTEAUT.*Algebraic Distinguishers against Symmetric Primitives*, in "Paris Crypto Day", Paris, France, June 2016, https://hal.inria.fr/hal-01401286.

[28] A. CANTEAUT.*Chiffrer mieux pour (dé)chiffrer plus*, in "Conférence d'Informatique de l'ENS", Paris, France, April 2016, https://hal.inria.fr/hal-01401333.

[29] A. CANTEAUT.*Comment concevoir un algorithme de chiffrement sûr et efficace : l'héritage de Shannon*, in "Théorie de l'information : nouvelles frontières (dans le cadre du centenaire de Claude Shannon)", Paris, France, IHP, October 2016, https://hal.inria.fr/hal-01401325.

[30] A. CANTEAUT, S. CARPOV, C. FONTAINE, T. LEPOINT, M. NAYA-PLASENCIA, P. PAILLIER, R. SIRDEY.*Stream Ciphers: A Practical Solution for Efficient Homomorphic-Ciphertext Compression*, in "CryptoAction Symposium 2016", Budapest, Hungary, April 2016, https://hal.inria.fr/hal-01401328.

[31] A. CHAILLOUX.*Cryptographie relativiste*, in "CCA 2016", Paris, France, July 2016, https://hal.inria.fr/hal-01409564.

[32] V. LALLEMAND.*Cryptanalysis of the FLIP Family of Stream Ciphers*, in "Paris Crypto Day", Paris, France, September 2016, https://hal.inria.fr/hal-01405423.

[33] G. LEURENT.*Breaking Symmetric Cryptosystems Using Quantum Period Finding*, in "TCCM-CACR 2016", Yinchuan, China, August 2016, https://hal.inria.fr/hal-01407929.

[34] G. LEURENT.*Transcript Collision Attacks*, in "Symmetric Cryptography (Dagstuhl Seminar 16021)", Dagstuhl, Germany, January 2016, https://hal.inria.fr/hal-01407921.

[35] A. LEVERRIER.*Distributing Secret Keys with Quantum Continuous Variables*, in "Recent Advances in Continuous-variable Quantum Information Theory", Barcelone, Spain, April 2016, https://hal.inria.fr/hal-01407434.

[36] A. LEVERRIER.*Quantum Expander Codes*, in "Beyond i.i.d. in Information Theory", Barcelone, Spain, July 2016, https://hal.inria.fr/hal-01407431.

[37] J.-P. TILLICH.*Attaining the capacity with Reed-Solomon codes through the* $(U|U+V)$*construction and Koetter-Vardy soft decoding*, in "Journée Claude Shannon", Paris, France, François Baccelli, Marc Lelarge, July 2016, https://hal.inria.fr/hal-01413503.

International Conferences with Proceedings

[38] M. BARDET, J. CHAULET, V. DRAGOI, A. OTMANI, J.-P. TILLICH.*Cryptanalysis of the McEliece Public Key Cryptosystem based on Polar Codes*, in "Post-Quantum Cryptography - PQCrypto 2016", Fukuoka, Japan, T. TAKAGI (editor), LNCS - Lecture Notes in Computer Science, Springer, February 2016, vol. 9606 [*DOI :* 10.1007/978-3-319-29360-8_9], https://hal.inria.fr/hal-01240856.

[39] M. BARDET, V. DRAGOI, A. OTMANI, J.-P. TILLICH.*Algebraic properties of polar codes from a new polynomial formalism*, in "International Symposium on Information Theory ISIT 2016", Barcelona, Spain, July 2016, p. 230 - 234 [*DOI :* 10.1109/ISIT.2016.7541295], https://hal.inria.fr/hal-01410210.

[40] K. BHARGAVAN, G. LEURENT.*On the Practical (In-)Security of 64-bit Block Ciphers: Collision Attacks on HTTP over TLS and OpenVPN*, 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.

[41] *Best Paper*
K. BHARGAVAN, G. LEURENT.*Transcript Collision Attacks: Breaking Authentication in TLS, IKE, and SSH*, in "Network and Distributed System Security Symposium – NDSS 2016", San Diego, United States, February 2016 [*DOI :* 10.14722/NDSS.2016.23418], https://hal.inria.fr/hal-01244855.

[42] C. BOURA, A. CANTEAUT.*Another View of the Division Property*, in "Crypto 2016 (part I) - 36th Annual International Cryptology Conference", Santa Barbara, United States, Lecture Notes in Computer Science, Springer, August 2016, vol. 9814, p. 654 - 682 [*DOI :* 10.1007/978-3-662-53018-4_24], https://hal.inria.fr/hal-01401016.

[43] C. BOURA, A. CHAKRABORTI, G. LEURENT, G. PAUL, D. SAHA, H. SOLEIMANY, V. SUDER.*Key Recovery Attack Against 2.5-Round Pi-Cipher*, in "FSE 2016 - 23rd International Conference Fast Software Encryption", Bochum, Germany, T. PEYRIN (editor), LNCS - Lecture Notes in Computer Science, Springer, March 2016, vol. 9783, p. 535 - 553 [*DOI :* 10.1007/978-3-662-52993-5_27], https://hal.inria.fr/hal-01404164.

[44] A. CANTEAUT, S. CARPOV, C. FONTAINE, T. LEPOINT, M. NAYA-PLASENCIA, P. PAILLIER, R. SIRDEY.*Stream ciphers: A Practical Solution for Efficient Homomorphic-Ciphertext Compression*, in "FSE 2016 : 23rd International Conference on Fast Software Encryption", Bochum, Germany, Fast Software Encryption 23rd International Conference, FSE 2016, Bochum, Germany, March 20-23, 2016,, Springer, March 2016, vol. 9783 - LNCS (Lecture Notes in Computer Science), p. 313-333 [*DOI :* 10.1007/978-3-662-52993-5_16], https://hal.archives-ouvertes.fr/hal-01280479.

[45] A. CANTEAUT, Y. ROTELLA.*Attacks Against Filter Generators Exploiting Monomial Mappings*, in "Fast Software Encrytion - FSE 2016", Bochum, Germany, Lecture Notes in Computer Science, Springer, March 2016, vol. 9783, p. 78 - 98 [*DOI :* 10.1007/978-3-662-52993-5_5], https://hal.inria.fr/hal-01401009.

[46] R. CANTO TORRES, N. SENDRIER.*Analysis of Information Set Decoding for a Sub-linear Error Weight*, in "Post-Quantum Cryptography - PQCrypto 2016", Fukuoka, Japan, February 2016, https://hal.inria.fr/hal-01244886.

[47] J. CHAULET, N. SENDRIER.*Worst case QC-MDPC decoder for McEliece cryptosystem*, in "IEEE International Symposium on Information Theory, ISIT 2016", Barcelone, Spain, ISIT 2016, proceedings, July 2016, 5 [*DOI :* 10.1109/ISIT.2016.7541522], https://hal.inria.fr/hal-01408633.

[48] S. DUVAL, V. LALLEMAND, Y. ROTELLA.*Cryptanalysis of the FLIP Family of Stream Ciphers*, 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. 9814, p. 457 - 475 [*DOI :* 10.1007/978-3-662-53018-4_17], https://hal.inria.fr/hal-01404145.

[49] P. GABORIT, A. HAUTEVILLE, J.-P. TILLICH.*RankSynd a PRNG Based on Rank Metric*, in "Post-Quantum Cryptography - 7th International Workshop, PQCrypto 2016", Fukuoka, Japan, T. TAKAGI (editor), Lecture Notes in Computer Science, Post-Quantum Cryptography - 7th International Workshop, PQCrypto 2016, Springer, February 2016, vol. 9606, p. 18-28 [*DOI :* 10.1007/978-3-319-29360-8_2], https://hal.inria.fr/hal-01289338.

[50] S. GUERON, N. MOUHA.*Simpira v2: A Family of Efficient Permutations Using the AES Round Function*, in "Advances in Cryptology - ASIACRYPT 2016", Hanoi, Vietnam, Lecture Notes in Computer Science, December 2016, vol. 10031, p. 95-125 [*DOI :* 10.1007/978-3-662-53887-6_4], https://hal.inria.fr/hal-01403414.

[51] 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.

[52] L. KHATI, N. MOUHA, D. VERGNAUD.*Full Disk Encryption: Bridging Theory and Practice*, in "CT-RSA 2017 - RSA Conference Cryptographers' Track", San Francisco, United States, Lecture Notes in Computer Science, February 2017, 16, https://hal.inria.fr/hal-01403418.

[53] B. LAC, M. BEUNARDEAU, A. CANTEAUT, J. J. A. FOURNIER, R. SIRDEY.*A First DFA on PRIDE: from Theory to Practice*, in "International Conference on Risks and Security of Internet and Systems - CRISIS 2016", Roscoff, France, Lecture Notes in Computer Science, September 2016, https://hal.inria.fr/hal-01401271.

[54] G. LEURENT.*Improved Differential-Linear Cryptanalysis of 7-Round Chaskey with Partitioning*, in "EUROCRYPT 2016 - 35th Annual International Conference on the Theory and Applications of Cryptographic", Vienna, Austria, M. FISCHLIN, J.-S. CORON (editors), Springer, May 2016, p. 344 - 371 [*DOI :* 10.1007/978-3-662-49890-3_14], https://hal.inria.fr/hal-01404221.

[55] I. MÁRQUEZ-CORBELLA, E. MARTÍNEZ-MORO.*Betti Numbers and Generalized Hamming Weights*, in "22nd Conference on Applications of Computer Algebra (ACA 2016)", Kassel, Germany, August 2016, https://hal.inria.fr/hal-01409298.

[56] I. MÁRQUEZ-CORBELLA, R. PELLIKAAN.*Is it hard to retrieve an error-correcting pair?*, in "22nd Conference on Applications of Computer Algebra (ACA 2016)", Kassel, Germany, August 2016, https://hal.inria.fr/hal-01409299.

[57] I. MÁRQUEZ-CORBELLA, J.-P. TILLICH.*Using Reed-Solomon codes in the $(U \mid U + V)$ construction and an application to cryptography*, in "International Symposium on Information Theory", Barcelona, Spain, July 2016, https://hal.inria.fr/hal-01410201.

[58] *Best Paper*
A. PHESSO, J.-P. TILLICH.*An Efficient Attack on a Code-Based Signature Scheme*, in "Post-Quantum Cryptography - 7th International Workshop, PQCrypto 2016", Fukuoka, Japan, T. TAKAGI (editor), Lecture Notes in Computer Science, Post-Quantum Cryptography - 7th International Workshop, PQCrypto 2016, Springer, February 2016, vol. 9606, p. 86-103 [*DOI :* 10.1007/978-3-319-29360-8_7], https://hal.inria.fr/hal-01289044.

Conferences without Proceedings

[59] A. LEVERRIER, J.-P. TILLICH, G. ZÉMOR.*Quantum Expander Codes*, in "19th International Conference on Quantum Information Processing", Banff, Canada, January 2016, https://hal.inria.fr/hal-01244685.

[60] J.-P. TILLICH.*Attaining the capacity with Reed-Solomon codes through the $(U|U+V)$construction and Koetter-Vardy soft decoding*, in "CohenFest 2016", Paris, France, July 2016, https://hal.inria.fr/hal-01413506.

Scientific Books (or Scientific Book chapters)

[61] P. CHARPIN, S. MESNAGER, S. SARKAR.*Dickson Polynomials that are Involutions*, in "Contemporary Developments in Finite Fields and Their Applications", A. CANTEAUT, G. EFFINGER, S. HUCZYNSKA, D. PANARIO, L. STORME (editors), World Scientific Press, 2016, p. 22-45 [*DOI :* 10.1142/9789814719261_0003], https://hal.inria.fr/hal-01237332.

Books or Proceedings Editing

[62] A. CANTEAUT, G. EFFINGER, S. HUCZYNSKA, D. PANARIO, L. STORME (editors). *Contemporary Developments in Finite Fields and Applications* , World Scientific, August 2016, 362 [*DOI :* 10.1142/9762], https://hal.inria.fr/hal-01401266.

[63] P. CHARPIN, T. JOHANSSON, G. M. KYUREGHYAN, N. SENDRIER, J.-P. TILLICH (editors). *Special issue on coding and cryptography*, Design, Codes and Cryptography - Special issue on coding and cryptography, Springer, 2016 [*DOI :* 10.1007/s10623-016-0307-3], https://hal.archives-ouvertes.fr/hal-01406954.

Research Reports

[64] A. CANTEAUT, S. DUVAL, L. PERRIN.*A generalisation of Dillon's APN permutation with the best known differential and linear properties for all fields of size* 2^{4k+2}, IACR Cryptology ePrint Archive, September 2016, n° 2016/887, 29, https://hal.inria.fr/hal-01401245.

Scientific Popularization

[65] A. CANTEAUT.*On the Origin of Trust: Struggle for Secure Cryptography*, in "Dot Security 2016", Paris, France, April 2016, https://hal.inria.fr/hal-01401311.

[66] A. CHAILLOUX.*L'ordinateur quantique*, in "Art, cerveau, futur", Mouans Sartoux, France, September 2016, https://hal.inria.fr/hal-01409565.

[67] N. SENDRIER, J.-P. TILLICH.*Code-Based Cryptography: New Security Solutions Against a Quantum Adversary*, in "ERCIM News", July 2016, vol. Special Theme Cybersecurity, n° 106, https://hal.archives-ouvertes.fr/hal-01410068.

Other Publications

[68] X. BONNETAIN.*Cryptanalyse quantique de primitives symétriques*, Télécom ParisTech ; Paris Diderot, September 2016, https://hal.inria.fr/hal-01409206.

[69] R. BRICOUT.*Protocole de mise en gage de bit relativiste*, MPRI, September 2016, https://hal.inria.fr/hal-01419367.

[70] R. BRICOUT, A. CHAILLOUX.*Recursive cheating strategies for the relativistic F_Q bit commitment protocol*, August 2016, working paper or preprint, https://hal.inria.fr/hal-01409563.

[71] K. CHAKRABORTY, A. CHAILLOUX, A. LEVERRIER.*Robust Relativistic Bit Commitment*, October 2016, International Conference for Young Quantum Information Scientists, Poster, https://hal.inria.fr/hal-01409527.

[72] K. CHAKRABORTY, A. CHAILLOUX, A. LEVERRIER.*Robust Relativistic Bit Commitment*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01407421.

[73] A. COUVREUR, I. MÁRQUEZ-CORBELLA, R. PELLIKAAN.*Cryptanalysis of McEliece Cryptosystem Based on Algebraic Geometry Codes and their subcodes*, March 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01280927.

[74] T. DEBRIS.*Décodage Statistique*, MPRI, September 2016, https://hal.inria.fr/hal-01413092.

[75] G. KACHIGAR.*Étude et conception d'algorithmes quantiques pour le décodage de codes linéaires* , Université de Rennes 1, France, September 2016, 127, https://hal.inria.fr/hal-01371018.

Team SERENA

Simulation for the Environment: Reliable and Efficient Numerical Algorithms

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
Earth, Environmental and Energy Sciences

Table of contents

Team SERENA

Creation of the Team: 2015 June 01

Keywords:

Computer Science and Digital Science:

2.1.3. - Functional programming
2.4.3. - Proofs
6.1.1. - Continuous Modeling (PDE, ODE)
6.1.5. - Multiphysics modeling
6.2.1. - Numerical analysis of PDE and ODE
6.2.5. - Numerical Linear Algebra
6.2.8. - Computational geometry and meshes
6.3.1. - Inverse problems
6.3.4. - Model reduction
6.3.5. - Uncertainty Quantification

Other Research Topics and Application Domains:

3.1. - Sustainable development
3.3.1. - Earth and subsoil
3.4.2. - Industrial risks and waste
3.4.3. - Pollution
4.2.1. - Fission
5.5. - Materials

The publications of the two members of SERENA from ENPC are not listed in the bibliography, since all administrative points of the creation of the SERENA team are not finalized for the moment. Consequently, A. Ern and L. Monasse as well as the Ph.D. students and post-docs from ENPC do not have yet the SERENA affiliation registered at the HAL preprint server.

1. Members

Research Scientists

Martin Vohralík [Team leader, Inria, Senior Researcher, HDR]
François Clément [Inria, Researcher]
Alexandre Ern [ENPC, Professor, HDR]
Michel Kern [Inria, Researcher]
Laurent Monasse [ENPC, Associate Professor]
Géraldine Pichot [Inria, Researcher]
Iain Smears [Inria, Starting Research position]
Pierre Weis [Inria, Senior Researcher]

PhD Students

Sarah Ali Hassan [Inria, Contract with ANDRA]
Amina Benaceur [Cifre EDF]
Nabil Birgle [Inria, until March 24th, 2016, Contract with ANDRA]
Pierre Cantin [Cifre EDF, until November 14th, 2016]
Karol Cascavita [ENPC, Labex MMCD]

Fatma Cheikh [ENIT-Lamsin (Tunisia), until October 12th, 2016]
Jad Dabaghi [Inria, ERC GATIPOR]
Patrik Daniel [Inria, ERC GATIPOR]
Cédric Josz [Cifre RTE, until July 13th, 2016]
Frédéric Marazzato [since October 1st, 2016, Contract with CEA]
Yannick Masson [ENPC, Labex MMCD]
Nicolas Pignet [Cifre EDF, since November 1st, 2016]
Mohamed Riahi [ENIT-Lamsin (Tunisia), until October 12th, 2016]
Rita Riedlbeck [Cifre EDF]

Post-Doctoral Fellows

Elyes Ahmed [Univ. Paris XIII, ANR DEDALES]
Thomas Boiveau [ENPC, since May 1st, 2016]
Matteo Cicuttin [ENPC, since February 1st, 2016]

Visiting Scientist

Markus Köppel [University of Stuttgart, PhD Student, Visiting Scientist]

Administrative Assistants

Cindy Crossouard [Inria, until November 30, 2016]
Virginie Collette [Inria, from December 1st, 2016]

Others

Hend Ben Ameur [IPEST and ENIT-Lamsin (Tunisia), Professor, External Collaborator, HDR]
Guy Chavent [Univ. Paris IX (retired), Professor Emeritus, External Collaborator, HDR]
Jérôme Jaffré [Inria (retired), External Collaborator, HDR]
Caroline Japhet [Univ. Paris XIII, Associate Professor, External Collaborator]
Vincent Martin [UT Compiègne, Associate Professor, External Collaborator]
Zoubida Mghazli [Université Ibn Tofail, Kenitra, Morocco, Professor, External Collaborator]
Jean-Elizabeth Roberts [Inria (retired), External Collaborator, HDR]

2. Overall Objectives

2.1. Overall Objectives

The 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.

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.

3.2. Structure-preserving discretizations 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 and IFP Energies Nouvelles). Our requirement is to derive *structure-preserving* methods, i.e., methods that mimic at the discrete level fundamental properties of the underlying PDEs, such as conservation principles and preservation of invariants. Here, the theoretical questions are closely linked to *differential geometry* for the lowest-order schemes. For the schemes we develop, we study existence, uniqueness, and stability questions, and derive a priori convergence estimates. Our special interest is in higher-order methods like the hybrid high-order method, which have recently begun to receive significant attention. Even though their use in practice may not be immediate, we believe that they represent the future generation of numerical methods for industrial simulations.

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. This in particular yields 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.

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 shall then address theoretically the question of convergence of the new algorithms and prove their numerical quasi-optimality, meaning that they need, up to a generic constant, the smallest possible number of degrees of freedom to achieve the given accuracy. We in particular seek to prove a guaranteed error reduction in terms of the number of degrees of freedom.

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. Environmental problems

We pursue *applications* of our theoretical results to current challenging *environmental problems* with numerous *academic collaborators* and with *industrial partners* such as ANDRA, IFP Energies Nouvelles, CEA, and EDF. We are traditionally interested in *porous media* for multiphase flows and transport of contaminants in the subsurface and concentrate on fractures, fracture networks, fractured porous media, subsurface depollution after chemical leakage, nuclear waste disposal in deep underground repositories, and geological sequestration of CO_2. Among our newer themes, we count complex inviscid flows interacting with a mechanical deformable structure and Navier–Stokes flows. Such problems are encountered in energy production (operation of nuclear reactors) and safety assessment (shock waves resulting from an explosion impinging on a structure).

5. New Software and Platforms

5.1. GEOFRACFLOW

GEOFRACFLOW

SCIENTIFIC DESCRIPTION GEOFRACFLOW 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.

- Participants: Géraldine Pichot, Jocelyne Erhel, and Jean-Raynald De Dreuzy
- Contact: Géraldine Pichot
- URL: https://bil.inria.fr/fr/software/view/2653/tab

5.2. SBM

Skew Brownian Motion

SCIENTIFIC 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.

- Participants: Antoine Lejay, Géraldine Pichot
- Contact: Antoine Lejay
- URL: https://gforge.inria.fr/projects/sbm

5.3. Sklml

The OCaml parallel skeleton system

SCIENTIFIC DESCRIPTION Writing parallel programs is not easy, and debugging them is usually a nightmare. To cope with these difficulties, the skeleton programming approach uses a set of predefined patterns for parallel computations. The skeletons are higher-order functional templates that describe the program underlying parallelism. Sklml is a new framework for parallel programming that embeds an innovative compositional skeleton algebra into the OCaml language. Thanks to its skeleton algebra, Sklml provides two evaluation regimes to programs: a regular sequential evaluation (merely used for prototyping and debugging) and a parallel evaluation obtained via a recompilation of the same source program in parallel mode. Sklml was specifically designed to prove that the sequential and parallel evaluation regimes coincide.
FUNCTIONAL DESCRIPTION Sklml is a functional parallel skeleton compiler and programming system for OCaml programs. Slogan is “easy coarse grain parallelization”.

- Participants: Pierre Weis and François Clément
- Contact: François Clément
- URL: http://sklml.inria.fr

6. New Results

6.1. Numerical algorithms for simulating diffusion processes in discontinuous media

Participant: Géraldine Pichot.

Grants: H2MN04 3

Software: SBM 5.2

Publications: [19]

We present several benchmark tests for Monte Carlo methods simulating diffusion in one-dimensional discontinuous media. These benchmark tests aim at studying the potential bias of the schemes and their impact on the estimation of micro- or macroscopic quantities (repartition of masses, fluxes, mean residence time,...). These benchmark tests are backed by a statistical analysis to filter out the bias from the unavoidable Monte Carlo error. We apply them on four different algorithms. The results of the numerical tests give a valuable insight of the fine behavior of these schemes, as well as rules to choose between them.

6.2. Locally space-time efficient estimates for parabolic problems

Participants: Martin Vohralík, Alexandre Ern, Iain Smears.

Grants: GATIPOR 8.3.1

Publications: [33]

In [33], we derive for the first time a posteriori error estimates for parabolic problems which are both globally reliable and locally space-time efficient. By this, one means that the error between a known approximate numerical solution and the unknown exact solution of a model parabolic PDE (the heat equation) is bounded from above on the whole space-time domain by a fully computable estimator, while this estimator does not overestimate significantly the error and localizes it both in space and in time. More precisely, the estimator also gives lower bounds on the error, up to a generic constant, and this on each time interval and in a small neighborhood of each space mesh element. We consider arbitrarily high-order conforming Galerkin spatial discretizations and arbitrarily high-order discontinuous Galerkin temporal discretizations, and the error is measured in a norm composed of the $L^2(H^1) \cap H^1(H^{-1})$-norm augmented by the temporal jumps of the numerical solution. The efficiency constant is robust with respect to (independent of) any mesh-size, time-step size, and the spatial and temporal polynomial degrees. The proposed estimators also have the practical advantage of not imposing any requirement on coarsening between the consecutive time steps.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- Three-parts contract Inria–EDF–Sciworks Technologies (November 2015–April 2016) on "Form-L for the formalization of constraints of complex systems".
- Contract Inria–IFP Energies Nouvelles (December 2016–December 2017) on "A posteriori error analysis for porous media flow problems with fractures".
- Numerous contracts accompanying Ph.D. theses and post-doc positions, see Section 1.

8. Partnerships and Cooperations

8.1. Regional Initiatives

GT Elfic (Labex DigiCosme, 2014–2016): "Formal proof for finite element programs", with TOCCATA (Inria Saclay - Île-de-France), CEA LIST, LIPN (Université de Paris 13), and LMAC (Université de Technologie de Compiègne).

8.2. National Initiatives

8.2.1. ANR

ANR DEDALES: "Algebraic and geometric domain decomposition for subsurface flow". The project aims at developing high performance software for the simulation of two phase flow in porous media. It specifically targets parallel computers where each node is itself composed of a large number of processing cores, such as are found in new generation many-core architectures. The project had its intermediate review in December 2016, and received excellent marks from the expert panel.

The partners are HIEPACS, Laboratoire Analyse, Géométrie et Application, University Paris 13, Maison de la Simulation, and ANDRA. SERENA representants are M. Kern (grant leader) and M. Vohralík, period 2014–2017.

ANR GEOPOR: "Geometrical approach for porous media flows: theory and numerics". A new approach to numerical methods for multiphase simulations based on the concept of gradient flows is investigated. With Laboratoire Jacques-Louis Lions, University Pierre and Marie Curie. SERENA representant is M. Vohralík, period 2013–2017.

ANR H2MNO4: "Original optimized object-oriented numerical model for heterogeneous hydrogeology". The project H2MNO4 develops numerical models for reactive transport in heterogeneous media. The objective is to design both Eulerian and Lagrangian models. Three applications are concerned: freshwater supply, remediation of mine drainage, and waste geological disposal. The project relies on a consortium of six partners, involving four public research laboratories (Inria, Geosciences Rennes, University of Lyon 1, University of Poitiers, Pprime Institute), one public institution (ANDRA), and one enterprise (ITASCA). International collaborations are pursued with University of San Diego (USA) and UPC (Spain). SERENA representant is G. Pichot, period 2012–2016.

ANR HHOMM: "Hybrid high-order methods on polyhedral meshes", Theoretical foundations and applications (up to software development) for the recently-devised Hybrid high-order methods. Coordinated by D. Di Pietro, University of Montpellier. SERENA representant is A. Ern, period 2015–2019.

C2S@Exa: "Computer and Computational Sciences at Exascale". This is an Inria Project Lab (IPL). This national initiative aims at the development of numerical modeling methodologies that fully exploit the processing capabilities of modern massively parallel architectures in the context of a number of selected applications related to important scientific and technological challenges for the quality and the security of life in our society. This project supported in particular the Ph.D. of N. Birgle in the framework of the Inria–ANDRA collaboration.

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

ERC GATIPOR: “Guaranteed fully adaptive algorithms with tailored inexact solvers for complex porous media flows”. The subject of this project 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), period 2015–2020.

EoCoE: “Energy Oriented Center of Excellence” This project is coordinated by Maison de la Simulationand gathers 23 partners from 13 countries to use the tremendous potential offered by the ever-growing computing infrastructure to foster and accelerate the European transition to a reliable low carbon energy supply using HPC (High Performance Computing). SERENA representant M. Kern, period 2015–2018.

8.3.2. Collaborations in European Programs, Except FP7 & H2020

8.3.2.1. ITEA 3

Program: ITEA 3

Project acronym: OPENCPS

Project title: Open cyber-physical system model-driven certified development

Duration: Dec 2015–Dec 2018

Coordinator: Magnus Eek

Other partners: AB SKF, CEA, ELTE-Soft Kft., ESI Group, EDF, Wqua Simulation AB, Ericsson, IncQuery Labs Kft., KTH, Linköping University, RTE, SICS, SIREHNA, Saab AB, Sherpa Engineering, Siemens Industrial Torbumachinery AB, VTT Technical Research Center of Finland Ltd.

Abstract: Cyber-physical systems put increasing demands on reliability, usability, and flexibility while, at the same time, lead time and cost efficiency are essential for industry competitiveness. Tools and environments for model-based development of cyber-physical systems are becoming increasingly complex and critical for the industry: tool interoperability, vendor lock-ins, and tool life-cycle support are some of the challenges. The project focuses on interoperability between the standards Modelica/UML/FMI, improved execution speed of (co-)simulation, and certified code generation.

8.3.2.2. ERC CZ

Program: Research, Development and Innovation Council of the Czech Republic

Project acronym: MoRe

Project title: Implicitly constituted material models: from theory through model reduction to efficient numerical methods

Duration: September 2012 – September 2017

Coordinator: Josef MÁLEK, Charles University in Prague. SERENA representant is M. Vohralík.

Other partners: Institute of Mathematics, Czech Academy of Sciences; University of Oxford

Abstract: A multidisciplinary project on nonlinear Navier–Stokes flows with implicit constitutive laws. It focuses on development of accurate, efficient, and robust numerical methods for simulations of the new class of implicit models.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

H. Ben Ameur, Professor at IPEST and member of ENIT-Lamsin, Tunis, Tunisia, November 1–15, 2016.

G. Hammond, Applied Systems Analysis and Research Sandia National Laboratories, USA, April 18, 2016.

M. Köppel, Ph.D. student, University of Stuttgart, Germany, October 1–December 31, 2016.

Z. Strakoš, Professor at the Charles University in Prague, April, 17–21, 2016.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

A. Ern co-organized with D. Di Pietro and L. Formaggia the IHP Thematic Quarter on Numerical methods for PDEs from September 5th to December 16th, 2016.

M. Vohralík co-organized, together with I. Vignon-Clémentel from the project-team REO, the monthly *Scientific computing, modeling, and numerical analysis* seminar ("Rencontres Inria-LJLL en calcul scientifique"), see the web page https://project.inria.fr/rencontresljll/, until July, 2016. Since September, 2016, I. Smears co-organizes this seminar together with C. Grandmont from the project-team REO.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

A. Ern and M. Vohralík are members of the Scientific Committee of the ENUMATH 2017 conference (Voss, Norway).

M. Vohralík is a member of the Scientific Committee of the Finite Volumes for Complex Applications 8 conference (Lille, France).

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

A. 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, Computational Methods in Applied Mathematics, and Journal de l'Ecole polytechnique, Mathématiques.

M. Vohralík is a member of the editorial boards of SIAM Journal on Numerical Analysis and of Acta Polytechnica.

9.1.3.2. Reviewer - Reviewing Activities

M. Kern was a reviewer for the journal Mathematics and Computers in Simulation, Electron. Trans. Numer. Anal., and Oil and Gas Science and Technology,

L. Monasse was a reviewer for J. Comput. Phys. and J. Comput. Particle Mech.

A. Ern and M. Vohralík served as reviewers for tens of papers in different journals.

9.1.4. Invited Talks

A. Ern, plenary speaker, ECCOMAS 2016, Crete.

A. Ern, plenary speaker, WONAPDE 2016, Concepción, Chile.

M. Kern, minisymposium speaker at the German Priority Programme, Software for Exascale Computing, Germany.

M. Vohralík, invited speaker, Adaptive algorithms for computational PDEs, Birmingham, Great Britain.

M. Vohralík, plenary speaker, WONAPDE 2016, Concepción, Chile.

M. Vohralík, plenary speaker, ALGORITMY 2016, Podbanske, Slovakia.

9.1.5. Leadership within the Scientific Community

M. Kern was a member of the nominating committee for the SIAM Activity Group on Geosciences.

M. Kern is a reviewer for the German Supercomputing Center JARA program.

M. Kern is a member of the Scientific Committee of Orap (ORganisation Associative du Parallélisme), of the Scientific Board of GDR Calcul and of the jury and executive board of Label C3I.

M. Vohralík is a member of the steering committee of GIS Géosciences franciliennes.

9.1.6. Research Administration

F. Clément is a member of the *Comité local d'hygiène, de sécurité et des conditions de travail* of the Inria Research Center of Paris.

F. Clément is the *correspondant Inria-entreprise* of the Inria Research Center of Paris for AMIES.

M. Kern is Deputy Director of Maison de la Simulation, a joint project between CEA, CNRS, Inria, Université de Paris 11, and Université de Versailles, focused on applications of high end computing.

M. Kern is a member of the *Comité de site* of the Inria center of Paris.

G. Pichot is a member of the *Comité local d'hygiène, de sécurité et des conditions de travail* of the Inria center of Paris.

G. Pichot is member of the Conseil de département MAM of Polytech Lyon.

G. Pichot is member of the Commission de developpement technologique (CDT) of the Inria center of Paris.

G. Pichot is a member of the CES commission of the Inria center of Paris.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

A. Ern, Analyse numérique et optimisation, 78h, L3, Ecole Polytechnique (professeur chargé de cours), France.

A. Ern, Méthodes de Galerkine Discontinu, 20h, M2 Mathématiques de la modélisation, Paris 6, France.

M. Kern, Modélisation et simulation des écoulements de fluides dans la géosphère (with E. Mouche, CEA), 30h, M2 Mathématiques et Applications (parcours Analyse, Modélisation et Simulation), Université Paris Saclay, France.

M. Kern, Eléments finis (avec D. Ryckelynck), 30h, 2nd year students, Ecole Mines-ParisTech, France.

M. Kern, Problèmes inverses, 24 h, 2nd year students, Ecole Mines-ParisTech, France.

M. Kern, Analyse numérique avancée, 20h, 3rd year students, MACS, Université Paris Nord, France.

L. Monasse, Analyse et Calcul Scientifique, 30h, L3, ENPC, France.

M. Vohralík, A posteriori error estimates for efficiency and error control in numerical simulations, 36h, M2, Charles University in Prague, Czech Republic.

M. Vohralík, A posteriori error estimates and adaptive error components balancing in numerical simulations, summer school "IHP quarter on Numerical Methods for PDEs", Cargèse, Corsica & Paris, France, 15h.

9.2.2. Supervision

PhD: N. Birgle, *Underground flow, numerical methods and high performance computing*, University Paris VI, defended on March 24th, 2016, advisor J. Jaffré, co-advisor J. E. Roberts.

PhD: F. Cheikh, *Identification of fractures in a porous medium by a method of indicators*, University Paris VI and University of Tunis El Manar, defended on October 12th, 2016, advisors J. E. Roberts and H. Ben Ameur, co-advisors V. Martin and F. Clément.

PhD: M. H. Riahi, *Identification of hydrogeological parameters in a porous medium*, University Paris VI and University of Tunis El Manar, defended on October 12th, 2016, advisors J. Jaffré and H. Ben Ameur.

PhD in progress: S. Ali Hassan, *A posteriori error estimates and stopping criteria for domain decomposition solvers with local time stepping*, University Paris VI, November 2013, advisor M. Vohralík, co-advisors C. Japhet and M. Kern.

PhD in progress: J. Dabaghi, *Adaptive modeling via complementarity of phase appearance and disappearance in fractured and porous media*, University Paris VI, November 2015, advisor M. Vohralík, co-advisor V. Martin.

PhD in progress: P. Daniel, *Adaptive multilevel solvers with a posteriori error control for porous media flows* University Paris VI, October 2015, advisor M. Vohralík, co-advisor A. Ern.

PhD students at ENPC are listed in Section 1.

9.2.3. Juries

A. Ern, external examiner of the PhD of Z. Dong, University of Leicester, November 24, 2016.

A. Ern, chair of jury of the PhD of M. Giacomini, Ecole Polytechnique, December 9, 2016.

M. Kern, jury member for the PhD of V. Groza, Identification de paramètres et analyses de sensibilité pour un modèle d'usinage par jet d'eau abrasif, Université de Nice, November 9, 2016.

M. Kern, jury member for the PhD of M. Massaro, Méthodes numériques pour les plasmas sur architectures multicoeurs, Université de Strasbourg, December 16, 2016.

M. Kern, jury member for the Habilitation of Y. Mesri, Méthodes numériques massivement parallèles à base de maillages non structurés adaptatifs et anisotropes pour la mécanique numérique, Université de Nice, December 19, 2016.

G. Pichot, jury member of the PhD of A. Botella, Unstructured volumetric meshing of geological models for physical phenomenon simulations, University of Lorraine, April, 1st, 2016.

G. Pichot, jury member of the PhD of A. Dartois, Study of the macro-dispersion of inert particles in highly heterogeneous 3D porous media, University of Poitiers, December, 14, 2016.

M. Vohralík, reviewer and jury member of the PhD of R. Tittarelli, University Lille 1, September 27, 2016.

M. Vohralík, chair of jury of the PhD of M. Groza, Université de Nice, November 10, 2016.

9.3. Popularization

F. Clément was member of the Organizing Committee of the *17e Salon Culture & Jeux Mathématiques*, held in Paris, 26–29 May, 2016. He was member of the Editorial Board of the *Maths Société Express* booklet distributed during the exhibition. He was also coordinator of the Maths-Enterprises booth for AMIES.

F. Clément realized, with the Communication Department of the Inria Research Center of Paris, an exhibition illustrating the results of the *Etude de l'impact socio-économique des mathématiques en France* sponsored by AMIES, FSMP, and FMJH.

F. Clément coordinated an article about AMIES in the magazine PLOT published by the *Association des Professeurs de Mathématiques de l'Enseignement Public*.

Major publications by the team in recent years: [1], [2], [3], [4], [5], [6], [7], [8], [9], [10].

10. 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] A. ERN, M. VOHRALÍK.*Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs*, in "SIAM J. Sci. Comput.", 2013, vol. 35, n^{o} 4, p. A1761–A1791, http://dx.doi.org/10.1137/120896918.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] I. SMEARS.*Robust and efficient preconditioners for the discontinuous Galerkin time-stepping method*, in "IMA Journal of Numerical Analysis", October 2016 [*DOI :* 10.1093/IMANUM/DRW050], https://hal.archives-ouvertes.fr/hal-01357497.

[10] M. VOHRALÍK, B. I. WOHLMUTH.*Mixed finite element methods: implementation with one unknown per element, local flux expressions, positivity, polygonal meshes, and relations to other methods*, in "Math. Models Methods Appl. Sci.", 2013, vol. 23, n^{o} 5, p. 803–838, http://www.worldscientific.com/doi/abs/10.1142/S0218202512500613.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] N. BIRGLE.*Underground flow, numerical methods and high performance computing*, Université Pierre et Marie Curie - Paris VI, March 2016, https://tel.archives-ouvertes.fr/tel-01366054.

[12] F. CHEIKH. *Identification of fractures in porous medium*, UPMC - Université Paris 6 Pierre et Marie Curie ; ENIT Tunis, October 2016, https://hal.inria.fr/tel-01415825.

[13] M. H. RIAHI.*Identification of hydrogeological parameters in a porous medium*, Universite Pierre & Marie Curie, October 2016, https://hal.inria.fr/tel-01423829.

Articles in International Peer-Reviewed Journal

[14] P. F. ANTONIETTI, P. HOUSTON, I. SMEARS.*A Note on Optimal spectral bounds for nonoverlapping domain decomposition preconditioners for hp–Version Discontinuous Galerkin methods*, in "International Journal of Numerical Analysis and Modeling", 2016, vol. 13, n^o 4, p. 513-524, https://hal.inria.fr/hal-01428749.

[15] E. CANCÈS, G. DUSSON, Y. MADAY, B. STAMM, M. VOHRALÍK.*A perturbation-method-based post-processing for the planewave discretization of Kohn–Sham models*, in "Journal of Computational Physics", February 2016, vol. 307, p. 446–459 [*DOI :* 10.1016/J.JCP.2015.12.012], http://hal.upmc.fr/hal-01140818.

[16] V. DOLEJŠÍ, A. ERN, M. VOHRALÍK.*hp-adaptation driven by polynomial-degree-robust a posteriori error estimates for elliptic problems*, in "SIAM Journal on Scientific Computing", October 2016, vol. 38, n^o 5, p. A3220-A3246, https://hal.inria.fr/hal-01165187.

[17] T.-T.-P. HOANG, C. JAPHET, M. KERN, J. E. ROBERTS.*Space-time Domain Decomposition and Mixed Formulation for solving reduced fracture models*, in "SIAM Journal on Numerical Analysis", February 2016, vol. 54, n^o 1, 29 [*DOI :* 10.1137/15M1009651], https://hal.inria.fr/hal-01113953.

[18] C. JOSZ, D. HENRION.*Strong duality in Lasserre's hierarchy for polynomial optimization*, in "Optimization Letters", January 2016, vol. 10, n^o 1, p. 3-10 [*DOI :* 10.1007/S11590-015-0868-5], https://hal.archives-ouvertes.fr/hal-00997726.

[19] A. LEJAY, G. PICHOT.*Simulating Diffusion Processes in Discontinuous Media: Benchmark Tests*, in "Journal of Computational Physics", June 2016, vol. 314, p. 348-413 [*DOI :* 10.1016/J.JCP.2016.03.003], https://hal.inria.fr/hal-01003853.

[20] I. SMEARS.*Robust and efficient preconditioners for the discontinuous Galerkin time-stepping method*, in "IMA Journal of Numerical Analysis", October 2016 [*DOI :* 10.1093/IMANUM/DRW050], https://hal.archives-ouvertes.fr/hal-01357497.

[21] I. SMEARS, E. SÜLI.*Discontinuous Galerkin finite element methods for time-dependent Hamilton-Jacobi-Bellman equations with Cordes coefficients*, in "Numerische Mathematik", May 2016, vol. 133, n^o 1, p. 141 - 176 [*DOI :* 10.1007/S00211-015-0741-6], https://hal.inria.fr/hal-01428647.

International Conferences with Proceedings

[22] S. BOLDO, F. CLÉMENT, F. FAISSOLE, V. MARTIN, M. MAYERO.*A Coq Formal Proof of the Lax–Milgram theorem*, in "6th ACM SIGPLAN Conference on Certified Programs and Proofs", Paris, France, January 2017, https://hal.inria.fr/hal-01391578.

[23] J.-R. DE DREUZY, G. PICHOT, P. LAUG, J. ERHEL.*Flow simulation in 3D Discrete Fracture Networks*, in "The XXI International Conference Computational Methods in Water Resources", Toronto, ON, Canada, June 2016, https://hal.inria.fr/hal-01387391.

Scientific Books (or Scientific Book chapters)

[24] M. KERN.*Méthodes numériques pour les problèmes inverses*, Collection Mathématiques et statisitques, ISTE Éditions, March 2016, 222, https://hal.inria.fr/hal-01297083.

Research Reports

[25] H. BEN AMEUR, G. CHAVENT, F. CHEIKH, F. CLÉMENT, V. MARTIN, J. E. ROBERTS.*First-Order Indicators for the Estimation of Discrete Fractures in Porous Media*, Inria Paris, February 2016, n^{o} RR-8857, 31, https://hal.inria.fr/hal-01279503.

[26] F. CLÉMENT, V. MARTIN.*The Lax-Milgram Theorem. A detailed proof to be formalized in Coq*, Inria Paris, July 2016, n^{o} RR-8934, https://hal.inria.fr/hal-01344090.

[27] G. PICHOT.*TECHNICAL REPORT N° 484 Algorithms for Gaussian random field generation*, Inria Paris, December 2016, n^{o} RT-0484, https://hal.inria.fr/hal-01414707.

[28] M. H. RIAHI, H. BEN AMEUR, J. JAFFRÉ, R. BOUHLILA.*Refinement indicators for estimating hydrogeologic parameters*, Inria Paris, February 2016, n^{o} RR-8877, 25, https://hal.inria.fr/hal-01285127.

Other Publications

[29] L. AMIR, M. KERN.*Preconditioning a coupled model for reactive transport in porous media*, June 2016, working paper or preprint, https://hal.inria.fr/hal-01327307.

[30] J. BLECHTA, J. MÁLEK, M. VOHRALÍK.*Localization of the $W^{-1,q}$ norm for local a posteriori efficiency*, June 2016, working paper or preprint, https://hal.inria.fr/hal-01332481.

[31] E. CANCÈS, G. DUSSON, Y. MADAY, B. STAMM, M. VOHRALÍK.*Guaranteed and robust a posteriori bounds for Laplace eigenvalues and eigenvectors: conforming approximations*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01194364.

[32] A. ERN, I. SMEARS, M. VOHRALÍK.*Discrete p-robust H(div)-liftings and a posteriori estimates for elliptic problems with H^{-1} source terms*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01377007.

[33] A. ERN, I. SMEARS, M. VOHRALÍK.*Guaranteed, locally space-time efficient, and polynomial-degree robust a posteriori error estimates for high-order discretizations of parabolic problems*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01377086.

[34] A. ERN, M. VOHRALÍK.*Stable broken H1 and H(div) polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01422204.

[35] T.-T.-P. HOANG, C. JAPHET, M. KERN, J. E. ROBERTS.*Space-time domain decomposition for advection-diffusion problems in mixed formulations*, March 2016, working paper or preprint, https://hal.inria.fr/hal-01296348.

[36] M. NAIT ABDALLAH.*On the logic resolution of the wave particle duality paradox in quantum mechanics (Extended abstract)*, April 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01303223.

[37] J. PAPEŽ, Z. STRAKOŠ, M. VOHRALÍK.*Estimating and localizing the algebraic and total numerical errors using flux reconstructions*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01312430.

[38] M. ČERMÁK, F. HECHT, Z. TANG, M. VOHRALÍK.*Adaptive inexact iterative algorithms based on polynomial-degree-robust a posteriori estimates for the Stokes problem*, January 2017, working paper or preprint, https://hal.inria.fr/hal-01097662.

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:

1.2.8. - Network security
3.4. - Machine learning and statistics
5.4. - Computer vision
6.2. - Scientific Computing, Numerical Analysis & Optimization
7.1. - Parallel and distributed algorithms
7.3. - Optimization
8.2. - Machine learning

Other Research Topics and Application Domains:

9.4.5. - Data science

1. Members

Research Scientists

Francis Bach [Team leader, Inria, Senior Researcher, HDR]
Alexandre d'Aspremont [CNRS, Senior Researcher, HDR]
Simon Lacoste-Julien [Inria, Starting Research Position, Until Aug 2016]

Technical Staff

Anton Osokin [Inria]
Fabian Pedregosa [Chaire Havas Dauphine]
Kevin Scaman [Inria, from Nov 2016]

PhD Students

Remi Leblond [Inria]
Jean-Baptiste Alayrac [Ecole Polytechnique]
Dmitry Babichev [Inria]
Anaël Bonneton [ENS Paris]
Alexandre Defossez [CIFRE Facebook]
Aymeric Dieuleveut [ENS Paris]
Christophe Dupuy [CIFRE Technicolor]
Nicolas Flammarion [ENS Lyon]
Damien Garreau [Inria]
Anastasia Podosinnikova [Inria, granted by Microsoft Research]
Antoine Recanati [CNRS]
Vincent Roulet [Ecole Polytechnique]
Damien Scieur [Inria]
Tatiana Shpakova [Inria]

Post-Doctoral Fellows

Amit Bermanis [Inria, until Jul 2016]
Nicolas Boumal [Research in Paris, until Jan 2016]
Pascal Germain [Inria]
Robert Gower [Inria, from Aug 2016]
Balamurugan Palaniappan [Inria]

Federico Vaggi [ENS Paris, from May 2016]

Visiting Scientists

Chiranjib Bhattacharyya [Associated Team Bigfoks2, Nov 2016]
Remi Lajugie [ENS Cachan, until Feb 2016]

Administrative Assistant

Lindsay Polienor

Others

Reza Babanezhad Harikandeh [Inria, from Jun 2016 until Sep 2016]
Fajwel Fogel [Chaire Havas Dauphine, until Mar 2016]
Gauthier Gidel [ENS Paris, from Apr 2016]
Samy Jelassi [Inria, from Feb 2016]
Senanayak Karri [Inria]
Horia Mania [Chaire Havas Dauphine, from May 2016 until Aug 2016]
Guillaume Obozinski [ENPC]

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. Application Domains

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. New Software and Platforms

5.1. DICA : Discrete Independent Component Analysis

FUNCTIONAL DESCRIPTION

Moment Matching for Latent Dirichlet Allocation (LDA) and Discrete Independent Component Analysis (DICA).

The DICA package contains Matlab and C++ (via Matlab mex files) implementations of estimation in the LDA and closely related DICA models.

The implementation consists of two parts. One part contains the efficient implementation for construction of the moment/cumulant tensors, while the other part contains implementations of several so called joint diagonalization type algorithms used for matching the tensors. Any tensor type (see below) can be arbitrarily combined with one of the diagonalization algorithms (see below) leading, in total, to 6 algorithms.

Two types of tensors are considered: (a) the LDA moments and (b) the DICA cumulants. The diagonalization algorithms include: (a) the orthogonal joint diagonalization algorithm based on iterative Jacobi rotations, (b) the spectral algorithm based on two eigen decompositions, and (c) the tensor power method.

- Contact: Anastasia Podosinnikova
- URL: https://github.com/anastasia-podosinnikova/dica

5.2. LinearFW: Implementation of linearly convergent versions of Frank-Wolfe

FUNCTIONAL DESCRIPTION

This is the code to reproduce all the experiments in the NIPS 2015 paper: "On the Global Linear Convergence of Frank-Wolfe Optimization Variants" by Simon Lacoste-Julien and Martin Jaggi, which covers the global linear convergence rate of Frank-Wolfe optimization variants for problems described as in Eq. (1) in the paper. It contains the implementation of Frank-Wolfe, away-steps Frank-Wolfe and pairwise Frank-Wolfe on two applications.

- Contact: Simon Lacoste-Julien
- URL: https://github.com/Simon-Lacoste-Julien/linearFW

5.3. cnn_head_detection: Context-aware CNNs for person head detection

FUNCTIONAL DESCRIPTION

Code for ICCV 2015 paper "Context-aware CNNs for person head detection": Person detection is a key problem for many computer vision tasks. While face detection has reached maturity, detecting people under a full variation of camera view-points, human poses, lighting conditions and occlusions is still a difficult challenge. In this work we focus on detecting human heads in natural scenes. Starting from the recent local R-CNN object detector, we extend it with two types of contextual cues. First, we leverage person-scene relations and propose a Global CNN model trained to predict positions and scales of heads directly from the full image. Second, we explicitly model pairwise relations among objects and train a Pairwise CNN model using a structured-output surrogate loss. The Local, Global and Pairwise models are combined into a joint CNN framework. To train and test our full model, we introduce a large dataset composed of 369,846 human heads annotated in 224,740 movie frames. We evaluate our method and demonstrate improvements of person head detection against several recent baselines in three datasets. We also show improvements of the detection speed provided by our model.

- Contact: Anton Osokin
- URL: https://github.com/aosokin/cnn_head_detection

5.4. Lightning: large-scale linear classification, regression and ranking in Python

FUNCTIONAL DESCRIPTION

Lightning is a Python library for large-scale machine learning. More specifically, the library focuses on linear models for classification, regression and ranking. Lightning is the first project to integrate scikit-learn-contrib, a repository of high-quality projects that follow the same API conventions as scikit-learn. Compared to scikit-learn, the main advantages of lightning are its scalability and its flexibility. Indeed, lightning implements cutting-edge optimization algorithms that allow to train models with millions of samples within seconds on commodity hardware. Furthermore, lightning can leverage prior knowledge thanks to so-called structured penalties, an area of research that has recently found applications in domains as diverse as biology, neuroimaging, finance or text processing. Lightning is available under the 3-clause BSD license at http://contrib.scikit-learn.org/lightning/.

- Contact: Fabian Pedregosa
- URL: http://contrib.scikit-learn.org/lightning/

6. New Results

6.1. Regularized Nonlinear Acceleration

In [34], describe a convergence acceleration technique for generic optimization problems. Our scheme computes estimates of the optimum from a nonlinear average of the iterates produced by any optimization method. The weights in this average are computed via a simple linear system, whose solution can be updated online. This acceleration scheme runs in parallel to the base algorithm, providing improved estimates of the solution on the fly, while the original optimization method is running. Numerical experiments are detailed on classical classification problems.

6.2. Harder, Better, Faster, Stronger Convergence Rates for Least-Squares Regression

In [20], we consider the optimization of a quadratic objective function whose gradients are only accessible through a stochastic oracle that returns the gradient at any given point plus a zero-mean finite variance random error. We present the first algorithm that achieves jointly the optimal prediction error rates for least-squares regression, both in terms of forgetting of initial conditions in $O(1/n^2)$, and in terms of dependence on the noise and dimension d of the problem, as $O(d/n)$. Our new algorithm is based on averaged accelerated regularized gradient descent, and may also be analyzed through finer assumptions on initial conditions and the Hessian matrix, leading to dimension-free quantities that may still be small while the " optimal " terms above are large. In order to characterize the tightness of these new bounds, we consider an application to non-parametric regression and use the known lower bounds on the statistical performance (without computational limits), which happen to match our bounds obtained from a single pass on the data and thus show optimality of our algorithm in a wide variety of particular trade-offs between bias and variance.

6.3. Stochastic Variance Reduction Methods for Saddle-Point Problems

In [12], we consider convex-concave saddle-point problems where the objective functions may be split in many components, and extend recent stochastic variance reduction methods (such as SVRG or SAGA) to provide the first large-scale linearly convergent algorithms for this class of problems which are common in machine learning. While the algorithmic extension is straightforward, it comes with challenges and opportunities: (a) the convex minimization analysis does not apply and we use the notion of monotone operators to prove convergence, showing in particular that the same algorithm applies to a larger class of problems, such as variational inequalities, (b) there are two notions of splits, in terms of functions, or in terms of partial derivatives, (c) the split does need to be done with convex-concave terms, (d) non-uniform sampling is key to an efficient algorithm, both in theory and practice, and (e) these incremental algorithms can be easily accelerated using a simple extension of the “catalyst” framework, leading to an algorithm which is always superior to accelerated batch algorithms.

6.4. Frank-Wolfe Algorithms for Saddle Point Problems

In [26], we extend the Frank-Wolfe (FW) optimization algorithm to solve constrained smooth convex-concave saddle point (SP) problems. Remarkably, the method only requires access to linear minimization oracles. Leveraging recent advances in FW optimization, we provide the first proof of convergence of a FW-type saddle point solver over polytopes, thereby partially answering a 30 year-old conjecture. We also survey other convergence results and highlight gaps in the theoretical underpinnings of FW-style algorithms. Motivating applications without known efficient alternatives are explored through structured predic- tion with combinatorial penalties as well as games over matching polytopes involving an exponential number of constraints.

6.5. Minding the Gaps for Block Frank-Wolfe Optimization of Structured SVM

In [10], we propose several improvements on the block-coordinate Frank-Wolfe (BCFW) algorithm from Lacoste-Julien et al. (2013) recently used to optimize the structured support vector machine (SSVM) objective in the context of structured prediction, though it has wider applications. The key intuition behind our improvements is that the estimates of block gaps maintained by BCFW reveal the block suboptimality that can be used as an adaptive criterion. First, we sample objects at each iteration of BCFW in an adaptive non-uniform way via gapbased sampling. Second, we incorporate pairwise and away-step variants of Frank-Wolfe into the block-coordinate setting. Third, we cache oracle calls with a cache-hit criterion based on the block gaps. Fourth, we provide the first method to compute an approximate regularization path for SSVM. Finally, we provide an exhaustive empirical evaluation of all our methods on four structured prediction datasets. The associated SOFTWARE is here: https://github.com/aosokin/gapBCFW

6.6. Asaga: Asynchronous Parallel Saga

In [29], we describe Asaga, an asynchronous parallel version of the incremental gradient algorithm Saga that enjoys fast linear convergence rates. We highlight a subtle but important technical issue present in a large fraction of the recent convergence rate proofs for asynchronous parallel optimization algorithms, and propose a simplification of the recently proposed “perturbed iterate” framework that resolves it. We thereby prove that Asaga can obtain a theoretical linear speedup on multi-core systems even without sparsity assumptions. We present results of an implementation on a 40-core architecture illustrating the practical speedup as well as the hardware overhead.

6.7. Convergence Rate of Frank-Wolfe for Non-Convex Objectives

In [28], we give a simple proof that the Frank-Wolfe algorithm obtains a stationary point at a rate of $O(1/\sqrt{t})$ on non-convex objectives with a Lipschitz continuous gradient. Our analysis is affine invariant and is the first, to the best of our knowledge, giving a similar rate to what was already proven for projected gradient methods (though on slightly different measures of stationarity).

6.8. Highly-Smooth Zero-th Order Online Optimization

The minimization of convex functions which are only available through partial and noisy infor- mation is a key methodological problem in many disciplines. In [3], we consider convex optimization with noisy zero-th order information, that is noisy function evaluations at any desired point. We focus on problems with high degrees of smoothness, such as logistic regression. We show that as opposed to gradient-based algorithms, high-order smoothness may be used to improve estimation rates, with a precise dependence of our upper-bounds on the degree of smoothness. In particular, we show that for infinitely differentiable functions, we recover the same dependence on sample size as gradient-based algorithms, with an extra dimension-dependent factor. This is done for both convex and strongly-convex functions, with finite horizon and anytime algorithms. Finally, we also recover similar results in the online optimization setting.

6.9. Slice Inverse Regression with Score Functions

Non-linear regression and related problems such as non-linear classification are core important tasks in machine learning and statistics. We consider the problem of dimension reduction in non-linear regression, which is often formulated as a non-convex optimization problem.

- We propose score function extensions to sliced inverse regression problems [38], [39], both for the first-order and second-order score functions, which provably improve estimation in the population case over the non-sliced versions; we study finite sample estimators and study their consistency given the exact score functions.
- We propose also to learn the score function as well (using score matching technique [37]) in two steps, i.e., first learning the score function and then learning the effective dimension reduction space, or directly, by solving a convex optimization problem regularized by the nuclear norm.

6.10. Inference and learning for log-supermodular distributions

In [11], we consider log-supermodular models on binary variables, which are probabilistic models with negative log-densities which are submodular. These models provide probabilistic interpretations of common combinatorial optimization tasks such as image segmentation. We make the following contributions:

- We review existing variational bounds for the log-partition function and show that the bound of T. Hazan and T. Jaakkola (On the Partition Function and Random Maximum A-Posteriori Perturbations, Proc. ICML, 2012), based on "perturb-and-MAP" ideas, formally dominates the bounds proposed by J. Djolonga and A. Krause (From MAP to Marginals: Variational Inference in Bayesian Submodular Models, Adv. NIPS, 2014).
- We show that for parameter learning via maximum likelihood the existing bound of J. Djolonga and A. Krause typically leads to a degenerate solution while the one based on "perturb-and-MAP" ideas and logistic samples does not.
- Given that the bound based on "perturb-and-MAP" ideas is an expectation (over our own randomization), we propose to use a stochastic subgradient technique to maximize the lower-bound on the log-likelihood, which can also be extended to conditional maximum likelihood.
- We illustrate our new results on a set of experiments in binary image denoising, where we highlight the flexibility of a probabilistic model for learning with missing data.

6.11. Beyond CCA: Moment Matching for Multi-View Models

In [31], we introduce three novel semi-parametric extensions of probabilistic canonical correlation analysis with identifiability guarantees. We consider moment matching techniques for estimation in these models. For that, by drawing explicit links between the new models and a discrete version of independent component analysis (DICA), we first extend the DICA cumulant tensors to the new discrete version of CCA. By further using a close connection with independent component analysis, we introduce generalized covariance matrices, which can replace the cumulant tensors in the moment matching framework, and, therefore, improve sample complexity and simplify derivations and algorithms significantly. As the tensor power method or orthogonal joint diagonalization are not applicable in the new setting, we use non-orthogonal joint diagonalization techniques for matching the cumulants. We demonstrate performance of the proposed models and estimation techniques on experiments with both synthetic and real datasets.

6.12. PAC-Bayesian Theory Meets Bayesian Inference

In [6], we exhibit a strong link between frequentist PAC-Bayesian bounds and the Bayesian marginal likelihood. That is, for the negative log-likelihood loss function, we show that the minimization of PAC-Bayesian generalization bounds maximizes the Bayesian marginal likelihood. This provides an alternative explanation to the Bayesian Occam's razor criteria, under the assumption that the data is generated by an *i.i.d.* distribution. Moreover, as the negative log-likelihood is an unbounded loss function, we motivate and propose a PAC-Bayesian theorem tailored for the sub-gamma loss family, and we show that our approach is sound on classical Bayesian linear regression tasks.

6.13. A New PAC-Bayesian Perspective on Domain Adaptation

In [7], we study the issue of PAC-Bayesian domain adaptation: We want to learn, from a source domain, a majority vote model dedicated to a target one. Our theoretical contribution brings a new perspective by deriving an upper-bound on the target risk where the distributions' divergence— expressed as a ratio—controls the trade-off between a source error measure and the target voters' disagreement. Our bound suggests that one has to focus on regions where the source data is informative. From this result, we derive a PAC-Bayesian generalization bound, and specialize it to linear classifiers. Then, we infer a learning algorithm and perform experiments on real data.

6.14. PAC-Bayesian Bounds based on the Rényi Divergence

In [13], we propose a simplified proof process for PAC-Bayesian generalization bounds, that allows to divide the proof in four successive inequalities, easing the "customization" of PAC-Bayesian theorems. We also propose a family of PAC-Bayesian bounds based on the Rényi divergence between the prior and posterior distributions, whereas most PAC-Bayesian bounds are based on the Kullback-Leibler divergence. Finally, we present an empirical evaluation of the tightness of each inequality of the simplified proof, for both the classical PAC-Bayesian bounds and those based on the Rényi divergence.

6.15. PAC-Bayesian theorems for multiview learning

In [27], we tackle the issue of multiview learning which aims to take advantages of multiple representations/views of the data. In this context, many machine learning algorithms exist. However, the majority of the theoretical studies focus on learning with exactly two representations. In this paper, we propose a general PAC-Bayesian theory for multiview learning with more than two views. We focus our study to binary classification models that take the form of a majority vote. We derive PAC-Bayesian generalization bounds allowing to consider different relations between empirical and true risks by taking into account a notion of diversity of the voters and views, and that can be naturally extended to semi-supervised learning.

6.16. A spectral algorithm for fast de novo layout of uncorrected long nanopore reads

Seriation is an optimization problem that seeks to reconstruct an ordering between n variables from pairwise similarity information. It can be formulated as a combinatorial problem over permutations and several algorithms have been derived from relaxations of this problem. We make the link between the seriation framework and the task of de novo genome assembly, which consists of reconstructing a whole DNA sequence from small pieces of it that are oversampled so as to cover the full genome. To achieve this task, one has to find the layout of small pieces of DNA sequences (reads). This layout step can be cast as a seriation problem. We show that a spectral algorithm for seriation can be efficiently applied to a genome assembly scheme.

New long read sequencers promise to transform sequencing and genome assembly by producing reads tens of kilobases long. However their high error rate significantly complicates assembly and requires expensive correction steps to layout the reads using standard assembly engines.

We present an original and efficient spectral algorithm to layout the uncorrected nanopore reads, and its seamless integration into a straightforward overlap/layout/consensus (OLC) assembly scheme. The method is shown to assemble Oxford Nanopore reads from several bacterial genomes into good quality ($\sim 99\%$ identity to the reference) genome-sized contigs, while yielding more fragmented assemblies from a *Sacharomyces cerevisiae* reference strain. See software in https://github.com/antrec/spectrassembler.

6.17. Using Deep Learning and Generative Adversarial Networks to Study Large Scale GFP Screens

Fluorescent imaging of GFP tagged proteins is one of the most widely used techniques to view the dynamics of proteins in live cells. By combining it with different perturbations such as RNAi or drug treatments we can understand how cells regulate complex processes such as mitosis or the cell cycle.

However, GFP imaging has certain limitations. There are only a limited number of different fluorescent proteins available, making imaging multiple proteins at the same time very challenging and expensive. Finally, analyzing complex screens can be very challenging: it's not always obvious a-priori what kind of features will predict the phenotypes we are interested in.

We discuss a new approach to studying large scale GFP screens using deep convolutional networks. We show that by using convolutional neural networks, we can greatly outperform traditional feature based approaches at different kind of prediction tasks. The networks learn flexible representations, which are suitable for multiple tasks, such as predicting the localization of Tea1 in fission yeast cells (blue signal, shown in image) in cells where only other proteins are tagged.

We then show that we can use generative adversarial neural networks to learn highly compact latent representations. Those latent representations can then be used to generate new realistic images, allowing us to simulate new phenotypes, and to predict the outcome of new perturbations (joint work between Federico Vaggi, Anton Osokin, Theophile Dalens).

6.18. SymPy: Symbolic computing in Python

SymPy is an open source computer algebra system written in pure Python. It is built with a focus on extensibility and ease of use, through both interactive and programmatic applications. These characteristics have led SymPy to become the standard symbolic library for the scientific Python ecosystem. This paper [30] presents the architecture of SymPy, a description of its features, and a discussion of select domain specific submodules. The supplementary materials provide additional examples and further outline details of the architecture and features of SymPy. As for the software, I am one of the main authors of the lightning machine learning library, that you can include if you want.

6.19. Robust Discriminative Clustering with Sparse Regularizers

Clustering high-dimensional data often requires some form of dimensionality reduction, where clustered variables are separated from "noise-looking" variables. In [24], we cast this problem as finding a low-dimensional projection of the data which is well-clustered. This yields a one-dimensional projection in the simplest situation with two clusters, and extends naturally to a multi-label scenario for more than two clusters. In this paper, (a) we first show that this joint clustering and dimension reduction formulation is equivalent to previously proposed discriminative clustering frameworks, thus leading to convex relaxations of the problem, (b) we propose a novel sparse extension, which is still cast as a convex relaxation and allows estimation in higher dimensions, (c) we propose a natural extension for the multi-label scenario, (d) we provide a new theoretical analysis of the performance of these formulations with a simple probabilistic model, leading to scalings over the form $d = O(\sqrt{n})$ for the affine invariant case and $d = O(n)$ for the sparse case, where n is the number of examples and d the ambient dimension, and finally, (e) we propose an efficient iterative algorithm with running-time complexity proportional to $O(nd^2)$, improving on earlier algorithms which had quadratic complexity in the number of examples.

6.20. Optimal Rates of Statistical Seriation

Given a matrix the seriation problem consists in permuting its rows in such way that all its columns have the same shape, for example, they are monotone increasing. In [23], we propose a statistical approach to this problem where the matrix of interest is observed with noise and study the corresponding minimax rate of estimation of the matrices. Specifically, when the columns are either unimodal or monotone, we show that the least squares estimator is optimal up to logarithmic factors and adapts to matrices with a certain natural structure. Finally, we propose a computationally efficient estimator in the monotonic case and study its performance both theoretically and experimentally. Our work is at the intersection of shape constrained estimation and recent work that involves permutation learning, such as graph denoising and ranking.

6.21. Breaking Sticks and Ambiguities with Adaptive Skip-gram

Recently proposed Skip-gram model is a powerful method for learning high-dimensional word representations that capture rich semantic relationships between words. However, Skip-gram as well as most prior work on learning word representations does not take into account word ambiguity and maintain only single representation per word. Although a number of Skip-gram modifications were proposed to overcome this limitation and learn multi-prototype word representations, they either require a known number of word meanings or learn them using greedy heuristic approaches. In [4], we propose the Adaptive Skip-gram model which is a nonparametric Bayesian extension of Skip-gram capable to automatically learn the required number of representations for all words at desired semantic resolution. We derive efficient online variational learning algorithm for the model and empirically demonstrate its efficiency on word-sense induction task.

6.22. Deep Part-Based Generative Shape Model with Latent Variables

The Shape Boltzmann Machine (SBM) and its multilabel version MSBM [5] have been recently introduced as deep generative models that capture the variations of an object shape. While being more flexible MSBM requires datasets with labeled parts of the objects for training. In [8], we present an algorithm for training MSBM using binary masks of objects and the seeds which approximately correspond to the locations of objects parts. The latter can be obtained from part-based detectors in an unsupervised manner. We derive a latent variable model and an EM-like training procedure for adjusting the weights of MSBM using a deep learning framework. We show that the model trained by our method outperforms SBM in the tasks related to binary shapes and is very close to the original MSBM in terms of quality of multilabel shapes.

6.23. Unsupervised Learning from Narrated Instruction Videos

In [2], we address the problem of automatically learning the main steps to complete a certain task, such as changing a car tire, from a set of narrated instruction videos. The contributions of this paper are three-fold. First, we develop a new unsupervised learning approach that takes advantage of the complementary nature of the input video and the associated narration. The method solves two clustering problems, one in text and one in video, applied one after each other and linked by joint constraints to obtain a single coherent sequence of steps in both modalities. Second, we collect and annotate a new challenging dataset of real-world instruction videos from the Internet. The dataset contains about 800,000 frames for five different tasks that include complex interactions between people and objects, and are captured in a variety of indoor and outdoor settings. Third, we experimentally demonstrate that the proposed method can automatically discover, in an unsupervised manner, the main steps to achieve the task and locate the steps in the input videos. The associated SOFTWARE is here: https://github.com/jalayrac/instructionVideos

6.24. Stochastic Optimization for Large-scale Optimal Transport

Optimal transport (OT) defines a powerful framework to compare probability distributions in a geometrically faithful way. However, the practical impact of OT is still limited because of its computational burden. In [5], we propose a new class of stochastic optimization algorithms to cope with large-scale OT problems. These methods can handle arbitrary distributions (either discrete or continuous) as long as one is able to draw samples from them, which is the typical setup in high-dimensional learning problems. This alleviates the need to discretize these densities, while giving access to provably convergent methods that output the correct distance without discretization error. These algorithms rely on two main ideas: *(a)* the dual OT problem can be re-cast as the maximization of an expectation; *(b)* the entropic regularization of the primal OT problem yields a smooth dual optimization which can be addressed with algorithms that have a provably faster convergence. We instantiate these ideas in three different setups: *(i)* when comparing a discrete distribution to another, we show that incremental stochastic optimization schemes can beat Sinkhorn's algorithm, the current state-of-the-art finite dimensional OT solver; *(ii)* when comparing a discrete distribution to a continuous density, a semi-discrete reformulation of the dual program is amenable to averaged stochastic gradient descent, leading to better performance than approximately solving the problem by discretization ; *(iii)* when dealing with two continuous densities, we propose a stochastic gradient descent over a reproducing kernel Hilbert space (RKHS). This is currently the only known method to solve this problem, apart from computing OT on finite samples. We backup these claims on a set of discrete, semi-discrete and continuous benchmark problems.

6.25. Online but Accurate Inference for Latent Variable Models with Local Gibbs Sampling

We study parameter inference in large-scale latent variable models. We first propose a unified treatment of online inference for latent variable models from a non-canonical exponential family, and draw explicit links between several previously proposed frequentist or Bayesian methods. We then propose a novel inference method for the frequentist estimation of parameters, that adapts MCMC methods to online inference of latent variable models with the proper use of local Gibbs sampling. Then, for latent Dirichlet allocation,we provide an extensive set of experiments and comparisons with existing work, where our new approach outperforms all previously proposed methods. In particular, using Gibbs sampling for latent variable inference is superior to variational inference in terms of test log-likelihoods. Moreover, Bayesian inference through variational methods perform poorly, sometimes leading to worse fits with latent variables of higher dimensionality.

In [22], we focus on methods that make a single pass over the data to estimate parameters. We make the following contributions:

1. We review and compare existing methods for online inference for latent variable models from a non-canonical exponential family, and draw explicit links between several previously proposed frequentist or Bayesian methods. Given the large number of existing methods, our unifying framework allows to understand differences and similarities between all of them.
2. We propose a novel inference method for the frequentist estimation of parameters, that adapts MCMC methods to online inference of latent variable models with the proper use of "local" Gibbs sampling. In our online scheme, we apply Gibbs sampling to the current observation, which is "local", as opposed to "global" batch schemes where Gibbs sampling is applied to the entire dataset.
3. After formulating LDA as a non-canonical exponential family, we provide an extensive set of experiments, where our new approach outperforms all previously proposed methods. In particular, using Gibbs sampling for latent variable inference is superior to variational inference in terms of test log-likelihoods. Moreover, Bayesian inference through variational methods perform poorly, sometimes leading to worse fits with latent variables of higher dimensionality.

6.26. Learning Determinantal Point Processes in Sublinear Time

In [21], we propose a new class of determinantal point processes (DPPs) which can be manipulated for inference and parameter learning in potentially sublinear time in the number of items. This class, based on a specific low-rank factorization of the marginal kernel, is particularly suited to a subclass of continuous DPPs and DPPs defined on exponentially many items. We apply this new class to modelling text documents as sampling a DPP of sentences, and propose a conditional maximum likelihood formulation to model topic proportions, which is made possible with no approximation for our class of DPPs. We present an application to document summarization with a DPP on 2^{500} items.

We make the following contributions:

- We propose a new class of determinantal point processes (DPPs) which is based on a particular low-rank factorization of the marginal kernel. Through the availability of a particular second-moment matrix, the complexity for inference and learning tasks is polynomial in the rank of the factorization and thus often sublinear in the total number of items (with exact likelihood computations).
- As shown in this work, these new DPPs are particularly suited to a subclass of continuous DPPs (infinite number of items), such as on $[0,1]^m$, and DPPs defined on the V-dimensional hypercube, which has 2^V elements.
- We propose a model of documents as sampling a DPP of sentences, and propose a conditional maximum likelihood formulation to model topic proportions. We present an application to document summarization with a DPP on 2^{500} items.

6.27. Decentralized Topic Modelling with Latent Dirichlet Allocation

Privacy preserving networks can be modelled as decentralized networks (e.g., sensors, connected objects, smartphones), where communication between nodes of the network is not controlled by a master or central node. For this type of networks, the main issue is to gather/learn global information on the network (e.g., by optimizing a global cost function) while keeping the (sensitive) information at each node. In this work, we focus on text information that agents do not want to share (e.g., text messages, emails, confidential reports). We use recent advances on decentralized optimization and topic models to infer topics from a graph with limited communication. We propose a method to adapt latent Dirichlet allocation (LDA) model to decentralized optimization and show on synthetic data that we still recover similar parameters and similar performance at each node than with stochastic methods accessing to the whole information in the graph.

In [14], we tackle the non-convex problem of topic modelling, where agents have sensitive text data at their disposal that they can not or do not want to share (e.g., text messages, emails, confidential reports). More precisely, we adapt the particular Latent Dirichlet Allocation (LDA) model to decentralized networks. We combine recent work of [22] on online inference for latent variable models, which adapts online EM with local Gibbs sampling in the case of intractable latent variable models (such as LDA) and recent advances on decentralized optimization.

7. Bilateral Contracts and Grants with Industry

7.1. 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/.

7.2. Bilateral Grants with Industry

- A. d’Aspremont: AXA, "mécénat scientifique, chaire Havas-Dauphine", machine learning.
- A. d’Aspremont: Société Générale - fondation ENS, "mécénat scientifique".
- A. d’Aspremont: Projet EMMA at Institut Louis Bachelier. Collaboration with Euroclear on REPO markets.
- S. Lacoste-Julien (with J. Sivic and I. Laptev in Willow project-team): Google Research Award “Structured Learning from Video and Natural Language”.
- F. Bach: Gift from Facebook AI Research.

8. Partnerships and Cooperations

8.1. European Initiatives

8.1.1. FP7 & H2020 Projects

8.1.1.1. SIPA

Title: Semidefinite Programming with Applications in Statistical Learning

Type: FP7

Instrument: ERC Starting Grant Duration: May 2011 - May 2016 Coordinator: A. d'Aspremont (CNRS)

Abstract: Interior point algorithms and a dramatic growth in computing power have revolutionized optimization in the last two decades. Highly nonlinear problems which were previously thought intractable are now routinely solved at reasonable scales. Semidefinite programs (i.e. linear programs on the cone of positive semidefinite matrices) are a perfect example of this trend: reasonably large, highly nonlinear but convex eigenvalue optimization problems are now solved efficiently by reliable numerical packages. This in turn means that a wide array of new applications for semidefinite programming have been discovered, mimicking the early development of linear programming. To cite only a few examples, semidefinite programs have been used to solve collaborative filtering problems (e.g. make personalized movie recommendations), approximate the solution of combinatorial programs, optimize the mixing rate of Markov chains over networks, infer dependence patterns from multivariate time series or produce optimal kernels in classification problems. These new applications also come with radically different algorithmic requirements. While interior point methods solve relatively small problems with a high precision, most recent applications of semidefinite programming in statistical learning for example form very large-scale problems with comparatively low precision targets, programs for which current algorithms cannot form even a single iteration. This proposal seeks to break this limit on problem size by deriving reliable first-order algorithms for solving large-scale semidefinite programs with a significantly lower cost per iteration, using for example subsampling techniques to considerably reduce the cost of forming gradients. Beyond these algorithmic challenges, the proposed research will focus heavily on applications of convex programming to statistical learning and signal processing theory where optimization and duality results quantify the statistical performance of coding or variable selection algorithms for example. Finally, another central goal of this work will be to produce efficient, customized algorithms for some key problems arising in machine learning and statistics.

8.1.1.2. MacSeNet

Title: Machine Sensing Training Network

Type: H2020

Instrument: Initial Training Network

Duration: January 2015 - January 2019

Coordinator: Mark Plumbley (University of Surrey)

Inria contact: Francis Bach

Abstract: The aim of this Innovative Training Network is to train a new generation of creative, entrepreneurial and innovative early stage researchers (ESRs) in the research area of measurement and estimation of signals using knowledge or data about the underlying structure. We will develop new robust and efficient Machine Sensing theory and algorithms, together methods for a wide range of signals, including: advanced brain imaging; inverse imaging problems; audio and music signals; and non-traditional signals such as signals on graphs. We will apply these methods to real-world problems, through work with non-Academic partners, and disseminate the results of this research to a wide range of academic and non-academic audiences, including through publications, data, software and public engagement events. MacSeNet is funded under the H2020-MSCA-ITN-2014 call and is part of the Marie Sklodowska- Curie Actions — Innovative Training Networks (ITN) funding scheme.

8.1.1.3. Spartan

Title: Sparse Representations and Compressed Sensing Training Network Type: FP7

Instrument: Initial Training Network

Duration: October 2014 to October 2018

Coordinator: Mark Plumbley (University of Surrey)

Inria contact: Francis Bach

Abstract: The SpaRTaN Initial Training Network will train a new generation of interdisciplinary researchers in sparse representations and compressed sensing, contributing to Europe's leading role in scientific innovation. By bringing together leading academic and industry groups with expertise in sparse representations, compressed sensing, machine learning and optimisation, and with an interest in applications such as hyperspectral imaging, audio signal processing and video analytics, this project will create an interdisciplinary, trans-national and inter-sectorial training network to enhance mobility and training of researchers in this area. SpaRTaN is funded under the FP7-PEOPLE-2013-ITN call and is part of the Marie Curie Actions — Initial Training Networks (ITN) funding scheme: Project number - 607290

8.1.1.4. SEQUOIA

Title: Robust algorithms for learning from modern data

Programm: H2020

Type: ERC

Duration: 2017-202

Coordinator: Inria

Inria contact: Francis BACH

8.2. International Initiatives

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

8.2.1.1. BigFOKS2

Title: Learning from Big Data: First-Order methods for Kernels and Submodular functions

International Partner (Institution - Laboratory - Researcher):

IISc Bangalore (India) - Computer Science Department - Chiranjib Bhattacharyya

Start year: 2016

See also: http://mllab.csa.iisc.ernet.in/indo-french.html

Recent advances in sensor technologies have resulted in large amounts of data being generated in a wide array of scientific disciplines. Deriving models from such large datasets, often known as "Big Data", is one of the important challenges facing many engineering and scientific disciplines. In this proposal we investigate the problem of learning supervised models from Big Data, which has immediate applications in Computational Biology, Computer vision, Natural language processing, Web, E-commerce, etc., where specific structure is often present and hard to take into account with current algorithms. Our focus will be on the algorithmic aspects. Often supervised learning problems can be cast as convex programs. The goal of this proposal will be to derive first-order methods which can be effective for solving such convex programs arising in the Big-Data setting. Keeping this broad goal in mind we investigate two foundational problems which are not well addressed in existing literature. The first problem investigates Stochastic Gradient Descent Algorithms in the context of First-order methods for designing algorithms for Kernel based prediction functions on Large Datasets. The second problem involves solving discrete optimization problems arising in Submodular formulations in Machine Learning, for which first-order methods have not reached the level of speed required for practical applications (notably in computer vision).

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

Alexandre d'Aspremont: Workshop preparation for les Houches in Feb. 2016: “Optimization without borders”, to celebrate Y. Nesterov's 60th birthday.

Francis Bach: organization of a workshop at IHES (with S. Arlot and A. Celisse), March 2016.

Francis Bach: co-organization of two NIPS workshops.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

Francis Bach: Area chair for ICML 2016

Simon Lacoste-Julien: Area chair for ICML 2016

Simon Lacoste-Julien: Area chair for NIPS 2016

9.1.3. Journal

9.1.3.1. Member of Editorial Boards

Alexandre d'Aspremont: Associate Editor, SIAM Journal on Optimization (2013-...).

F. Bach: Action Editor, Journal of Machine Learning Research.

F. Bach: Information and Inference, Associate Editor.

F. Bach: SIAM Journal on Imaging Sciences, Associate Editor.

F. Bach: Electronic Journal of Statistics, Associate Editor.

9.1.4. Invited Talks

Alexandre d'Aspremont: Regularized Nonlinear Acceleration, BIRS workshop, Oaxaca, October 2016.

Alexandre d'Aspremont: Optimal Affine Invariant Smooth Minimization Algorithms, Institut des hautes études scientifiques, June 2016.

Alexandre d'Aspremont: Optimal Affine Invariant Smooth Minimization Algorithms, Nexus of Information and Computation Theories, Institut Henri Poincaré, March 2016.

Alexandre d'Aspremont: Optimal Affine Invariant Smooth Minimization Algorithms, Workshop on Algorithms and Dynamics for Games and Optimization, Santiago Chile, January 2016.

Francis Bach: Winter School on Signal processing, Bonn, January 2016.

Francis Bach: "Optimization without borders", Les Houches, February 2016.

Francis Bach: Oberwolfach, March 2016.

Francis Bach: Dali meeting, Sestri Levante, Italy, March 2016.

Francis Bach: ETH Computer Science Colloquium, April 2016.

Francis Bach: Workshop San Servolo, May 2016.

Francis Bach: Machine Learning summer school Cadiz, May 2016.

Francis Bach: Summer school, Bangalore, July 2016.

Francis Bach: ICCOPT conference, plenary speaker, August 2016.

Francis Bach: Workshop, Haifa, Septembre 2016.

Francis Bach: Statistics Seminar, Cambridge, October 2016.

Francis Bach: BIRS Oaxaca, October 2016.

Francis Bach: NIPS workshops (three presentations), December 2016.

Damien Garreau: “Consistent multiple change-point detection with kernels”, Group meeting of Geometrica Inria project team, Saclay (February 18, 2016).

Damien Garreau: “Consistent multiple change-point detection with kernels”, Inria Junior Seminar, Paris (March 15, 2016).

Damien Garreau: “Consistent multiple change-point detection with kernels”, Colloque final de l’ANR Calibration, Nice (April 7, 2016).

Damien Garreau: “Consistent multiple change-point detection with kernels”, Colloque Jeunes probabilistes et Statisticiens, Les Houches (April 18, 2016).

Pascal Germain: “A Representation Learning Approach for Domain Adaptation”, Tao Seminars, Université Paris-Sud, Paris, France, March 2016.

Pascal Germain: “A Representation Learning Approach for Domain Adaptation”, Data Intelligence Group Seminars, Université Jean-Monnet, Saint-Étienne, France, March 2016.

Pascal Germain: “Variations on the PAC-Bayesian Bound”, Bayes in Paris Seminar at ENSAE, Paris, France, June 2016.

Pascal Germain: “Variations on the PAC-Bayesian Bound”, Séminaires du département d’informatique et de génie logiciel, Université Laval, Quebec, Canada, July 2016.

Simon Lacoste-Julien: "On the Global Linear Convergence of Frank-Wolfe Optimization Variants", invited talk in the Conic and Polynomial Optimization cluster at ICCOPT 2016, Tokyo, Japan, August 2016..

Simon Lacoste-Julien: "On the Global Linear Convergence of Frank-Wolfe Optimization Variants", invited talk in the Learning and Optimization workshop of DALI meeting, Sestri Levante, Italy, April 2016.

Simon Lacoste-Julien: "Modern Optimization for Structured Machine Learning", CS & OR Department Colloquium, Université de Montréal, Montreal, Canada, February 2016.

Antoine Recanati: Presentation at the group meeting of Mines ParisTech Centre for Computational Biology (CBIO) at Institut Curie, October, 18th 2016.

9.1.5. Leadership within the Scientific Community

Alexandre d’Aspremont: Porteur de l’IRIS PSL “Science des données, données de la science”.

Alexandre d’Aspremont: Co-scientific director of Master’s program MASH (Mathématiques, Apprentissage et Sciences Humaines), with ENS - Paris Dauphine.

Alexandre d’Aspremont: Scientific committee, programme Gaspard Monge pour l’Optimisation.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: A. d’Aspremont, M1 course on Optimization: ENS Paris, 21h

Master: A. d’Aspremont, M2 course on Optimization: MVA, ENS Cachan, 21h

Master : F. Bach (together with J.-P. Vert), “Apprentissage statistique”, 35h, M1, Ecole Normale Supérieure.

Master : F. Bach (together with G. Obozinski), "Graphical models", 30h, M2 (MVA), ENS Cachan.

Master : F. Bach , 20h, M2 (Mathématiques de l’aléatoire), Université Paris-Sud.

Mastere (M1): S. Lacoste-Julien, F. Vogel, “Projets informatiques”, 10h, Université de Paris-Dauphine, Master M2: Mathématiques, Apprentissage et Sciences Humaines (MASH)

Master : A. Osokin (together with K. Alahari), “The introduction to discrete optimization”, 30h, M2, Centrale Supélec

Master: Fabian Pedredoga, Machine learning with scikit-learn, Master Mathématiques, Apprentissage et Sciences Humaines (MASH), Paris Dauphine.

9.2.2. Supervision

PhD: Anastasia Podosinnikova, November 2016, co-advised by Francis Bach and Simon Lacoste-Julien

PhD: Thomas Schatz, September 2016, co-advised by and E. Dupoux (ENS, cognitive sciences).

PhD: Sesh Kumar, September 2016, advised by F. Bach.

PhD in progress : Nom du doctorant, titre (provisoire) du mémoire, date du début de la thèse, encadrant(s)

PhD in progress : Jean-Baptiste Alayrac, co-advised by Simon Lacoste-Julien, Josef Sivic and Ivan Laptev, started Sept. 2014.

PhD in progress : Rémi Leblond, advised by Simon Lacoste-Julien, started Sept. 2015.

PhD in progress : Gauthier Gidel, advised by Simon Lacoste-Julien, started Sept. 2016.

PhD in progress : Vincent Roulet, directed by Alexandre d'Aspremont, started as a PhD on Oct. 1 2014.

PhD in progress : Nicolas Flammarion, co-directed by Alexandre d'Aspremont and Francis Bach, started Sept. 2013.

PhD in progress : Damien Scieur, co-directed with Alexandre d'Aspremont and Francis Bach, started Sept. 2015.

PhD in progress : Antoine Recanati, directed by Alexandre d'Aspremont, started Sept. 2015.

PhD in progress: Rafael Rezende, September 2013, F. Bach, co-advised with J. Ponce.

PhD in progress: PhD in progress: Christophe Dupuy, January 2014, co-advised by F. Bach and C. Diot (Technicolor).

PhD in progress: Damien Garreau, September 2014, co-advised by S. Arlot and G. Biau.

PhD in progress: Anaël Bonneton, December 2014, co- advised by F. Bach, located in Agence nationale de la sécurité des systèmes d'information (ANSSI).

PhD in progress: Dmitry Babichev, September 2015, co-advised by F. Bach and A. Judistky (Univ. Grenoble).

PhD in progress: Tatiana Shpakova, September 2015, advised by F. Bach.

9.2.3. Juries

Alexandre d'Aspremont: PhD Committee for Igor Colin, Nov. 2016.

Francis Bach: PhD Committee for Alain Durmus, Dec. 2016.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] F. FOGEL, I. WALDSPURGER, A. D'ASPREMONT.*Phase retrieval for imaging problems*, in "Mathematical Programming Computations", September 2016, vol. 8, n^o 3, p. 311-335, https://hal.archives-ouvertes.fr/hal-00907529.

International Conferences with Proceedings

[2] J.-B. ALAYRAC, P. BOJANOWSKI, N. AGRAWAL, J. SIVIC, I. LAPTEV, S. LACOSTE-JULIEN.*Unsupervised Learning from Narrated Instruction Videos*, in "CVPR2016 - 29th IEEE Conference on Computer Vision and Pattern Recognition", Las Vegas, United States, June 2016, https://hal.inria.fr/hal-01171193.

[3] F. BACH, V. PERCHET.*Highly-Smooth Zero-th Order Online Optimization Vianney Perchet*, in "Conference on Learning Theory (COLT)", New York, United States, June 2016, https://hal.archives-ouvertes.fr/hal-01321532.

[4] S. BARTUNOV, D. KONDRASHKIN, A. OSOKIN, D. VETROV.*Breaking Sticks and Ambiguities with Adaptive Skip-gram*, in "Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS)", Cadiz, Spain, May 2016, p. 130–138, https://hal.archives-ouvertes.fr/hal-01404056.

[5] A. GENEVAY, M. CUTURI, G. PEYRÉ, F. BACH.*Stochastic Optimization for Large-scale Optimal Transport*, in "NIPS 2016 - Thirtieth Annual Conference on Neural Information Processing System", Barcelona, Spain, NIPS (editor), Proc. NIPS 2016, December 2016, https://hal.archives-ouvertes.fr/hal-01321664.

[6] P. GERMAIN, F. BACH, A. LACOSTE, S. LACOSTE-JULIEN.*PAC-Bayesian Theory Meets Bayesian Inference*, in "Neural Information Processing Systems (NIPS 2016)", Barcelone, Spain, Proceedings of the Neural Information Processing Systems Conference, December 2016, https://hal.archives-ouvertes.fr/hal-01324072.

[7] P. GERMAIN, A. HABRARD, F. LAVIOLETTE, E. MORVANT.*A New PAC-Bayesian Perspective on Domain Adaptation*, in "33rd International Conference on Machine Learning (ICML 2016)", New York, NY, United States, Proceedings of the 33rd International Conference on Machine Learning, June 2016, https://hal.archives-ouvertes.fr/hal-01307045.

[8] A. KIRILLOV, M. GAVRIKOV, E. LOBACHEVA, A. OSOKIN, D. VETROV.*Deep Part-Based Generative Shape Model with Latent Variables*, in "27th British Machine Vision Conference (BMVC 2016)", York, United Kingdom, September 2016, https://hal.archives-ouvertes.fr/hal-01404071.

[9] R. LAJUGIE, P. BOJANOWSKI, P. CUVILLIER, S. ARLOT, F. BACH.*A weakly-supervised discriminative model for audio-to-score alignment*, in "41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP)", Shanghai, China, Proceedings of the 41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP), March 2016, https://hal.archives-ouvertes.fr/hal-01251018.

[10] A. OSOKIN, J.-B. ALAYRAC, I. LUKASEWITZ, P. K. DOKANIA, S. LACOSTE-JULIEN.*Minding the Gaps for Block Frank-Wolfe Optimization of Structured SVMs*, in "International Conference on Machine Learning (ICML 2016)", New York, United States, 2016, Appears in Proceedings of the 33rd International Conference on Machine Learning (ICML 2016). 31 pages, https://hal.archives-ouvertes.fr/hal-01323727.

[11] T. SHPAKOVA, F. BACH.*Parameter Learning for Log-supermodular Distributions*, in "NIPS 2016 - Thirtieth Annual Conference on Neural Information Processing System", Barcelona, Spain, December 2016, https://hal.inria.fr/hal-01354789.

Conferences without Proceedings

[12] P. BALAMURUGAN, F. BACH.*Stochastic Variance Reduction Methods for Saddle-Point Problems*, in "Neural Information Processing Systems (NIPS)", Barcelona, Spain, Advances in Neural Information Processing Systems, December 2016, https://hal.archives-ouvertes.fr/hal-01319293.

[13] L. BÉGIN, P. GERMAIN, F. LAVIOLETTE, J.-F. ROY.*PAC-Bayesian Bounds based on the Rényi Divergence*, in "International Conference on Artificial Intelligence and Statistics (AISTATS 2016)", Cadiz, Spain, Proceedings of the 19th International Conference on Artificial Intelligence and Statistics, May 2016, https://hal.inria.fr/hal-01384783.

[14] I. COLIN, C. DUPUY.*Decentralized Topic Modelling with Latent Dirichlet Allocation*, in "NIPS 2016 - 30th Conference on Neural Information Processing Systems", Barcelone, Spain, December 2016, https://hal.archives-ouvertes.fr/hal-01383111.

[15] A. GOYAL, E. MORVANT, P. GERMAIN, M.-R. AMINI.*Théorèmes PAC-Bayésiens pour l'apprentissage multi-vues*, in "Conférence Francophone sur l'Apprentissage Automatique (CAp)", Marseille, France, July 2016, https://hal.archives-ouvertes.fr/hal-01329763.

[16] L. LANDRIEU, G. OBOZINSKI.*Cut Pursuit: fast algorithms to learn piecewise constant functions*, in "19th International Conference on Artificial Intelligence and Statistics (AISTATS 2016)", Cadix, Spain, May 2016, https://hal.archives-ouvertes.fr/hal-01306786.

Research Reports

[17] P. GERMAIN, A. HABRARD, F. LAVIOLETTE, E. MORVANT.*PAC-Bayesian Theorems for Domain Adaptation with Specialization to Linear Classifiers*, Université Jean Monnet, Saint-Étienne (42) ; Département d'Informatique et de Génie Logiciel, Université Laval (Québec) ; ENS Paris ; IST Austria, August 2016, This report is a long version of our paper entitled A PAC-Bayesian Approach for Domain Adaptation with Specialization to Linear Classifiers published in the proceedings of the International Conference on Machine Learning (ICML) 2013. We improved our main results, extended our experiments, and proposed an extension to multisource domain adaptation, https://hal.archives-ouvertes.fr/hal-01134246.

Other Publications

[18] D. BABICHEV, F. BACH.*Slice inverse regression with score functions*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01388498.

[19] F. BACH.*Submodular Functions: from Discrete to Continous Domains*, February 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01222319.

[20] A. DIEULEVEUT, N. FLAMMARION, F. BACH.*Harder, Better, Faster, Stronger Convergence Rates for Least-Squares Regression*, February 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01275431.

[21] C. DUPUY, F. BACH.*Learning Determinantal Point Processes in Sublinear Time*, October 2016, Under review for AISTATS 2017, https://hal.archives-ouvertes.fr/hal-01383742.

[22] C. DUPUY, F. BACH.*Online but Accurate Inference for Latent Variable Models with Local Gibbs Sampling*, July 2016, Under submission in JMLR, https://hal.inria.fr/hal-01284900.

[23] N. FLAMMARION, C. MAO, P. RIGOLLET.*Optimal Rates of Statistical Seriation*, November 2016, V2 corrects an error in Lemma A.1, v3 corrects appendix F on unimodal regression where the bounds now hold with polynomial probability rather than exponential, https://hal.archives-ouvertes.fr/hal-01405738.

[24] N. FLAMMARION, B. PALANIAPPAN, F. BACH.*Robust Discriminative Clustering with Sparse Regularizers*, August 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01357666.

[25] D. GARREAU, S. ARLOT.*Consistent change-point detection with kernels*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01416704.

[26] G. GIDEL, T. JEBARA, S. LACOSTE-JULIEN.*Frank-Wolfe Algorithms for Saddle Point Problems*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01403348.

[27] A. GOYAL, E. MORVANT, P. GERMAIN, M.-R. AMINI.*PAC-Bayesian Theorems for Multiview Learning*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01336260.

[28] S. LACOSTE-JULIEN.*Convergence Rate of Frank-Wolfe for Non-Convex Objectives*, June 2016, 6 pages, https://hal.inria.fr/hal-01415335.

[29] R. LEBLOND, F. PEDREGOSA, S. LACOSTE-JULIEN.*Asaga: Asynchronous Parallel Saga*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01407833.

[30] A. MEURER, C. P. SMITH, M. PAPROCKI, O. ČERTÍK, S. B. KIRPICHEV, M. ROCKLIN, A. KUMAR, S. IVANOV, J. K. MOORE, S. SINGH, T. RATHNAYAKE, S. VIG, B. E. GRANGER, R. P. MULLER, F. BONAZZI, H. GUPTA, S. VATS, F. JOHANSSON, F. PEDREGOSA, M. J. CURRY, A. R. TERREL, Š. ROUČKA, A. SABOO, I. FERNANDO, S. KULAL, R. CIMRMAN, A. SCOPATZ.*SymPy: Symbolic computing in Python*, May 2016, working paper or preprint [*DOI :* 10.7287/PEERJ.PREPRINTS.2083V3], https://hal.inria.fr/hal-01404156.

[31] A. PODOSINNIKOVA, F. BACH, S. LACOSTE-JULIEN.*Beyond CCA: Moment Matching for Multi-View Models*, March 2016, working paper or preprint, https://hal.inria.fr/hal-01291060.

[32] V. ROULET, F. FOGEL, A. D'ASPREMONT, F. BACH.*Learning with Clustering Structure*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01239305.

[33] M. SCHMIDT, N. LE ROUX, F. BACH.*Minimizing Finite Sums with the Stochastic Average Gradient*, May 2016, Revision from January 2015 submission. Major changes: updated literature follow and discussion of subsequent work, additional Lemma showing the validity of one of the formulas, somewhat simplified presentation of Lyapunov bound, included code needed for checking proofs rather than the polynomials generated by the code, added error regions to the numerical experiments, https://hal.inria.fr/hal-00860051.

[34] D. SCIEUR, A. D'ASPREMONT, F. BACH.*Regularized Nonlinear Acceleration*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01384682.

[35] G. SEGUIN, P. BOJANOWSKI, R. LAJUGIE, I. LAPTEV.*Instance-level video segmentation from object tracks*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01255765.

[36] K. S. SESH KUMAR, F. BACH.*Active-set Methods for Submodular Minimization Problems*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01161759.

References in notes

[37] A. HYVÄRINEN.*Estimation of non-normalized statistical models by score matching*, in "Journal of Machine Learning Research", 2005, vol. 6, p. 695–709.

[38] K.-C. LI.*Sliced Inverse Regression for Dimensional Reduction*, in "Journal of the American Statistical Association", 1991, vol. 86, p. 316–327.

[39] T. M. STOKER.*Consistent estimation of scaled coefficients*, in "Econometrica", 1986, vol. 54, p. 1461–1481.

Team TAPDANCE

Theory and Practice of Nanoscale Computing Engines

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
Computational Biology

Table of contents

Team TAPDANCE

Creation of the Team: 2016 June 03

Keywords:

Computer Science and Digital Science:

1.1.12. - Non-conventional architectures
1.3. - Distributed Systems
2.2. - Compilation
7.2. - Discrete mathematics, combinatorics

Other Research Topics and Application Domains:

5.3. - Nanotechnology
5.6. - Robotic systems

1. Members

Research Scientist

Damien Woods [Team leader, Inria, Advanced Research position, from Jun 2016]

2. Overall Objectives

2.1. Overall Objectives

In biological systems we see extraordinarily sophisticated growth processes, where molecular self-assembly is combined with active molecular components. Indeed, biological systems consume energy (e.g. ATP) and exhibit phenomena such as rapid growth in cell size and numbers, reconfiguration of internal components, molecular motors that push and pull large structures around, as well as molecular complexes, cells and whole organs that actively respond to the environment. Computer science gives us tools and methodologies to think about and design systems with large number of interacting components. Our goal is to bring these ideas together to design computational molecular systems.

The work of the newly-created TAPDANCE team will be concerned with the theory and practice of active DNA nanostructures that build structures and compute, all at the nanoscale.

We will focus on:

1. Proposing and analysing models of computation for nanoscale bimolecular systems. This includes finding new models for the systems we wish to build, proving theorems (e.g. about their computational power), as well as developing the theory of existing models.
2. Implementing these models in the wet-lab, primarily using DNA.
3. Software to design these kinds of systems (e.g. DNA sequence design) as well as coarse-grained molecular models for system analysis. Software tools are one of the main ways we bridge the gap between theory and experiments.

3. Research Program

3.1. Ongoing work

Recent theoretical work (Meunier, Woods "The non-cooperative tile assembly model is not intrinsically universal or capable of bounded Turing machine simulation") to be published in 2017 by has centered on the power of a model of self-assembly. In this model, called the noncooperative (or temperature 1) abstract Tile Assembly Model, square tiles assemble structures, called assemblies, in the discrete plane where each tile binds to a growing structure if one of its 4 coloured edges matches the colour of some available site on a growing assembly. It has been conjectured since 2000 that this model is not capable of computation or other sophisticated forms of growth. We show two results. One of our results states that time-bounded Turing machine computation is impossible in this model if we require the simulation to occur in a bounded rectangle in the plane. This result has a short proof that essentially follows from our other main result which states that this model is not "intrinsically universal". This latter result means that there is no single tileset in this model that can simulate any instance of the model, answering a question from and contrasting a result for the more general cooperative (temperature 2) model.

Other work by Woods has focused on experimentally implementing a wide class of Boolean circuits of a certain form. Experiments were mostly carried out at Caltech, and the work is in collaboration with colleagues at Caltech, UC Davis, Harvard and Cambridge and a publication is in preparation with [Woods, Doty, Myhrvold, Hui, Zhou, Yin, Winfree]. Details will be described in a future report subsequent to publication.

Work published earlier in 2016 (Erik D Demaine, Matthew J Patitz, Trent A Rogers, Robert T Schweller Scott M Summers and Damien Woods, "The two-handed tile assembly model is not intrinsically universal", Algorithmica 74:2, pages 812–850 (2016). not on HAL) shows results on a hierarchal model of algorithmic self-assembly called the two-handed self-assembly model (2HAM). Specifically, that the model is not intrinsically universal. In fact, we show that for all $\tau' < \tau$, each temperature-τ' 2HAM tile system does not simulate at least one temperature-τ 2HAM tile system. This impossibility result proves that the 2HAM is not intrinsically universal and stands in contrast to the fact that the (single-tile addition) abstract Tile Assembly Model is intrinsically universal. On the positive side, we prove that, for every fixed temperature $\tau \geq 2$, temperature-τ 2HAM tile systems are indeed intrinsically universal. In other words, for each τ there is a single intrinsically universal 2HAM tile set U_τ that, when appropriately initialized, is capable of simulating the behavior of any temperature-τ 2HAM tile system. As a corollary, we find an infinite set of infinite hierarchies of 2HAM systems with strictly increasing simulation power within each hierarchy. Finally, we show that for each τ , there is a temperature-τ 2HAM system that simultaneously simulates all temperature-τ 2HAM systems.

There are a number of projects being designed along the lines of topics above in Overall Objectives.

4. Highlights of the Year

4.1. Highlights of the Year

TAPDANCE Team created in June 2016.

A Starting Research Fellow, Pierre-Étiene Meunier, was hired by Inria to begin work with TAPDANCE in January 2017.

5. Partnerships and Cooperations

5.1. International Research Visitors

5.1.1. Visits of International Scientists

Prof. David Doty from UC Davis, California, was hosted for 1 week in 2016.

5.1.2. *Visits to International Teams*

Woods visited Caltech for several weeks in 2016.

Woods visited Dagstuhl 3-8 July 2016 for Caltech for several weeks in 2016. Dagstuhl Seminar 16271 Algorithmic Foundations of Programmable Matter. Collaborative work with workshop attendees. Invited talk.

6. Dissemination

6.1. Promoting Scientific Activities

6.1.1. *Scientific Events Organisation*

6.1.1.1. Chair of Conference Program Committees

Woods. Program committee (PC) co-chair for DNA22: The 22nd International Conference on DNA Computing and Molecular Programming, 2016. Munich, Germany (co-chairing with Yannick Rondelez, CNRS, ESPCI)

6.1.1.2. Member of the Conference Program Committees

Woods. AUTOMATA 2016. 22nd International Workshop on Cellular Automata & Discrete Complex Systems, ETH Zürich, Switzerland

6.1.1.3. Reviewer

Woods was reviewer for several conferences and journals (not listed for confidentiality reasons).

6.1.2. *Invited Talks*

- Woods. Transversal aspects of tilings, month-long workshop/course, Oléron, France. Week 1 lectures on Theory and Experiments with Algorithmic Self-Assembly. Invited lecture series.
- Woods. Dagstuhl Seminar 16271 on Algorithmic Foundations of Programmable Matter, 3-8 July 2016 (Germinay).
- Woods. Oxford University, Department of Computer Science, UK, 2016.
- Woods. Journées GT COA, Bordeaux. Evaluating a large class of Boolean circuits via algorithmic self-assembly of DNA strands. 28-29 Nov, 2016.
- Woods. 15éme Journées de la Matière Condensée, Bordeaux 22-26 Aug 2016 (JMC15). Evaluating a large class of Boolean circuits via algorithmic self-assembly of DNA strands

6.2. Teaching - Supervision - Juries

6.2.1. *Teaching*

Woods made preparations, including visits, to teach a 1-week school at ENS Lyon showing students both theoretical results and wet-lab experimental results. Also, students took part in wet-lab experiments, as well as carrying out projects in teams (involving both theory and experiments). The school occurred in the week of Jan 16-20, 2017.

6.2.2. *Juries*

In 2016 Woods was PhD examiner for: Frits Dannenberg. Oxford University, 2016 (Supervisors: Marta Kwiatkowska & Andrew Turberfield) Thesis title: Modelling and verification for DNA nanotechnology

7. Bibliography

Major publications by the team in recent years

[1] E. D. DEMAINE, M. J. PATITZ, T. A. ROGERS, R. T. SCHWELLER, S. M. SUMMERS, D. WOODS.*The two-handed tile assembly model is not intrinsically universal*, in "Algorithmica", 2016, vol. 74, n^{o} 2, p. 812–850.

[2] P.-É. MEUNIER, D. WOODS.*The non-cooperative tile assembly model is not intrinsically universal or capable of bounded Turing machine simulation*, In preparation.

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

Université Pierre et Marie Curie (Paris 6)

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:

1. - Architectures, systems and networks
1.1.1. - Multicore
2. - Software
2.1.6. - Concurrent programming
2.1.10. - Domain-specific languages
2.1.11. - Proof languages
2.2.1. - Static analysis
2.2.3. - Run-time systems
2.3.1. - Embedded systems
2.3.3. - Real-time systems
2.4. - Verification, reliability, certification
2.4.3. - Proofs
2.5. - Software engineering
2.6. - Infrastructure software
2.6.1. - Operating systems
2.6.2. - Middleware
2.6.3. - Virtual machines

Other Research Topics and Application Domains:

5. - Industry of the future
5.2.1. - Road vehicles
5.2.3. - Aviation
5.2.4. - Aerospace
6.1. - Software industry
6.1.1. - Software engineering
6.1.2. - Software evolution, maintenance
6.3.3. - Network Management
6.5. - Information systems
6.6. - Embedded systems

1. Members

Research Scientists

Gilles Muller [Team leader, Inria, Senior Researcher, HDR]
Julia Lawall [Inria, Senior Researcher, HDR]
Pierre-Évariste Dagand [CNRS, Researcher]

Faculty Member

Bertil Folliot [Univ. Paris VI, Professor, HDR]

Technical Staff

Quentin Lambert [Inria, until Oct 2016]

PhD Students

Cédric Courtaud [Thales, from Mar 2016, granted by CIFRE]
Redha Gouicem [Univ. Paris VI, from Oct 2016]
Antoine Blin [Inria, until July 2016, Engineer from November 2016]

Visiting Scientist

Greg Kroah-Hartman [Linux Foundation, from October 2016]

Administrative Assistants

Helene Milome [Inria]
Eugène Kamdem [UPMC]

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], [49], [87]. 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 [1], [8], [9] for transforming C programs (see Section 6.2), or domain-specific languages such as Devil [7] and Bossa [6] 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 [33], [36], [39] or dynamic analysis [57], [61], [67]. 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* [36]. 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* [48] 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 [76]. More recently, more general forms of symbolic execution [33] 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 [29], [31].

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. [41] 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 [34], 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 [21]. 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) [53], 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 [52].

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 [5] [62].

3.1.2.1. Traditional approach.

Using little languages to aid in software development is a tried-and-trusted technique [79] 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 [7]. 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 [6] 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) [3] 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, [22] 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 [42] and its descendant Pharo [26] – or multi-paradigm languages – such as the Scala programming language [64] – 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 [23].

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 [32], [27] 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 [3], 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 [4]. Interestingly, our development crucially relied on an existing Coq formalization, due to Braibant and Pous, [28] 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 [24], [25], [65], [68], [72], [74], [78][10], [20], 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 4500 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 [17], 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 start in March 2017, 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 [7] to interact with devices, Bossa [6] 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 [3]. 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. Dagand's recent work using the Coq proof assistant as an x86 macro-assembler [4] is a step in that direction, which belongs to a larger trend of hosting DSLs in dependent type theories [27], [63], [32]. 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 [49], 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, v4.9, comprises over 14 million lines of code, and supports 31 different families of CPU architectures, 73 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. [40], numerous research tools have been applied to the Linux kernel, typically for finding bugs [39], [56], [69], [80] or for computing software metrics [46], [85]. In our work, we have studied generic C bugs in Linux code [9], bugs in function protocol usage [50], [51], issues related to the processing of bug reports [73] and crash dumps [45], and the problem of backporting [68], 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. Section 6.3 presents our dissemination efforts to the Linux community.

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 [7], a domain-specific language for describing the low-level interface of a device. Chipounov *et al.* propose RevNic, [31] a template-based approach for porting device drivers from one OS to another. Ryzhyk *et al.* propose Termite, [70], [71] 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 [21], Coverity [40], CP-Miner, [55] PR-Miner [56], and Coccinelle [8]. 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 main highlight of the year is the continuous spreading of Coccinelle within the developer community of the Linux kernel. We submitted the first patches to the Linux kernel based on Coccinelle in 2007. Since then, over 4500 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. Another testimonial of the impact of our work is the signature of a Memorendum Of Understanding (MOU) with the Linux Foundation. As part of the MOU, Greg Kroah-Hartman will spend a year with Whisper starting in October 2016. Kroah-Hartman is one of the leading developers of the Linux kernel, and is one of only a few developers employed by the Linux Foundation, with another being Linus Torvalds. Greg participated in the activities of the Whisper team around the use of Coccinelle and research projects related to the Linux kernel, and he is a convinced ambassador of our research work.

Our work on Remote Core Locking (RCL) [10] was accepted in ACM Transaction in Computer Systems (TOCS) which is the most prestigious journal in systems. RCL is currently one of the most efficient locks for multicore architectures.

6. New Software and Platforms

6.1. Prequel

KEYWORDS: Code quality - Evolution - Infrastructure software

FUNCTIONAL DESCRIPTION

The commit history of a large, actively developed 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 imply 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 have developed Prequel (Patch Query Language) [20]. Prequel builds on the semantic patch lamguage SmPL developed for Coccinelle, which is now well known to the Linux kernel developer community, to allow developers to scan the changes in a source code development history, taking into account not only the specific changes made, but also the context in which these changes occur. As the history of a code base under active development quickly becomes large, with the Linux kernel incorporating around 13,000 commits on each 2-3 month release cycle, a particular goal in the development of Prequel has been to provide reasonable performance. Currently, most queries in our experiments complete in under minute when running on a single core on a standard laptop. So far, we have applied Prequel to the problem of understanding how to eliminate uses of deprecated functions [20], and are investigating how it may be useful in a systematic driver porting methodology.

Prequel is publicly available under GPLv2. The development of Prequel is supported by OSADL, and Julia Lawall presented Prequel at the 2016 OSADL networking day (https://www.osadl.org/OSADL-Networking-Day-2016.networking-day-2016.0.html).

- Participants: Julia Lawall and Gilles Muller
- Partners: IRILL - LIP6
- Contact: Julia Lawall
- URL: http://prequel-pql.gforge.inria.fr/

6.2. Coccinelle

KEYWORDS: Code quality - Evolution - Infrastructure software
FUNCTIONAL DESCRIPTION

Coccinelle is a tool for C code program matching and transformation that has been developed by members of the Whisper team over the last 10 years [8]. Coccinelle is widely used by the Linux kernel developer community and for other C software projects. Over the last three years, Coccinelle has benefited from the support of an engineer from the SED. Major improvements in 2016 include support for Python 3, independence from a no-longer-supported interface between Python and OCaml, better support for parallelism, and better support for integrating arbitrary predicates into the matching process. These features significantly improve performance and improve the uniformity of the rule specification language, thus providing a better experience for users. Coccinelle is at the foundation of much of our research work, including the ANR ITrans project, and these improvements will enhance and facilitate our research, accordingly.

Coccinelle is publicly available under GPLv2. In 2016, Julia Lawall presented Coccinelle in an invited keynote at the Linux Security Summit (http://events.linuxfoundation.org/events/archive/2016/linux-security-summit) and at a “birds of a feather” session at Linuxcon Europe (http://events.linuxfoundation.org/events/LinuxCon-europe).

- Participants: Julia Lawall, Gilles Muller, and Thierry Martinez
- Partners: IRILL - LIP6
- Contact: Julia Lawall
- URL: http://coccinelle.lip6.fr

6.3. Hector (BtrLinux)

KEYWORDS: Code quality - Evolution - Infrastructure software
FUNCTIONAL DESCRIPTION

A major source of errors in systems code is resource-release omission, which can lead to memory leaks and to crashes, if the system ends up in an inconsistent state. Currently, many tools exist that detect common patterns in software and detect faults as deviations from those patterns, but most suffer from high rates of false positives. Hector takes the novel approach of detecting inconsistencies local to a single function, and thus has been able to find over 300 faults in Linux kernel code and other C infrastructure software, with a rate of false positives of only 23%. Hector was originally the subject of the PhD thesis of Suman Saha [75]. Over the past two years, improving the robustness of the implementation of Hector has been the focus of ADT (young engineer position) BtrLinux supported by Inria, with the goal of making Hector publicly available and popularizing its use in the Linux kernel developer community. Some Linux kernel patches based on the use of Hector have been integrated into the Linux kernel, and the public release of Hector is in progress. The ADT position also involved the creation and maintenance of the website https://btrlinux.inria.fr/ as a showcase for the work of the Whisper team around Linux kernel development tools.

Building on his experience acquired in the ADT position, Quentin Lambert has recently been offered a position as an engineer at Wolfram MathCore AB.

- Participants: Quentin Lambert, Julia Lawall, and Gilles Muller
- Partners: IRILL - LIP6
- Contact: Julia Lawall
- URL: https://btrlinux.inria.fr/

6.4. ssrbit

FUNCTIONAL DESCRIPTION

ssrbit is a Coq library offering an efficient formalization of bit vectors, a refinement framework for abstractly reasoning about bitsets, and a trustworthy extraction of bit vectors to OCaml integers. Initially developed by Whisper members (Pierre-Évariste Dagand, Julia Lawall), the development has attracted an external contributor (Emilio Jesús Gallego Arias, postdoctoral researcher in CRI Mines-ParisTech), which led to significant improvements. We plan to improve the overall support and documentation so as to provide a full-featured library.

- Participants: Pierre Évariste Dagand, Julia Lawall, and Emilio Jesús Gallego Arias
- Contact: Pierre Évariste Dagand
- URL: https://github.com/ejgallego/ssrbit/

7. New Results

7.1. Software engineering for infrastructure software

Our main work in this area has focused on driver porting. We aim at fully automating the backporting (or symmetrically forward porting) process: given any driver for one Linux kernel version, one would like to obtain a driver that has the same functionality for another kernel version. This requires identifying the changes that are needed, obtaining examples of how to carry these changes out, and inferring from these examples a change that is appropriate for the given driver code. We have carried out a preliminary study in this direction with David Lo of Singapore Management University; this work, published at ICSME 2016 [17], 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.

More general automation of backporting requires more extensive search for relevant examples. This raises issues of scalability, because the Linux kernel code history is very large, and of expressivity, because we need to be able to express complex patterns to obtain change examples that are most relevant to a particular backporting problem. To this end, we have been adapted the notation used by Coccinelle, which describes how a change should be carried out, into a *patch query language* that allows describing patterns of changes that have been previously performed. The associated tool, Prequel, can find patches that match a particular pattern among several hundred thousand commits, often in tens of seconds [20]. This work is supported in part by OSADL, a consortium of companies, mostly in Germany, supporting the use and development of open source software in automation and other industries.

We will continue research in this direction over the next three years as part of the ANR PRCI ITrans project, awarded in 2016 and to be carried out in 2017-2020.

7.2. Developing infrastructure software using Domain Specific Languages

To bootstrap our long-term effort in designing safe and composable domain-specific languages, we have initiated two exploratory actions involving a combination of advanced type-theoretic concepts and domain-specific compilation techniques. Both actions are complementary, the first adopts a bottom-up approach – going from low-level artifacts to high-level abstractions – while the second follows a top-down approach – offering a safe translation of high-level guarantees to low-level executable code.

Our first line of inquiry, of which some early results have been published at FLOPS 2016 [13], aims at bridging the formalization gap between low-level, bit-twiddling code and high-level, mathematical abstractions. As such, it provided us with an opportunity to experiment with using an interactive theorem prover to design abstractions in a bottom-up manner. We have developed a library (`ssrbit`, publicly available under an open-source license) for modeling and computing with bit vectors in the Coq [35] proof assistant. Because ease of proving and efficiency in computing are often incompatible objectives, this library offers a two pronged approach by offering an abstract specification for proving and an efficient implementation for computing; we have shown that the latter is correct with respect to the former. Using this model of bit-level operations, we have implemented a bitset library and proved its correctness with respect to the formalization of sets of finite types provided by the Ssreflect library [43], which is part of the Mathematical Components framework developed at the MSR-Inria joint center. This library thus enables a seamless interaction of sets for computing and sets for proving. This library also supports the trustworthy extraction of bitsets down to OCaml's machine integers: we gained greater confidence in our model by adopting a methodology based on exhaustive testing. This enabled us to implement three bit-twiddling applications in Coq (Bloom filter, n-queens, and the efficient enumeration of all k-combinations of a set), prove their correctness and obtain efficient low-level OCaml code.

Our second line of inquiry is influenced by the realization that domain-specific languages are often treating the symptoms rather than providing a cure. Infrastructure software is often developed in C, which suffers from many semantic kludges and is, as a result, hardly amenable to formal reasoning. Many domain-specific languages are born out of the frustration of being unable to guarantee static properties of one's code: more often than not, the resulting language is little more than a domain-specific variant of Pascal supporting custom static analyses and some form of transliteration to C. To achieve safety and composability, we believe that a more holistic approach is called for, involving not only the design of a domain-specific *syntax* but also of a domain-specific *semantics*. Concretely, we are exploring the design of *certified domain-specific compilers* that integrate, from the ground up, a denotational and domain-specific semantics as part of the design of a domain-specific language. This vision is illustrated by our work on the safe compilation of Coq programs into secure OCaml code [14], [18]. It combines ideas from gradual typing – through which types are compiled into run-time assertions – and the theory of ornaments [37] – through which Coq datatypes can be related to OCaml datatypes. Within this formal framework, we enable a secure interaction, termed *dependent interoperability*, between correct-by-construction software and untrusted programs, be it system calls or legacy libraries. To do so, we trade static guarantees for runtime checks, thus allowing OCaml values to be safely coerced to dependently-typed Coq values and, conversely, to expose dependently-typed Coq programs defensively as OCaml programs. Our framework is developed in Coq: it is constructive and verified in the strictest sense of the terms. It thus becomes possible to internalize and hand-tune the extraction of dependently-typed programs to interoperable OCaml programs within Coq itself. This work is part of a collaboration with Eric Tanter, from the University of Chile, and Nicolas Tabareau, from the Ascola Inria project-team.

To further explore the realm of domain-specific compilers, we have been involved in the design and implementation of a certified compiler for the Lustre [30] synchronous dataflow language. Synchronous dataflow languages are widely used for the design of embedded systems: they allow a high-level description of the system and naturally lend themselves to a hierarchical design. This on-going work, in collaboration with members of the Parkas team and Gallium team of Inria Paris, formalizes the compilation of a synchronous data-flow language into an imperative sequential language, which is eventually translated to Cminor [54], one of CompCert's intermediate languages. This project illustrates perfectly our methodological position: the design of synchronous dataflow languages is first governed by semantic considerations (Kahn process networks and the synchrony hypothesis) that are then reifed into syntactic artefacts. The implementation of a certified compiler highlights this dependency on semantics, forcing us to give as crisp a semantics as possible for the proof effort to be manageable. This work is part of an on-going collaboration with Marc Pouzet and Tim Bourke, from the Parkas team of Inria Paris, Lionel Rieg, postdoc at Collège de France, and Xavier Leroy, from the Gallium Inria project-team.

In terms of DSL design for domains where correctness is critical, our current focus is on process scheduling and multicore architectures. Ten years ago, we developed Bossa, targeting process scheduling on unicore processors, and primarily focusing on the correctness of a scheduling policy with respect to the requirements of the target kernel. At that time, the main use cases were soft real-time applications, such as video playback. Bossa was and still continues to be used in teaching, because the associated verifications allow a student to develop a kernel-level process scheduling policy without the risk of a kernel crash. Today, however, there is again a need for the development of new scheduling policies, now targeting multicore architectures. As identified by Lozi *et al.* [59], large-scale server applications, having specific resource access properties, can exhibit pathological properties when run with the Linux kernel's various load balancing heuristics. We are working on a new domain-specific language, Ipanema, to allow expressing load balancing properties, and to enable verification of critical scheduling properties such as liveness; for the latter, we are exploring the use of tools such as the Z3 theorem prover from Microsoft, and the Leon theorem prover from EPFL. A first version of the language has been designed and we expect to have a prototype of Ipanema working next year. The work around Ipanema is the subject of a very active collaboration between researchers at four institutions (Inria, University of Nice, University of Grenoble, and EPFL (groups of V. Kuncak and W. Zwaenepoel)). Baptiste Lepers (EPFL) will be supported in 2017 as a postdoc as part of the Inria-EPFL joint laboratory.

Finally, in the context of the Multicore IPL, we are working with Jens Gustedt and Mariem Saeid of the Inria Camus project-team on developing a domain-specific language that eases programming with the ordered read-write lock (ORWL) execution model. The goal of this work is to provide a single execution model for parallel programs and to allow them to be deployed on multicore machines with varying architectures [16].

7.3. Run-time environments for multicore architectures

In the recent past, we acquired a solid expertise in multicore systems through the PhD of Jean-Pierre Lozi [60] and Florian David [38]. This expertise has led us to initiate several collaborations with industry partners, in the form of CIFRE PhD support. We first targeted real-time multicore systems with the goal of improving resource usage, through a cooperation with Renault and the PhD of Antoine Blin. Recently, we have started another cooperation on multicore real-time systems for avionics and space with Thales TRT, that is the topic of the PhD of Cédric Courtaud.

The PhD of Jean-Pierre Lozi [60] was on improving the performance locks on large multicore architectures. In an paper published at Usenix ATC 2012 [58], and more recently in an article published in 2016 in ACM Transactions on Computer Systems (TOCS) [10], we proposed a new locking technique, Remote Core Locking (RCL), that aims to accelerate the execution of critical sections in legacy applications on multicore architectures. RCL is currently one of the most efficient locking technique and the ATC 2012 paper has currently 67 citations on Google scholar. The idea of RCL is to replace lock acquisitions by optimized remote procedure calls to a dedicated server hardware thread. RCL limits the performance collapse observed with other lock algorithms when many threads try to acquire a lock concurrently and removes the need to transfer lock-protected shared data to the hardware thread acquiring the lock because such data can typically remain in the server's cache. Eighteen applications were used to evaluate RCL from standard multicore benchmark suites, such as SPLASH-2 and Phoenix 2. By using RCL instead of Linux POSIX locks, performance is improved by up to 2.5 times on Memcached, and up to 11.6 times on Berkeley DB with the TPC-C client. On a SPARC machine with two Sun Ultrasparc T2+ processors and 128 hardware threads, performance is improved by up to 1.3 times with respect to Solaris POSIX locks on Memcached, and up to 7.9 times on Berkeley DB with the TPC-C client.

The PhD of Antoine Blin is on modern complex embedded systems that involve a mix of real-time and best-effort applications. The recent emergence of low-cost multicore processors raises the possibility of running both kinds of applications on a single machine, with virtualization ensuring isolation. Nevertheless, memory contention can introduce other sources of delay, that can lead to missed deadlines. We first investigated the source of memory contention for the Mibench benchmark in a paper published at NETYS 2016 [12]. Then, in a paper published at ECRTS 2016 [11], we present a combined offline/online memory bandwidth monitoring approach. Our approach estimates and limits the impact of the memory contention incurred by the best-effort applications on the execution time of the real-time application. Using our approach, the system designer can limit the overhead on the real-time application to under 5% of its expected execution time, while still enabling progress of the best-effort applications.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Renault, 2014-2016, 45 000 euros. The purpose of this contract is to develop solutions for running a mix of real-time and best-effort applications on a small embedded multicore architecture. Our goal is to optimize the usage of the processor resource. The PhD of Antoine Blin is supported by a CIFRE fellowship with Renault.
- Orange Labs, 2016-2017, 60 000 euros. The purpose of this contract is to apply the techniques developed in the context of the PhD of Antoine Blin to the domain of Software Defined Networks where network functions are run using virtual machines on commodity multicore machines.
- Thales Research, 2016-2018, 45 000 euros. The purpose of this contract is to enable the usage of multicore architectures in avionics systems. More precisely, our goal is to develop optimizations for a software TDMA hypervisor developed by Thales that provides full time-isolation of tasks. The PhD of Cédric Courtaud is supported by a CIFRE fellowship with Thales Research.
- OSADL, 2016-2017, development of the Prequel patch query language, 20 000 euros. OSADL is an organization headquartered in Germany that promotes and supports the use of open source software in the automation and machine industry. The project is in the context of the OSADL project SIL2LinuxMP bringing together various companies in automotive and embedded sytems with the goal of developing methodologies for certifying the basic components of a GNU/Linux-based RTOS.

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.

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

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.

Chronos network, Time and Events in Computer Science, Control Theory, Signal Processing, Computer Music, and Computational Neurosciences and Biology - (2015 - 2016)

Coordinator: Gerard Berry

Whisper member: Gilles Muller

Funding: ANR 2014, Défi “Société de l’information et de la communication”.

The Chronos interdisciplinary network aims at placing in close contact and cooperation researchers of a variety of scientific fields: computer science, control theory, signal processing, computer music, neurosciences, and computational biology. The scientific object of study will be the understanding, modeling, and handling of time- and event-based computation across the fields.

Chronos will work by organizing a regular global seminar on subjects ranging from open questions to concrete solutions in the research fields, workshops gathering subsets of the Chronos researchers to address specific issues more deeply, a final public symposium presenting the main contributions and results, and an associated compendium.

9.2.2. Multicore Inria Project Lab

The Multicore IPL is an Inria initiative, led by Gilles Muller, whose goal is to develop techniques for deploying parallel programs on heterogeneous multicore machines while preserving scalability and performance. The IPL brings together researchers from the ALF, Algorille, CAMUS, Compsys, DALI, REGAL, Runtime and Whisper Inria Teams. These connections provide access to a diversity of expertise on open source development and parallel computing, respectively. In this context, Gilles Muller and Julia Lawall are working with Jens Gustedt and Mariem Saeid of Inria Lorraine on developing a domain-specific language that eases programming with the ordered read-write lock (ORWL) execution model. The goal of this work is to provide a single execution model for parallel programs and to allow them to be deployed on multicore machines with varying architectures.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

We collaborate with David Lo and Lingxiao Jiang of Singapore Management University, who are experts in software mining, clone detection, and information retrieval techniques. Our work with Lo and/or Jiang has led to 7 joint publications since 2013 [66], [77], [81], [82], [83], [86], [84], at conferences including ASE and ICSME.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

9.4.1.1. Internships

Natacha Crooks, PhD student at the University of Austin, Texas, spent three months in Whisper from May to August working on Ipanema.

Derek Palinski, undergraduate at Oberlin College, USA, spent January and June to August working on understanding of device driver evolution, including the evaluation of Prequel.

Vatika Harlalka, undergraduate at the International Institute of Information Technology - Hyderabad, India, spent May to July working on strategies for improving the performance of multicore real-time systems.

Denis Merigoux, final-year student from Ecole Polytechnique, spent March to August working on inference of Coccinelle semantic patches from examples.

Roman Delgado, undergraduate at UPMC, spent June to August working with Pierre-Évariste Dagand on implementing dependent induction in type theory.

Swaraj Dash, undergraduate at Cambridge University, spent August to September working with Pierre-Évariste Dagand on the derivative of indexed datatypes.

Redha Gouicem, Master 2 at UPMC, spent March to August working on memory access control for multicore real-time systems.

Axelle Piot, Master 2 at ENS, spent March to July working on Ipanema.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Selection

10.1.1.1. Member of the Conference Program Committees

- Pierre-Évariste Dagand: TyDe 2016 PC.
- Julia Lawall: FLOPS 2016 PC, PLDI 2016 EPC, LCTES 2016 PC, OOPSLA 2016 PC, ASE 2016 ERP, Middleware 2016 PC, GPCE 2016 PC, SLE 2016 PC, ICSME 2016 ERA.
- Gilles Muller: EDCC 2016 PC, OPODIS 2016 PC, DSN 2016 PC, ASPLOS 2016 PC.

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

- Julia Lawall: Editorial board of Science of Computer Programing (2008 - present).

10.1.2.2. Reviewer - Reviewing Activities

- Pierre-Évariste Dagand: Journal of Logical and Algebraic Methods in Programming (journal), Journal of Functional Programming (journal), Type-driven Development (workshop)
- Julia Lawall: Automated Software Engineering (journal).
- Gilles Muller: IEEE Transactions on Computers, Operating Systems Review.

10.1.3. Invited Talks

- Julia Lawall: PPL workshop (Japan) 2016, Linux Security Summit 2016, SPLASH 2016 Programming Languages Mentoring Workshop, IFIP WG 2.4 (Software Implementation Technology).

10.1.4. Research Administration

- Pierre-Évariste Dagand: Member of the steering committee for the Colloquium d'Informatique de L'UPMC Sorbonne Universités
- Julia Lawall: IFIP TC secretary (2012 - present).

  Hiring committees: Inria Paris (CR2, 2016), Bordeaux (MdC, 2016), CNAM (MdC, 2016)
- Gilles Muller: EuroSys steering committee (2013-2016), elected member of WG 10.4 (Dependability), representative of Inria in Sorbonne University's advisory committee for research, member of the project committee board of the Inria Paris Center, member of the Paris committee for allocating post-docs, PhD stipends and sabbaticals.
- Bertil Folliot: Elected member of the IFIP WG10.3 working group (Concurrent systems)

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Licence: Pierre-Évariste Dagand, Distributed cooperating objects, L3, UPMC, France
- Professional Licence: Bertil Folliot, Programmation C, L2, UPMC, France
- Professional Licence: Bertil Folliot, Lab projects, L2, UPMC, France

10.2.2. Supervision

- PhD in progress : Antoine Blin, CIFRE Renault, Vers une utilisation efficace des multi-coeurs dans des systèmes embarqués à criticités multiples, 30 Janvier 2017, Gilles Muller, Julien Sopéna (Regal)
- PhD in progress : Mariem Saeid, Jens Gustedt (Camus), soutenance en 2017, Gilles Muller.
- PhD in progress : Cédric Courtaud, CIFRE Thalès, 2016-2018, Gilles Muller, Julien Sopéna (Regal).
- PhD in progress : Redha Gouicem, 2016-2018, Gilles Muller, Julien Sopéna (Regal).

10.2.3. Juries

- Pierre-Évariste Dagand: member of the Jury of Mahsa Najafzadeh (UPMC)
- Julia Lawall: PhD reporter for Pierre Wilke (Rennes), Alan Charpentier (Bordeaux), Guido Lena Cota (Milan, defense in 2017), Krishna Narasimhan (Frankfurt, defense in 2017). PhD jury for Ripon Saha (UT Austin).
- Gilles Muller: President of the PhD thesis of T. Tigori (U. of Nantes), Member of the Jury of V. Trigonakis (EPFL, Switzerland), Reporter of the PhD of A. Walker (U. of New South Wales, Australia).

10.3. Popularization

10.3.1. Talks

- Julia Lawall: Coccinelle: invited talk, Linux Security Summit, 2016.

- Julia Lawall: Coccinelle BoF: Linuxcon Europe 2016.
- Julia Lawall: Outreachy intern panel, Linuxcon Europe 2016.
- Julia Lawall: Prequel, 2016 OSADL networking day.

11. Bibliography

Major publications by the team in recent years

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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

[10] J.-P. Lozi, F. David, G. Thomas, J. 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.

International Conferences with Proceedings

[11] A. BLIN, C. COURTAUD, J. SOPENA, J. LAWALL, G. MULLER.*Maximizing Parallelism without Exploding Deadlines in a Mixed Criticality Embedded System*, in "28th EUROMICRO Conference on Real-Time Systems (ECRTS'16)", Toulouse, France, July 2016, https://hal.inria.fr/hal-01346979.

[12] A. BLIN, C. COURTAUD, J. SOPENA, J. LAWALL, G. MULLER.*Understanding the Memory Consumption of the MiBench Embedded Benchmark*, in "Netys", Marakech, Morocco, May 2016, https://hal.inria.fr/hal-01349421.

[13] A. BLOT, P.-É. DAGAND, J. LAWALL.*From Sets to Bits in Coq*, in "FLOPS 2016", Kochi, Japan, March 2016, https://hal.archives-ouvertes.fr/hal-01251943.

[14] P.-E. DAGAND, N. TABAREAU, É. TANTER.*Partial Type Equivalences for Verified Dependent Interoperability*, in "ICFP 2016 - 21st ACM SIGPLAN International Conference on Functional Programming", Nara, Japan, September 2016, p. 298-310 [*DOI :* 10.1145/2951913.2951933], https://hal.inria.fr/hal-01328012.

[15] K. NARASIMHAN, C. REICHENBACH, J. LAWALL.*Interactive Data Representation Migration: Exploiting Program Dependence to Aid Program Transformation*, in "PEPM 2017 Workshop on Partial Evaluation and Program Manipulation", Paris, France, January 2017, https://hal.inria.fr/hal-01408266.

[16] M. SAIED, J. GUSTEDT, G. MULLER.*Automatic Code Generation for Iterative Multi-dimensional Stencil Computations*, in "High Performance Computing, Data, and Analitics", Hydarabat, India, A. BENOÎT (editor), IEEE, December 2016, https://hal.inria.fr/hal-01337093.

[17] F. THUNG, D. X. B. LE, D. LO, J. 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.

National Conferences with Proceeding

[18] T. BOURKE, P.-E. DAGAND, M. POUZET, L. RIEG.*Vérification de la génération modulaire du code impératif pour Lustre*, in "JFLA 2017 - Vingt-huitième Journées Francophones des Langages Applicatifs", Gourettes, France, January 2017, https://hal.inria.fr/hal-01403830.

Research Reports

[19] A. BLIN, C. COURTAUD, J. SOPENA, J. LAWALL, G. MULLER.*Maximizing Parallelism without Exploding Deadlines in a Mixed Criticality Embedded System*, Inria, February 2016, n° RR-8838, https://hal.inria.fr/hal-01268078.

[20] J. LAWALL, Q. LAMBERT, G. MULLER.*Prequel: A Patch-Like Query Language for Commit History Search*, Inria Paris, June 2016, n° RR-8918, https://hal.inria.fr/hal-01330861.

References in notes

[21] 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.

[22] 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.

[23] 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.

[24] T. F. BISSYANDÉ, L. RÉVEILLÈRE, J. 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.

[25] T. F. BISSYANDÉ, L. RÉVEILLÈRE, J. 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.

[26] A. P. BLACK, S. DUCASSE, O. NIERSTRASZ, D. POLLET.*Pharo by Example*, Square Bracket Associates, 2010.

[27] 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.

[28] 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.

[29] 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.

[30] P. CASPI, N. HALBWACHS, D. PILAUD, J. PLAICE.*Lustre: a declarative language for programming synchronous systems*, in "14th ACM Symposium on Principles of Programming Languages", ACM, 1987.

[31] V. CHIPOUNOV, G. CANDEA.*Reverse Engineering of Binary Device Drivers with RevNIC*, in "EuroSys", 2010, p. 167–180.

[32] A. CHLIPALA.*The Bedrock Structured Programming System: Combining Generative Metaprogramming and Hoare Logic in an Extensible Program Verifier*, in "ICFP", 2013, p. 391–402.

[33] 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.

[34] 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.

[35] COQ DEVELOPMENT TEAM.*The Coq proof assistant reference manual*, 2015, http://coq.inria.fr.

[36] P. COUSOT, R. COUSOT.*Abstract Interpretation: Past, Present and Future*, in "CSL-LICS", 2014, p. 2:1–2:10.

[37] P.-É. DAGAND.*Reusability and Dependent Types*, University of Strathclyde, 2013.

[38] F. DAVID.*Continuous and Efficient Lock Profiling for Java on Multicore Architectures*, Université Pierre et Marie Curie - Paris VI, July 2015, https://hal.inria.fr/tel-01263203.

[39] I. DILLIG, T. DILLIG, A. AIKEN.*Sound, complete and scalable path-sensitive analysis*, in "PLDI", June 2008, p. 270–280.

[40] 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.

[41] 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.

[42] A. GOLDBERG, D. ROBSON.*Smalltalk-80: The Language and Its Implementation*, Addison-Wesley, 1983.

[43] G. GONTHIER, A. MAHBOUBI, E. TASSI.*A Small Scale Reflection Extension for the Coq system*, Inria Saclay Ile de France, 2015, n^{o} RR-6455.

[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] G. A. KILDALL.*A Unified Approach to Global Program Optimization*, in "POPL", 1973, p. 194–206.

[49] 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.

[50] 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.

[51] 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.

[52] 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.

[53] 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.

[54] X. LEROY.*Formal verification of a realistic compiler*, in "Communications of the ACM", 2009, vol. 52, n^{o} 7, p. 107–115.

[55] 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.

[56] 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.

[57] D. LO, S. KHOO.*SMArTIC: towards building an accurate, robust and scalable specification miner*, in "FSE", 2006, p. 265–275.

[58] J.-P. LOZI, F. DAVID, G. THOMAS, J. L. LAWALL, G. MULLER.*Remote Core Locking: migrating critical-section execution to improve the performance of multithreaded applications*, in "USENIX Annual Technical Conference", Boston, MA, USA, June 2012, p. 65–76.

[59] J. LOZI, B. LEPERS, J. R. FUNSTON, F. GAUD, V. QUÉMA, A. FEDOROVA.*The Linux scheduler: a decade of wasted cores*, in "Proceedings of the Eleventh European Conference on Computer Systems, EuroSys 2016, London, United Kingdom, April 18-21, 2016", C. CADAR, P. PIETZUCH, K. KEETON, R. RODRIGUES (editors), ACM, 2016, p. 1:1–1:16, http://doi.acm.org/10.1145/2901318.2901326.

[60] J.-P. LOZI.*Towards more scalable mutual exclusion for multicore architectures*, Université Pierre et Marie Curie - Paris VI, July 2014, https://tel.archives-ouvertes.fr/tel-01067244.

[61] 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.

[62] 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.

[63] 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.

[64] M. ODERSKY, T. ROMPF.*Unifying functional and object-oriented programming with Scala*, in "Commun. ACM", 2014, vol. 57, n^{o} 4, p. 76–86.

[65] 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.

[66] K. PAVNEET SINGH, F. THUNG, D. LO, J. LAWALL.*An Empirical Study on the Adequacy of Testing in Open Source Projects*, in "21st Asia-Pacific Software Engineering Conference", Jeju, South Korea, December 2014, https://hal.inria.fr/hal-01096132.

[67] 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.

[68] L. R. RODRIGUEZ, J. 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.

[69] 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.

[70] L. RYZHYK, P. CHUBB, I. KUZ, E. LE SUEUR, G. HEISER.*Automatic device driver synthesis with Termite*, in "SOSP", 2009, p. 73–86.

[71] 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.

[72] 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.

[73] 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.

[74] S. SAHA, J.-P. LOZI, G. THOMAS, J. 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.

[75] S. SAHA.*Improving the Quality of Error-Handling Code in Systems Software using Function-Local Information*, Université Pierre et Marie Curie - Paris VI, March 2013, https://tel.archives-ouvertes.fr/tel-00937807.

[76] D. A. SCHMIDT.*Data Flow Analysis is Model Checking of Abstract Interpretations*, in "POPL", 1998, p. 38–48.

[77] P. SENNA, L. RÉVEILLÈRE, L. JIANG, D. LO, J. LAWALL, G. MULLER.*Understanding the genetic makeup of Linux device drivers*, in "PLOS'13 - 7th Workshop on Programming Languages and Operating Systems", Nemacolin Woodlands Resort, Pennsylvania, United States, ACM, November 2013 [*DOI :* 10.1145/2525528.2525536], https://hal.inria.fr/hal-00927070.

[78] P. SENNA TSCHUDIN, J. 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.

[79] M. SHAPIRO.*Purpose-built languages*, in "Commun. ACM", 2009, vol. 52, n^{o} 4, p. 36–41.

[80] 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.

[81] F. THUNG, D. LO, J. L. LAWALL.*Automated library recommendation*, in "WCRE 2013 - 20th Working Conference on Reverse Engineering", Koblenz, Germany, R. LÄMMEL, R. OLIVETO, R. ROBBES (editors), IEEE, October 2013, p. 182-191 [*DOI :* 10.1109/WCRE.2013.6671293], https://hal.inria.fr/hal-00918076.

[82] F. THUNG, S. WANG, D. LO, J. LAWALL.*Automatic recommendation of API methods from feature requests*, in "ASE 2013 - 28th IEEE/ACM International Conference on Automated Software Engineering", Palo Alto, California, United States, E. DENNEY, T. BULTAN, A. ZELLER (editors), IEEE, November 2013, https://hal.inria.fr/hal-00918828.

[83] Y. TIAN, D. LO, J. LAWALL.*Automated construction of a software-specific word similarity database*, in "2014 Software Evolution Week - IEEE Conference on Software Maintenance, Reengineering, and Reverse Engineering, CSMR-WCRE", Antwerp, Belgium, IEEE, February 2014, p. 44-53, https://hal.inria.fr/hal-01086077.

[84] Y. TIAN, D. LO, J. LAWALL.*SEWordSim: software-specific word similarity database*, ACM, May 2014, p. 568-571, ICSE Companion 2014 - Companion Proceedings of the 36th International Conference on Software Engineering, Poster [*DOI :* 10.1145/2591062.2591071], https://hal.inria.fr/hal-01086079.

[85] W. WANG, M. GODFREY.*A Study of Cloning in the Linux SCSI Drivers*, in "Source Code Analysis and Manipulation (SCAM)", IEEE, 2011.

[86] S. WANG, D. LO, J. LAWALL.*Compositional Vector Space Models for Improved Bug Localization*, in "30th International Conference on Software Maintenance and Evolution", Victoria, Canada, IEEE, September 2014, p. 171-180, https://hal.inria.fr/hal-01086084.

[87] J. YANG, C. HAWBLITZEL.*Safe to the Last Instruction: Automated Verification of a Type-safe Operating System*, in "PLDI", 2010, p. 99–110.

Project-Team WILLOW

Models of visual object recognition and scene understanding

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

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Paris

RESEARCH CENTER
Paris

THEME
Vision, perception and multimedia interpretation

Table of contents

Project-Team WILLOW

Creation of the Project-Team: 2007 June 01

Keywords:

Computer Science and Digital Science:

3.1.1. - Modeling, representation
3.4. - Machine learning and statistics
5.3. - Image processing and analysis
5.4. - Computer vision
8. - Artificial intelligence
8.1. - Knowledge
8.2. - Machine learning

Other Research Topics and Application Domains:

9.4.1. - Computer science
9.4.5. - Data science

1. Members

Research Scientists

Minsu Cho [Inria, Starting Research position, until Aug 2016]
Ivan Laptev [Inria, Senior Researcher, HDR]
Josef Sivic [Inria, Senior Researcher, HDR]

Faculty Member

Jean Ponce [Team leader, ENS Paris, Professor]

Technical Staff

Jonathan Chemla [Inria, until Aug 2016]
Petr Gronat [Inria]
Antony Marion [Inria, until Mar 2016]
Ignacio Rocco Spremolla [Inria, from Apr 2016]

PhD Students

Guilhem Cheron [Inria]
Theophile Dalens [Inria]
Vadim Kantorov [Inria]
Antoine Miech [Inria, from Oct 2016]
Maxime Oquab [Inria]
Julia Peyre [Inria]
Rafael Sampaio de Rezende [Inria]
Guillaume Seguin [ENS Paris, until Aug 2016]
Matthew Trager [Inria]
Gul Varol Simsekli [Inria]
Piotr Bojanowski [Inria, until Mar 2016]
Tuan Hung Vu [Inria]

Post-Doctoral Fellows

Anton Osokin [Inria]
Relja Arandjelovic [Inria, until Jul 2016]

Andrei Bursuc [Inria, until Oct 2016]
Bumsub Ham [Inria, until Jul 2016]
Suha Kwak [Inria, until Mar 2016]

Visiting Scientists

John Canny [UC Berkeley, Professor, Inria International Chair]
Alexei Efros [UC Berkeley, May-Jun 2016]
Sergiu Irimie [Inria, until Aug 2016]
Phillip Isola [UC Berkeley, Jun 2016]
Oleh Rybkin [Inria, Sep 2016]
Richard Zhang [UC Berkeley, Jun 2016]

Administrative Assistants

David Dinis [Inria, until Apr 2016]
Sarah Le [Inria, from Jul 2016]

Others

Kai Han [Inria]
Mathieu Aubry [ENPC]
Gunnar Atli Sigurdsson [Carnegie Mellon University]
Pavel Trutman [Czech Technical University]

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 Rocquencourt, Ecole Normale Supérieure (ENS) and Centre National de la Recherche Scientifique (CNRS).

This year we have hired two new Phd students: Antoine Miech (Inria) and Ignacio Rocco (inria). Alexei Efros (Professor, UC Berkeley, USA) visited Willow during May-June with his postdoc Phillip Isola and Phd student Richard Zhang. John Canny (Professor, UC Berkeley, USA) visited Willow within the framework of Inria's International Chair program.

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.

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

- Jean Ponce (together with Svetlana Lazebnik and Cordelia Schmid) received the Longuet-Higgins Prize for "Fundamental contributions in Computer Vision", awarded at the IEEE Conference on Computer Vision and Pattern Recognition, 2016.

6. New Software and Platforms

6.1. NetVLAD: CNN architecture for weakly supervised place recognition

Open source release of the software package for our paper "NetVLAD: CNN architecture for weakly supervised place recognition" [9]. It provides a full implementation of the method, including code for weakly supervised training of the CNN representation, testing on standard datasets, as well as trained models. Links to all of these are available at our project page http://www.di.ens.fr/willow/research/netvlad/.

6.2. Unsupervised learning from narrated instruction videos

Open source release of the software package for our paper "Unsupervised learning from narrated instruction videos" . It provides a full implementation of the method, including code for weakly supervised training from instruction video, as well as trained models. Links to all of these are available at our project page http://www.di.ens.fr/willow/research/instructionvideos/.

6.3. ContextLocNet: Context-aware deep network models for weakly supervised localization

Open source release of code reproducing the results in our "ContextLocNet: Context-aware deep network models for weakly supervised localization" [11]. It provides code for training models, testing on standard datasets and trained models. It can be found online at https://github.com/vadimkantorov/contextlocnet.

6.4. Long-term Temporal Convolutions for Action Recognition

Open source release of the software package for our paper "Long-term Temporal Convolutions for Action Recognition" [20]. It provides code for training models, testing on standard datasets and trained models. Links are available at our project page http://www.di.ens.fr/willow/research/ltc/.

7. New Results

7.1. 3D object and scene modeling, analysis, and retrieval

7.1.1. Trinocular Geometry Revisited

Participants: Jean Ponce, Martial Hebert, Matthew Trager.

When do the visual rays associated with triplets of point correspondences converge, that is, intersect in a common point? Classical models of trinocular geometry based on the fundamental matrices and trifocal tensor associated with the corresponding cameras only provide partial answers to this fundamental question, in large part because of underlying, but seldom explicit, general configuration assumptions. In this project, we use elementary tools from projective line geometry to provide necessary and sufficient geometric and analytical conditions for convergence in terms of transversals to triplets of visual rays, without any such assumptions. In turn, this yields a novel and simple minimal parameterization of trinocular geometry for cameras with non-collinear or collinear pinholes, which can be used to construct a practical and efficient method for trinocular geometry parameter estimation. This work has been published at CVPR 2014, and a revised version that includes numerical experiments using synthetic and real data has been published in IJCV [7] and example results are shown in figure 1.

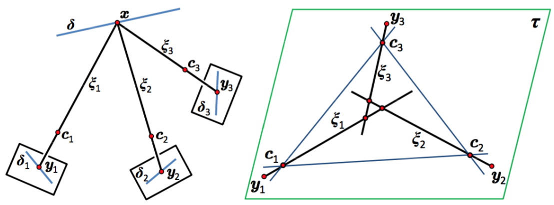

Figure 1. Left: Visual rays associated with three (correct) correspondences. Right: Degenerate epipolar constraints associated with three coplanar, but non-intersecting rays lying in the trifocal plane.

7.1.2. Consistency of silhouettes and their duals

Participants: Matthew Trager, Martial Hebert, Jean Ponce.

Silhouettes provide rich information on three-dimensional shape, since the intersection of the associated visual cones generates the "visual hull", which encloses and approximates the original shape. However, not all silhouettes can actually be projections of the same object in space: this simple observation has implications in object recognition and multi-view segmentation, and has been (often implicitly) used as a basis for camera calibration. In this paper, we investigate the conditions for multiple silhouettes, or more generally arbitrary closed image sets, to be geometrically "consistent". We present this notion as a natural generalization of traditional multi-view geometry, which deals with consistency for points. After discussing some general results, we present a "dual" formulation for consistency, that gives conditions for a family of planar sets to be sections of the same object. Finally, we introduce a more general notion of silhouette "compatibility" under partial knowledge of the camera projections, and point out some possible directions for future research. This work has been published in [16] and example results are shown in 2.

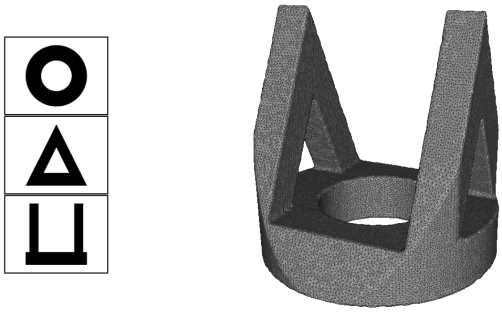

Figure 2. Geometrically consistent silhouettes are feasible projections of a single object.

7.1.3. Congruences and Concurrent Lines in Multi-View Geometry

Participants: Jean Ponce, Bernd Sturmfels, Matthew Trager.

We present a new framework for multi-view geometry in computer vision. A camera is a mapping between P^3 and a line congruence. This model, which ignores image planes and measurements, is a natural abstraction of traditional pinhole cameras. It includes two-slit cameras, pushbroom cameras, catadioptric cameras, and many more. We study the concurrent lines variety, which consists of n-tuples of lines in P^3 that intersect at a point. Combining its equations with those of various congruences, we derive constraints for corresponding images in multiple views. We also study photographic cameras which use image measurements and are modeled as rational maps from P^3 to P^2 or $P^1 \times P^1$. This work has been accepted for publication in [19] and example results are shown in 3.

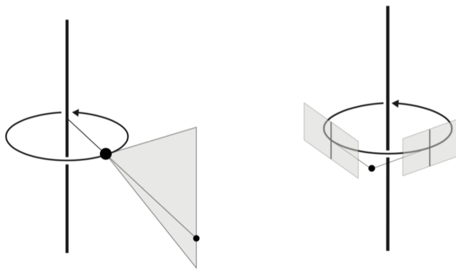

Figure 3. Non-central panoramic (left) and stereo panoramic cameras (right) are examples of non-linear cameras that can be modeled using line congruences.

7.1.4. NetVLAD: CNN architecture for weakly supervised place recognition

Participants: Relja Arandjelović, Petr Gronat, Akihiko Torii, Tomas Pajdla, Josef Sivic.

In [9], we tackle the problem of large scale visual place recognition, where the task is to quickly and accurately recognize the location of a given query photograph. We present the following three principal contributions. First, we develop a convolutional neural network (CNN) architecture that is trainable in an end-to-end manner directly for the place recognition task. The main component of this architecture, NetVLAD, is a new generalized VLAD layer, inspired by the "Vector of Locally Aggregated Descriptors" image representation commonly used in image retrieval. The layer is readily pluggable into any CNN architecture and amenable to training via backpropagation. Second, we develop a training procedure, based on a new weakly supervised ranking loss, to learn parameters of the architecture in an end-to-end manner from images depicting the same places over time downloaded from Google Street View Time Machine. Finally, we show that the proposed architecture obtains a large improvement in performance over non-learnt image representations as well as significantly outperforms off-the-shelf CNN descriptors on two challenging place recognition benchmarks. This work has been published at CVPR 2016 [9]. Figure 4 shows some qualitative results.

(a) Mobile phone query (b) Retrieved image of same place

Figure 4. Our trained NetVLAD descriptor correctly recognizes the location (b) of the query photograph (a) despite the large amount of clutter (people, cars), changes in viewpoint and completely different illumination (night vs daytime).

7.1.5. Pairwise Quantization

Participants: Artem Babenko, Relja Arandjelović, Victor Lempitsky.

We consider the task of lossy compression of high-dimensional vectors through quantization. We propose the approach that learns quantization parameters by minimizing the distortion of scalar products and squared distances between pairs of points. This is in contrast to previous works that obtain these parameters through the minimization of the reconstruction error of individual points. The proposed approach proceeds by finding a linear transformation of the data that effectively reduces the minimization of the pairwise distortions to the minimization of individual reconstruction errors. After such transformation, any of the previously-proposed quantization approaches can be used. Despite the simplicity of this transformation, the experiments demonstrate that it achieves considerable reduction of the pairwise distortions compared to applying quantization directly to the untransformed data. This work has been published on arXiv [18] and submitted to Neurocomputing journal.

7.1.5.1. Learning and Calibrating Per-Location Classifiers for Visual Place Recognition

Participants: Petr Gronat, Josef Sivic, Guillaume Obozinski [ENPC / Inria SIERRA], Tomáš Pajdla [CTU in Prague].

The aim of this work is to localize a query photograph by finding other images depicting the same place in a large geotagged image database. This is a challenging task due to changes in viewpoint, imaging conditions and the large size of the image database. The contribution of this work is two-fold. First, we cast the place recognition problem as a classification task and use the available geotags to train a classifier for each location in the database in a similar manner to per-exemplar SVMs in object recognition. Second, as only few positive training examples are available for each location, we propose a new approach to calibrate all the per-location SVM classifiers using *only* the negative examples. The calibration we propose relies on a significance measure essentially equivalent to the p-values classically used in statistical hypothesis testing. Experiments are performed on a database of 25,000 geotagged street view images of Pittsburgh and demonstrate improved place recognition accuracy of the proposed approach over the previous work. This work has been published at CVPR 2013, and a revised version that includes additional experimental results has been published at IJCV [3].

7.2. Category-level object and scene recognition

7.2.1. Proposal Flow

Participants: Bumsub Ham, Minsu Cho, Cordelia Schmid, Jean Ponce.

Finding image correspondences remains a challenging problem in the presence of intra-class variations and large changes in scene layout, typical in scene flow computation. In [10], we introduce a novel approach to this problem, dubbed proposal flow, that establishes reliable correspondences using object proposals. Unlike prevailing scene flow approaches that operate on pixels or regularly sampled local regions, proposal flow benefits from the characteristics of modern object proposals, that exhibit high repeatability at multiple scales, and can take advantage of both local and geometric consistency constraints among proposals. We also show that proposal flow can effectively be transformed into a conventional dense flow field. We introduce a new dataset that can be used to evaluate both general scene flow techniques and region-based approaches such as proposal flow. We use this benchmark to compare different matching algorithms, object proposals, and region features within proposal flow with the state of the art in scene flow. This comparison, along with experiments on standard datasets, demonstrates that proposal flow significantly outperforms existing scene flow methods in various settings. This work has been published at CVPR 2016 [10]. The proposed method and its qualitative result are illustrated in Figure 5.

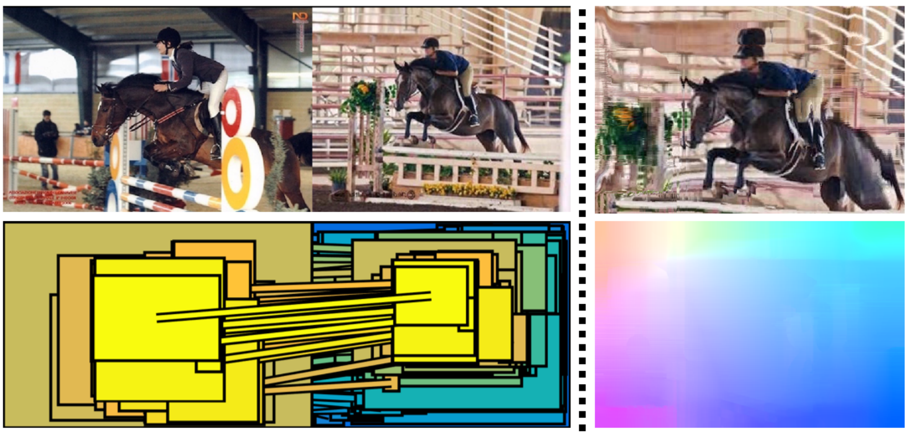

Figure 5. Proposal flow generates a reliable scene flow between similar images by establishing geometrically consistent correspondences between object proposals. (Left) Region-based scene flow by matching object proposals. (Right) Color-coded dense flow field generated from the region matches, and image warping using the flow.

7.2.1.1. Learning Discriminative Part Detectors for Image Classification and Cosegmentation

Participants: Jian Sun, Jean Ponce.

In this work, we address the problem of learning discriminative part detectors from image sets with category labels. We propose a novel latent SVM model regularized by group sparsity to learn these part detectors. Starting from a large set of initial parts, the group sparsity regularizer forces the model to jointly select and optimize a set of discriminative part detectors in a max-margin framework. We propose a stochastic version of a proximal algorithm to solve the corresponding optimization problem. We apply the proposed method to image classification and cosegmentation, and quantitative experiments with standard bench- marks show that it matches or improves upon the state of the art. The first version of this work has appeared at CVPR 2013. An extended version has been published at IJCV [6].

7.2.2. ContextLocNet: Context-aware deep network models for weakly supervised localization

Participants: Vadim Kantorov, Maxime Oquab, Minsu Cho, Ivan Laptev.

In [11] we aim to localize objects in images using image-level supervision only. Previous approaches to this problem mainly focus on discriminative object regions and often fail to locate precise object boundaries. In [11] we address this problem by introducing two types of context-aware guidance models, additive and contrastive models, that leverage their surrounding context regions to improve localization. The additive model encourages the predicted object region to be supported by its surrounding context region. The contrastive model encourages the predicted object region to be outstanding from its surrounding context region. Our approach benefits from the recent success of convolutional neural networks for object recognition and extends Fast R-CNN to weakly supervised object localization. Extensive experimental evaluation on the PASCAL VOC 2007 and 2012 benchmarks shows hat our context-aware approach significantly improves weakly supervised localization and detection. A high-level architecture of our model is presented in Figure 6, the project webpage is at http://www.di.ens.fr/willow/research/contextlocnet/.

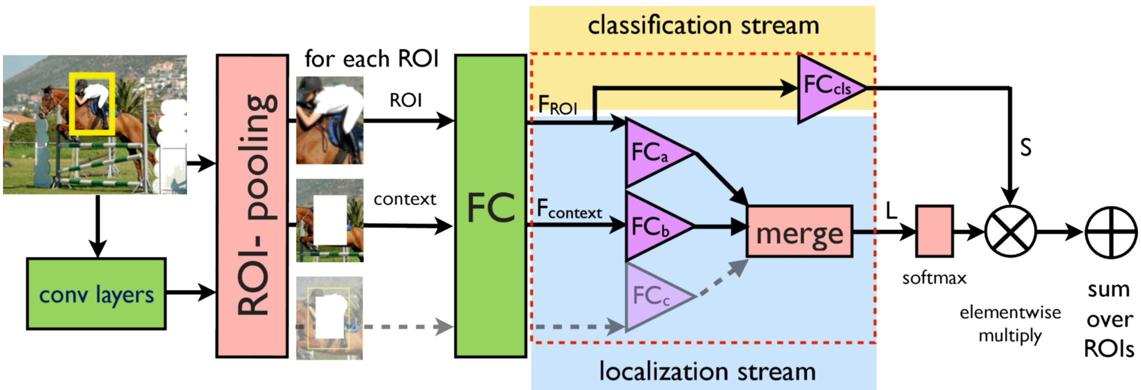

Figure 6. ContextLocNet improves localization by comparing an object score between a proposal and its context.

7.2.3. Faces In Places: Compound query retrieval

Participants: Yujie Zhong, Relja Arandjelović, Andrew Zisserman.

The goal of this work is to retrieve images containing both a target person and a target scene type from a large dataset of images. At run time this compound query is handled using a face classifier trained for the person, and an image classifier trained for the scene type. We make three contributions: first, we propose a hybrid convolutional neural network architecture that produces place-descriptors that are aware of faces and their corresponding descriptors. The network is trained to correctly classify a combination of face and scene classifier scores. Second, we propose an image synthesis system to render high quality fully-labelled face-and-place images, and train the network only from these synthetic images. Last, but not least, we collect and annotate a dataset of real images containing celebrities in different places, and use this dataset to evaluate the retrieval system. We demonstrate significantly improved retrieval performance for compound queries using the new face-aware place-descriptors. This work has been published at BMVC 2016 [17]. Figure 7 shows some qualitative results.

Figure 7. Examples of the top two retrieved images for various compound queries.

7.3. Image restoration, manipulation and enhancement

7.3.1. Robust Guided Image Filtering Using Nonconvex Potentials

Participants: Bumsub Ham, Minsu Cho, Jean Ponce.

Filtering images using a guidance signal, a process called joint or guided image filtering, has been used in various tasks in computer vision and computational photography, particularly for noise reduction and joint upsampling. The aim is to transfer the structure of the guidance signal to an input image, restoring noisy or altered image structure. The main drawbacks of such a data-dependent framework are that it does not consider differences in structure between guidance and input images, and it is not robust to outliers. We propose a novel SD (for static/dynamic) filter to address these problems in a unified framework by jointly leveraging structural information of guidance and input images. Joint image filtering is formulated as a nonconvex optimization problem, which is solved by the majorization-minimization algorithm. The proposed algorithm converges quickly while guaranteeing a local minimum. The SD filter effectively controls the underlying image structure at different scales and can handle a variety of types of data from different sensors. It is robust to outliers and other artifacts such as gradient reversal and global intensity shifting, and has good edge-preserving smoothing properties. We demonstrate the flexibility and effectiveness of the SD filter in a great variety of applications including depth upsampling, scale-space filtering, texture removal, flash/non-flash denoising, and RGB/NIR denoising. This has been published at CVPR 2015. A new revised version is currently in submission [4]. The SD filter is illustrated in Figure 8.

7.4. Human activity capture and classification

7.4.1. Hollywood in Homes: Crowdsourcing Data Collection for Activity Understanding

Participants: Gunnar A. Sigurdsson, Gül Varol, Xiaolong Wang, Ali Farhadi, Ivan Laptev, Abhinav Gupta.

Computer vision has a great potential to help our daily lives by searching for lost keys, watering flowers or reminding us to take a pill. To succeed with such tasks, computer vision methods need to be trained from real and diverse examples of our daily dynamic scenes. While most of such scenes are not particularly exciting, they typically do not appear on YouTube, in movies or TV broadcasts. So how do we collect sufficiently many diverse but boring samples representing our lives? We propose a novel Hollywood in Homes approach to collect such data. Instead of shooting videos in the lab, we ensure diversity by distributing and crowdsourcing the whole process of video creation from script writing to video recording and annotation. Following this procedure we collect a new dataset, *Charades*, with hundreds of people recording videos in their own homes, acting out casual everyday activities (see Figure 9). The dataset is composed of 9,848 annotated videos with an

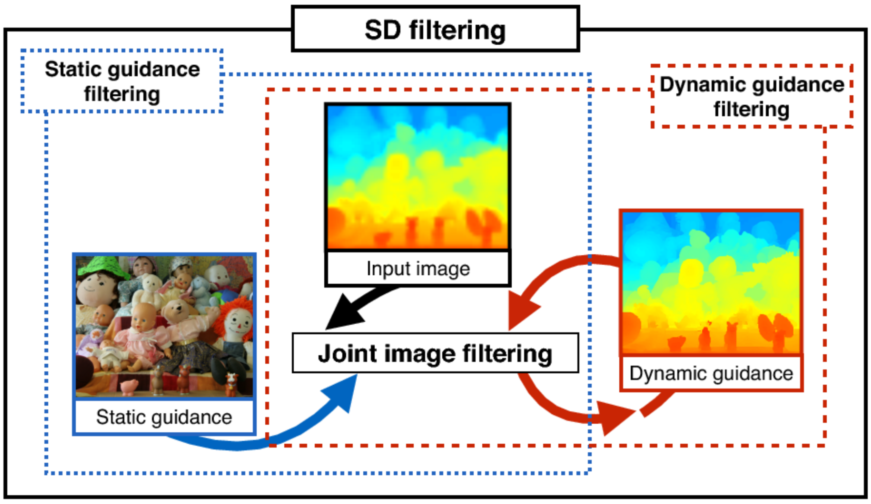

Figure 8. Sketch of joint image filtering and SD filtering: Static guidance filtering convolves an input image with a weight function computed from static guidance, as in the dotted blue box. Dynamic guidance filtering uses weight functions that are repeatedly obtained from regularized input images, as in the dotted red box. We have observed that static and dynamic guidance complement each other, and exploiting only one of them is problematic, especially in the case of data from different sensors (e.g., depth and color images). The SD filter takes advantage of both, and addresses the problems of current joint image filtering.

average length of 30 seconds, showing activities of 267 people from three continents. Each video is annotated by multiple free-text descriptions, action labels, action intervals and classes of interacted objects. In total, Charades provides 27,847 video descriptions, 66,500 temporally localized intervals for 157 action classes and 41,104 labels for 46 object classes. Using this rich data, we evaluate and provide baseline results for several tasks including action recognition and automatic description generation. We believe that the realism, diversity, and casual nature of this dataset will present unique challenges and new opportunities for computer vision community. This work has been published at ECCV 2016 [15].

Figure 9. Comparison of actions in the Charades dataset and on YouTube: Reading a book, Opening a refrigerator, Drinking from a cup. YouTube returns entertaining and often atypical videos, while Charades contains typical everyday videos.

7.4.2. Unsupervised learning from narrated instruction videos

Participants: Jean-Baptiste Alayrac, Piotr Bojanowski, Nishant Agrawal, Josef Sivic, Ivan Laptev, Simon Lacoste-Julien.

In [8], we address the problem of automatically learning the main steps to complete a certain task, such as changing a car tire, from a set of narrated instruction videos. The contributions of this paper are three-fold. First, we develop a new unsupervised learning approach that takes advantage of the complementary nature of the input video and the associated narration. The method solves two clustering problems, one in text and one in video, applied one after each other and linked by joint constraints to obtain a single coherent sequence of steps in both modalities. Second, we collect and annotate a new challenging dataset of real-world instruction videos from the Internet. The dataset contains about 800,000 frames for five different tasks that include complex interactions between people and objects, and are captured in a variety of indoor and outdoor settings. Third, we experimentally demonstrate that the proposed method can automatically discover, in an unsupervised manner, the main steps to achieve the task and locate the steps in the input videos. This work has been published at CVPR 2016 [8].

7.4.3. Long-term Temporal Convolutions for Action Recognition

Participants: Gul Varol, Ivan Laptev, Cordelia Schmid.

Typical human actions such as hand-shaking and drinking last several seconds and exhibit characteristic spatio-temporal structure. Recent methods attempt to capture this structure and learn action representations with convolutional neural networks. Such representations, however, are typically learned at the level of single frames or short video clips and fail to model actions at their full temporal scale. In [20], we learn video representations using neural networks with long-term temporal convolutions. We demonstrate that CNN models with increased temporal extents improve the accuracy of action recognition despite reduced spatial resolution. We also study the impact of different low-level representations, such as raw values of video pixels and optical flow vector fields and demonstrate the importance of high-quality optical flow estimation for learning accurate action models. We report state-of-the-art results on two challenging benchmarks for human action recognition UCF101 and HMDB51. This work is under review. The results for the proposed method are illustrated in Figure 10.

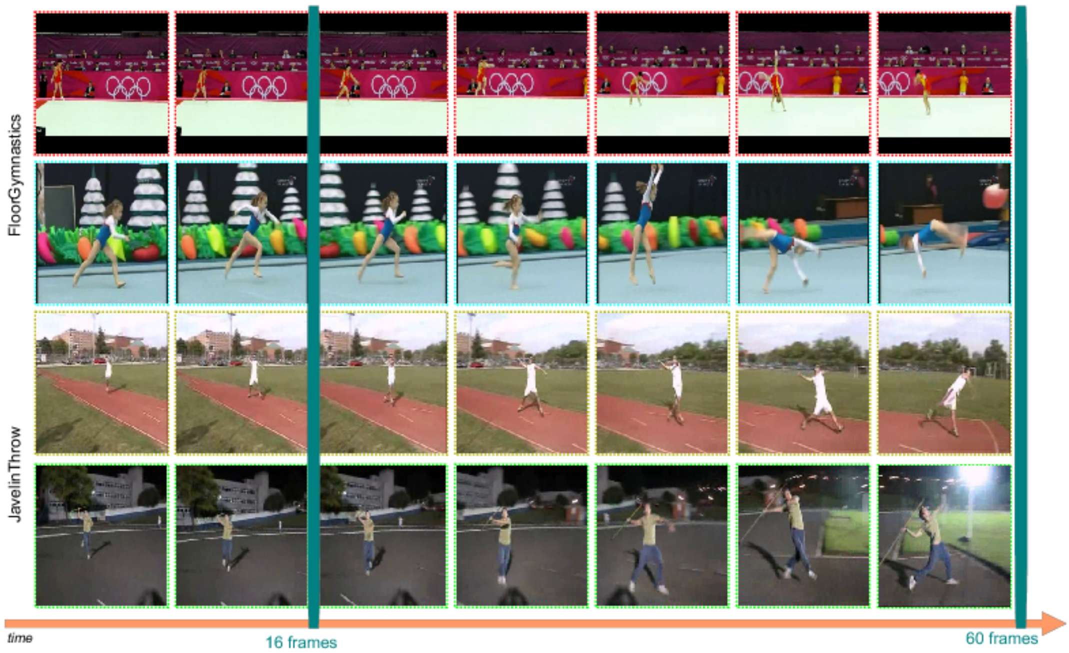

Figure 10. The highest improvement of long-term temporal convolutions in terms of class accuracy is for "JavelinThrow". For 16-frame network, it is mostly confused with "FloorGymnastics" class. We visualize sample videos with 7 frames extracted at every 8 frames. The intuitive explanation is that both classes start by running for a few seconds and then the actual action takes place. Long-term temporal convolutions with 60 frames can capture this interval, whereas 16-frame networks fail to recognize such long-term activities.

7.4.4. Thin-Slicing forPose: Learning to Understand Pose without Explicit Pose Estimation

Participants: Suha Kwak, Minsu Cho, Ivan Laptev.

In [12], we address the problem of learning a pose-aware, compact embedding that projects images with similar human poses to be placed close-by in the embedding space (Figure 11). The embedding function is built on a deep convolutional network, and trained with a triplet-based rank constraint on real image data. This architecture allows us to learn a robust representation that captures differences in human poses by effectively factoring out variations in clothing, background, and imaging conditions in the wild. For a variety of pose-related tasks, the proposed pose embedding provides a cost-efficient and natural alternative to explicit pose estimation, circumventing challenges of localizing body joints. We demonstrate the efficacy of the embedding on pose-based image retrieval and action recognition problems. This work has been published at CVPR 2016 [12].

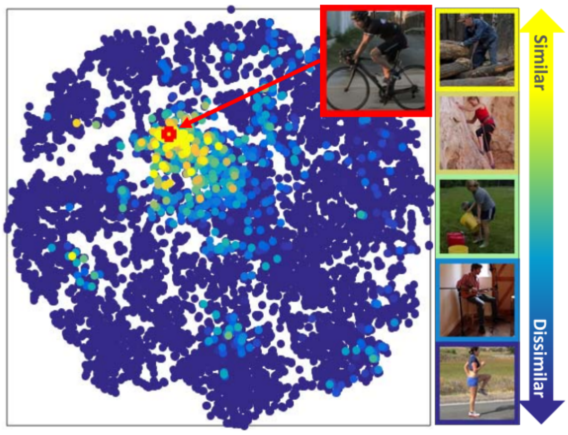

Figure 11. The manifold of our pose embedding visualized using t-SNE. Each point represents a human pose image. To better show correlation between the pose embedding and annotated pose, we color-code pose similarities in annotation between an arbitrary target image (red box) and all the other images. Selected examples of color-coded images are illustrated in the right-hand side. Images similar with the target in annotated pose are colored in yellow, otherwise in blue. As can be seen, yellow images lie closer by the target in general, which indicates that a position on the embedding space implicitly represents a human pose.

7.4.5. Instance-level video segmentation from object tracks

Participants: Guillaume Seguin, Piotr Bojanowski, Rémi Lajugie, Ivan Laptev.

In [14], we address the problem of segmenting multiple object instances in complex videos. Our method does not require manual pixel-level annotation for training, and relies instead on readily-available object detectors or visual object tracking only. Given object bounding boxes at input as shown in Figure 12, we cast video segmentation as a weakly-supervised learning problem. Our proposed objective combines (a) a discriminative clustering term for background segmentation, (b) a spectral clustering one for grouping pixels of same object instances, and (c) linear constraints enabling instance-level segmentation. We propose a convex relaxation of this problem and solve it efficiently using the Frank-Wolfe algorithm. We report results and compare our method to several baselines on a new video dataset for multi-instance person segmentation. This work has been published at CVPR 2016.

8. Bilateral Contracts and Grants with Industry

8.1. Facebook AI Research Paris: Weakly-supervised interpretation of image and video data (Inria)

Figure 12. Results of our method applied to multi-person segmentation in a sample video from our database. Given an input video together with the tracks of object bounding boxes (left), our method finds pixel-wise segmentation for each object instance across video frames (right).

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. Google: Learning to annotate videos from movie scripts (Inria)

Participants: Josef Sivic, Ivan Laptev, Jean Ponce.

The goal of this project is to automatically generate annotations of complex dynamic events in video. We wish to deal with events involving multiple people interacting with each other, objects and the scene, for example people at a party in a house. The goal is to generate structured annotations going beyond simple text tags. Examples include entire text sentences describing the video content as well as bounding boxes or segmentations spatially and temporally localizing the described objects and people in video. This is an extremely challenging task due to large intra-class variation of human actions. We propose to learn joint video and text representations enabling such annotation capabilities from feature length movies with coarsely aligned shooting scripts. Building on our previous work in this area, we aim to develop structured representations of video and associated text enabling to reason both spatially and temporally about scenes, objects and people as well as their interactions. Automatic understanding and interpretation of video content is a key-enabling factor for a range of practical applications such as content-aware advertising or search. Novel video and text representations are needed to enable breakthrough in this area.

8.3. 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.

8.4. MSR-Inria joint lab: Image and video mining for science and humanities (Inria)

Participants: Leon Bottou [Facebook], Ivan Laptev, Maxime Oquab, Jean Ponce, Josef Sivic, Cordelia Schmid [Inria Lear].

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 Science” 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 2013-2016 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.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. Agence Nationale de la Recherche (ANR): SEMAPOLIS

Participants: Mathieu Aubry, Josef Sivic.

The goal of the SEMAPOLIS project is to develop advanced large-scale image analysis and learning techniques to semantize city images and produce semantized 3D reconstructions of urban environments, including proper rendering. Geometric 3D models of existing cities have a wide range of applications, such as navigation in virtual environments and realistic sceneries for video games and movies. A number of players (Google, Microsoft, Apple) have started to produce such data. However, the models feature only plain surfaces, textured from available pictures. This limits their use in urban studies and in the construction industry, excluding in practice applications to diagnosis and simulation. Besides, geometry and texturing are often wrong when there are invisible or discontinuous parts, e.g., with occluding foreground objects such as trees, cars or lampposts, which are pervasive in urban scenes. This project will go beyond the plain geometric models by producing semantized 3D models, i.e., models which are not bare surfaces but which identify architectural elements such as windows, walls, roofs, doors, etc. Semantic information is useful in a larger number of scenarios, including diagnosis and simulation for building renovation projects, accurate shadow impact taking into account actual window location, and more general urban planning and studies such as solar cell deployment. Another line of applications concerns improved virtual cities for navigation, with object-specific rendering, e.g., specular surfaces for windows. Models can also be made more compact, encoding object repetition (e.g., windows) rather than instances and replacing actual textures with more generic ones according to semantics; it allows cheap and fast transmission over low- bandwidth mobile phone networks, and efficient storage in GPS navigation devices.

This is a collaborative effort with LIGM / ENPC (R. Marlet), University of Caen (F. Jurie), Inria Sophia Antipolis (G. Drettakis) and Acute3D (R. Keriven).

9.2. European Initiatives

9.2.1. European Research Council (ERC) Advanced Grant: "VideoWorld" - Jean Ponce

Participants: Jean Ponce, Ivan Laptev, Josef Sivic.

WILLOW will be funded in part from 2011 to 2016 by the ERC Advanced Grant "VideoWorld" awarded to Jean Ponce by the European Research Council.

‘Digital video is everywhere, at home, at work, and on the Internet. Yet, effective technology for organizing, retrieving, improving, and editing its content is nowhere to be found. Models for video content, interpretation and manipulation inherited from still imagery are obsolete, and new ones must be invented. With a new convergence between computer vision, machine learning, and signal processing, the time is right for such an endeavor. Concretely, we will develop novel spatio-temporal models of video content learned from training data and capturing both the local appearance and nonrigid motion of the elements—persons and their surroundings—that make up a dynamic scene. We will also develop formal models of the video interpretation process that leave behind the architectures inherited from the world of still images to capture the complex interactions between these elements, yet can be learned effectively despite the sparse annotations typical of video understanding scenarios. Finally, we will propose a unified model for video restoration and editing that builds on recent advances in sparse coding and dictionary learning, and will allow for unprecedented control of the video stream. This project addresses fundamental research issues, but its results are expected to serve as a basis for groundbreaking technological advances for applications as varied as film post-production, video archival, and smart camera phones.’

9.2.2. European Research Council (ERC) Starting Grant: "Activia" - Ivan Laptev

Participant: Ivan Laptev.

WILLOW will be funded in part from 2013 to 2017 by the ERC Starting Grant "Activia" awarded to Ivan Laptev by the European Research Council.

‘Computer vision is concerned with the automated interpretation of images and video streams. Today’s research is (mostly) aimed at answering queries such as ’Is this a picture of a dog?’, (classification) or sometimes ’Find the dog in this photo’ (detection). While categorisation and detection are useful for many tasks, inferring correct class labels is not the final answer to visual recognition. The categories and locations of objects do not provide direct understanding of their function i.e., how things work, what they can be used for, or how they can act and react. Such an understanding, however, would be highly desirable to answer currently unsolvable queries such as ’Am I in danger?’ or ’What can happen in this scene?’. Solving such queries is the aim of this proposal. My goal is to uncover the functional properties of objects and the purpose of actions by addressing visual recognition from a different and yet unexplored perspective. The main novelty of this proposal is to leverage observations of people, i.e., their actions and interactions to automatically learn the use, the purpose and the function of objects and scenes from visual data. The project is timely as it builds upon the two key recent technological advances: (a) the immense progress in visual recognition of objects, scenes and human actions achieved in the last ten years, as well as (b) the emergence of a massive amount of public image and video data now available to train visual models. ACTIVIA addresses fundamental research issues in automated interpretation of dynamic visual scenes, but its results are expected to serve as a basis for ground-breaking technological advances in practical applications. The recognition of functional properties and intentions as explored in this project will directly support high-impact applications such as detection of abnormal events, which are likely to revolutionise today’s approaches to crime protection, hazard prevention, elderly care, and many others.’

9.2.3. European Research Council (ERC) Starting Grant: "Leap" - Josef Sivic

Participant: Josef Sivic.

The contract has begun on Nov 1st 2014. WILLOW will be funded in part from 2014 to 2018 by the ERC Starting Grant "Leap" awarded to Josef Sivic by the European Research Council.

‘People constantly draw on past visual experiences to anticipate future events and better understand, navigate, and interact with their environment, for example, when seeing an angry dog or a quickly approaching car. Currently there is no artificial system with a similar level of visual analysis and prediction capabilities. LEAP is a first step in that direction, leveraging the emerging collective visual memory formed by the unprecedented amount of visual data available in public archives, on the Internet and from surveillance or personal cameras - a complex evolving net of dynamic scenes, distributed across many different data sources, and equipped with plentiful but noisy and incomplete metadata. The goal of this project is to analyze dynamic patterns in this shared visual experience in order (i) to find and quantify their trends; and (ii) learn to predict future events in dynamic scenes. With ever expanding computational resources and this extraordinary data, the main scientific challenge is now to invent new and powerful models adapted to its scale and its spatio-temporal, distributed and dynamic nature. To address this challenge, we will first design new models that generalize across different data sources, where scenes are captured under vastly different imaging conditions such as camera viewpoint, temporal sampling, illumination or resolution. Next, we will develop a framework for finding, describing and quantifying trends that involve measuring long-term changes in many related scenes. Finally, we will develop a methodology and tools for synthesizing complex future predictions from aligned past visual experiences. Our models will be automatically learnt from large-scale, distributed, and asynchronous visual data, coming from different sources and with different forms of readily-available but noisy and incomplete metadata such as text, speech, geotags, scene depth (stereo sensors), or gaze and body motion (wearable sensors). Breakthrough progress on these problems would have profound implications on our everyday lives as well as science and commerce, with safer cars that anticipate the behavior of pedestrians on streets; tools that help doctors monitor, diagnose and predict patients’ health; and smart glasses that help people react in unfamiliar situations enabled by the advances from this project.’

9.3. International Initiatives

9.3.1. IARPA FINDER Visual geo-localization (Inria)

Participants: Josef Sivic, Petr Gronat, Relja Arandjelovic.

Finder is an IARPA funded project aiming to develop technology to geo-localize images and videos that do not have geolocation tag. It is common today for even consumer-grade cameras to tag the images that they capture with the location of the image on the earth’s surface (“geolocation"). However, some imagery does not have a geolocation tag and it can be important to know the location of the camera, image, or objects in the scene. Finder aims to develop technology to automatically or semi-automatically geo-localize images and video that do not have the geolocation tag using reference data from many sources, including overhead and ground-based images, digital elevation data, existing well-understood image collections, surface geology, geography, and cultural information.

Partners: ObjectVideo, DigitalGlobe, UC Berkeley, CMU, Brown Univ., Cornell Univ., Univ. of Kentucky, GMU, Indiana Univ., and Washington Univ.

9.3.2. Inria CityLab initiative

Participants: Josef Sivic, Jean Ponce, Ivan Laptev, Alexei Efros [UC Berkeley].

Willow participates in the ongoing CityLab@Inria initiative (co-ordinated by V. Issarny), which aims to leverage Inria research results towards developing “smart cities" by enabling radically new ways of living in, regulating, operating and managing cities. The activity of Willow focuses on urban-scale quantitative visual analysis and is pursued in collaboration with A. Efros (UC Berkeley).

Currently, map-based street-level imagery, such as Google Street-view provides a comprehensive visual record of many cities worldwide. Additional visual sensors are likely to be wide-spread in near future: cameras will be built in most manufactured cars and (some) people will continuously capture their daily visual experience using wearable mobile devices such as Google Glass. All this data will provide large-scale, comprehensive and dynamically updated visual record of urban environments.

The goal of this project is to develop automatic data analytic tools for large-scale quantitative analysis of such dynamic visual data. The aim is to provide quantitative answers to questions like: What are the typical architectural elements (e.g., different types of windows or balconies) characterizing a visual style of a city district? What is their geo-spatial distribution (see figure 1)? How does the visual style of a geo-spatial area evolve over time? What are the boundaries between visually coherent areas in a city? Other types of interesting questions concern distribution of people and their activities: How do the number of people and their activities at particular places evolve during a day, over different seasons or years? Are there tourists sightseeing, urban dwellers shopping, elderly walking dogs, or children playing on the street? What are the major causes for bicycle accidents?

Break-through progress on these goals would open-up completely new ways smart cities are visualized, modeled, planned and simulated, taking into account large-scale dynamic visual input from a range of visual sensors (e.g., cameras on cars, visual data from citizens, or static surveillance cameras).

9.4. International Research Visitors

9.4.1. Visits of International Scientists

Prof. Alexei Efros (UC Berkeley, USA) visited Willow during May-June with his postdoc Phillip Isola and Phd student Richard Zhang. Prof. John Canny (UC Berkeley) has visited Willow in 2016 within the framework of Inria's International Chair program.

9.4.1.1. Internships

P. Trutman and O. Rybkin have visited Willow from Czech Technical University in Prague.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- I. Laptev will be program co-chair of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

10.1.1.2. Member of the Organizing Committees

- M. Trager is an organizer of "Minisymposium" on "Algebraic Vision" at the SIAM conference on Applied Algebraic Geometry (Atlanta, July 31st-August 4th 2017).

10.1.2. Scientific Events Selection

10.1.2.1. Area chairs

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016 (I. Laptev).
- Asian Conference on Computer Vision (ACCV), 2016 (I. Laptev).
- International Conference on Computer Vision (ICCV), 2017 (J. Sivic).

10.1.2.2. Member of the Conference Program Committees

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016 (R. Arandjelovic, A. Bursuc, P. Bojanowski, M. Cho, J. Sivic, G. Cheron).
- European Conference on Computer Vision (ECCV), 2016 (R. Arandjelovic, A. Bursuc, P. Bojanowski, G. Cheron, M. Cho, S. Kwak, J. Sivic, G. Cheron, I. Laptev).
- International Conference on Learning Representations, 2016 (J. Sivic).

10.1.3. Journals

10.1.3.1. Member of the editorial board

- 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).
- I. Laptev and J. Sivic co-edit a special issue on "Video representations for visual recognition" in the International Journal of Computer Vision.
- J. Sivic co-edits a special issue on "Advances in Large-Scale Media Geo-Localization" in the International Journal of Computer Vision.

10.1.3.2. Reviewer

- International Journal of Computer Vision (M. Cho, G. Cheron, R. Arandjelovic).
- IEEE Transactions on Pattern Analysis and Machine Intelligence (R. Arandjelovic, P. Bojanowski, M. Cho, S. Kwak, G. Cheron, A. Bursuc).
- IEEE Transactions on Circuits and Systems for Video Technology (P. Bojanowski, B. Ham).
- IEEE Transactions on Image Processing (B. Ham).
- IEEE Signal Processing Letters (B. Ham).
- Computer Vision and Image Understanding (M. Cho, A. Bursuc).
- Elsevier Neurocomputing (B. Ham).
- EURASIP Journal on Image and Video Processing (B. Ham).

10.1.4. Others

- J. Sivic is senior fellow of the Neural Computation and Adaptive Perception program of the Canadian Institute of Advanced Research.
- R. Arandjelovic and J. Sivic obtained the outstanding reviewer award at IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

10.1.5. Invited Talks

- V. Kantorov, Speaker at Second Christmas Colloquium on Computer Vision (December 2016, Moscow, SkolTech).
- I. Laptev, Invited talk, MailRu, Moscow, May 2016.
- I. Laptev, Invited talk, Skolkovo Robotics, Moscow, May 2016.
- I. Laptev, Invited talk, Deep Machine Intelligence and its Applications, SkolTech, Moscow, June 2016.
- I. Laptev, Invited talk, MSR-Inria Offsite Meeting, Paris, September, 2016.
- I. Laptev, Invited talk, University of Central Florida, Orlando, September, 2016.
- I. Laptev, Invited talk, Georgia Institute of Technology, Atlanta, September, 2016.
- I. Laptev, Invited talk, ECCV'16 Workshop on Brave new ideas for motion representations in videos, Amsterdam, October, 2016.
- I. Laptev, Invited talk, Open Day AI Innovation Factory , December, 2016.
- J. Ponce, Invited talk, New York University, January 2016.
- J. Ponce, Invited talk, Université Marne la Vallée, Mars 2016.
- J. Ponce, Invited talk, Workshop on Algebraic Vision, San Jose, May 2016.
- J. Ponce, Invited talk, Colloque LORIA, Nancy, May 2016.
- J. Ponce, Invited talk, Parthenos Workshop, Bordeaux, November 2016.

- J. Sivic, seminar, UC Berkeley, May, December, 2016.
- J. Sivic, invited talk, Brno University of Technology, April 2016.
- J. Sivic, Invited talk, the CIFAR workshop, Barcelona, December 2016.
- J. Sivic, Invited talk, Colloquium on Perspectives and New Challenges in Data Science, Ecole de Ponts ParisTech, 2016.
- M. Trager, invited speaker, AIM workshop “Algebraic Vision" (San Jose, May 2-6, 2016).

10.1.6. Leadership within the Scientific Community

- Member, advisory board, IBM Watson AI Xprize (J. Ponce).
- Member, steering committee, "France Intelligence Artificielle" initiative (J. Ponce).
- Member, advisory board, Computer Vision Foundation (J. Sivic).

10.1.7. Scientific Expertise

- J. Sivic gave an overview of state-of-the-art in computer vision at the seminar on artificial intelligence, Direction Generale des Entreprises (DGE) du Ministere de l’Economie, de l’Industrie et du Numerique, September, 2016.

10.1.8. Research Administration

- Member, Bureau du comité des projets, Inria, Paris (J. Ponce)
- Director, Department of Computer Science, Ecole normale supérieure (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 developpement technologique (CDT), 2012- (J. Sivic).
- Member of ANR evaluation committee (I. Laptev).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Master : J. Ponce, "Introduction to computer vision", M1, Ecole normale superieure, 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 : I. Laptev, J. Ponce and J. Sivic, Cours PSL-ITI - Informatique, mathematiques appliques pour le traitement du signal et l’imagerie, 20h.

10.2.2. Supervision

PhD : Piotr Bojanowski, "Learning to annotate dynamic video scenes", graduated in 2016, I. Laptev, J. Ponce, C. Schmid and J. Sivic.

PhD : Guillaume Seguin, "Person analysis in stereoscopic movies", graduated in 2016, 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 Phd thesis proposal", started in Oct 2016, I. Laptev, J. Sivic, P. Bojanowski (Facebook AI Research).

PhD in progress : Gul Varol, “Deep learning methods for video interpretation”, started in Oct 2015, I. Laptev, C. Schmid.

PhD in progress : Julia Peyre, “Learning to reason about scenes from images and language", started in Oct 2015, C. Schmid, I. Laptev, J. Sivic.

PhD in progress : Jean-Baptiste Alayrac, "Structured learning from video and natural language", started in 2014, I. Laptev, J. Sivic and S. Lacoste-Julien (Inria SIERRA / U. Montreal).

PhD in progress : Rafael Sampaio de Rezende, started in 2013, J.Ponce.

PhD in progress : Guilhem Cheron, "Structured modeling and recognition of human actions in video", started in 2014, I. Laptev and C. Schmid.

PhD in progress : Theophile Dalens, "Learning to analyze and reconstruct architectural scenes", starting in Jan 2015, M. Aubry and J. Sivic.

PhD in progress : Vadim Kantorov, "Large-scale video mining and recognition", started in 2012, I. Laptev.

PhD in progress : Maxime Oquab, "Learning to annotate dynamic scenes with convolutional neural networks", started in Jan 2014, L. Bottou (Facebook AI Research), I. Laptev and J. Sivic.

PhD in progress : Matthew Trager, "Projective geometric models in vision", started in 2014, J. Ponce and M. Hebert (CMU).

PhD in progress : Tuang Hung VU, "Learning functional description of dynamic scenes", started in 2013, I. Laptev.

10.2.3. Juries

- PhD thesis committee:
  - Stavros Tsogkas, Ecole Centrale, France, 2016 (J. Sivic, examinateur).
  - Sesh Karri, Ecole Normale Superieure, France, 2016 (J. Sivic, examinateur).
  - Elliot Crowley, University of Oxford, UK, 2016, (J. Sivic, external examiner)
  - Olivier Frigo, Universite Paris Descartes, France, 2016 (J. Sivic, rapporteur).
  - Mattis Paulin, Inria Grenoble, France, 2017 (J. Sivic, rapporteur).
  - Francesco Massa, ENPC, France 2017 (J. Sivic, examinateur).
  - Philippe Weinzaepfel, Universite Grenoble Alpes, France, 2015 (I. Laptev, rapporteur).
  - Guillaume Seguin, Ecole Normale Superieure, France, 2016 (I. Laptev, J.Ponce, J. Sivic, examinateurs).
  - Piotr Bojanowski, Ecole Normale Superieure, France, 2016 (I. Laptev, J.Ponce, J. Sivic, examinateurs).
  - Ala Aboudib, Télécom Bretagne, France, 2016 (J. Ponce).
  - Philippe Weinzaepfel, Universite Grenoble Alpes, France, 2016 (J. Ponce).

10.3. Popularization

- Participation to the round table on "L'IA est-elle réservée aux GAFA", NUMA, June 2016 (J. Ponce).
- Participation to the round table on "Fictions, magie numerique et realites", Post-digital program, ENS/PSL Research University, October 2016 (J. Ponce).
- Debate with Jacques Attali, "Intelligence Artificielle, science avec conscience?", "Intelligence Artificielle : de la technique au business" Conference, December 2016 (J. Ponce).
- Participation to the round table on AI, Liberté Living Lab, December 2016 (J. Ponce).
- Participation to the round table on ethics at the Senate's public hearing on Artificial Intelligence, January 2017 (J. Ponce).
- Interview on Nolife 56Kast (https://www.youtube.com/watch?v=8UgH8_J2ugU) (J. Ponce).
- Interview in Le Monde (http://www.lemonde.fr/pixels/article/2016/01/08/intelligence-artificielle-ce-que-voient-les-machines_4843858_4408996.html) (J. Ponce).

- Interview in Télérama (http://www.telerama.fr/monde/trouver-le-calme-reconstituer-palmyre-ou-choisir-un-traitement-grace-a-l-ia,141131.php) (J. Ponce).

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] P. BOJANOWSKI.*Learning to annotate dynamic video scenes*, Ecole normale supérieure, June 2016, https://hal.inria.fr/tel-01364560.

[2] G. SEGUIN.*Person analysis in stereoscopic movies*, Ecole normale superieure, April 2016, https://tel.archives-ouvertes.fr/tel-01311143.

Articles in International Peer-Reviewed Journal

[3] P. GRONÁT, G. OBOZINSKI, J. SIVIC, T. PAJDLA.*Learning and calibrating per-location classifiers for visual place recognition*, in "International Journal of Computer Vision", April 2016, vol. 118, n^{o} 3, p. 319-336 [*DOI :* 10.1007/S11263-015-0878-X], https://hal.inria.fr/hal-01418239.

[4] B. HAM, M. CHO, J. PONCE.*Robust Guided Image Filtering Using Nonconvex Potentials*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2017, Accepted pending minor revision, https://hal.archives-ouvertes.fr/hal-01279857.

[5] H. R. IDREES, A. R. ZAMIR, Y.-G. JIANG, A. R. GORBAN, I. R. LAPTEV, R. R. SUKTHANKAR, M. R. SHAH.*The THUMOS challenge on action recognition for videos "in the wild"*, in "Computer Vision and Image Understanding", 2016 [*DOI :* 10.1016/J.CVIU.2016.10.018], https://hal.inria.fr/hal-01431525.

[6] J. SUN, J. PONCE.*Learning Dictionary of Discriminative Part Detectors for Image Categorization and Cosegmentation*, in "International Journal of Computer Vision", March 2016 [*DOI :* 10.1007/S11263-016-0899-0], https://hal.archives-ouvertes.fr/hal-01064637.

[7] M. TRAGER, J. PONCE, M. HEBERT.*Trinocular Geometry Revisited*, in "International Journal on Computer Vision (IJCV)", 2016, https://hal.archives-ouvertes.fr/hal-01152348.

International Conferences with Proceedings

[8] J.-B. ALAYRAC, P. BOJANOWSKI, N. AGRAWAL, J. SIVIC, I. LAPTEV, S. LACOSTE-JULIEN.*Unsupervised Learning from Narrated Instruction Videos*, in "CVPR2016 - 29th IEEE Conference on Computer Vision and Pattern Recognition", Las Vegas, United States, June 2016, https://hal.inria.fr/hal-01171193.

[9] R. ARANDJELOVIĆ, P. GRONAT, A. TORII, T. PAJDLA, J. SIVIC.*NetVLAD: CNN architecture for weakly supervised place recognition*, in "CVPR 2016 - 29th IEEE Conference on Computer Vision and Pattern Recognition", Las Vegas, United States, June 2016, https://hal.inria.fr/hal-01242052.

[10] B. HAM, M. CHO, C. SCHMID, J. PONCE.*Proposal Flow*, in "CVPR 2016 - IEEE Conference on Computer Vision & Pattern Recognition", LAS VEGAS, United States, June 2016, https://hal.archives-ouvertes.fr/hal-01240281.

[11] V. KANTOROV, M. OQUAB, M. CHO, I. LAPTEV.*ContextLocNet: Context-Aware Deep Network Models for Weakly Supervised Localization*, in "ECCV 2016", Amsterdam, Netherlands, Springer, October 2016, p. 350 - 365 [*DOI :* 10.1007/978-3-319-46454-1_22], https://hal.inria.fr/hal-01421772.

[12] S. KWAK, M. CHO, I. LAPTEV.*Thin-Slicing for Pose: Learning to Understand Pose without Explicit Pose Estimation*, in "CVPR 2016 - IEEE Conference on Computer Vision and Pattern Recognition", Las Vegas, United States, June 2016, https://hal.inria.fr/hal-01242724.

[13] R. LAJUGIE, P. BOJANOWSKI, P. CUVILLIER, S. ARLOT, F. BACH.*A weakly-supervised discriminative model for audio-to-score alignment*, in "41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP)", Shanghai, China, Proceedings of the 41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP), March 2016, https://hal.archives-ouvertes.fr/hal-01251018.

[14] G. SEGUIN, P. BOJANOWSKI, R. LAJUGIE, I. LAPTEV.*Instance-level video segmentation from object tracks*, in "CVPR 2016", Las Vegas, United States, Proceedings of the 29th IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, June 2016, https://hal.inria.fr/hal-01255765.

[15] G. A. SIGURDSSON, G. VAROL, X. WANG, A. FARHADI, I. LAPTEV, A. GUPTA.*Hollywood in Homes: Crowdsourcing Data Collection for Activity Understanding*, in "Computer Vision – ECCV 2016", Amsterdam, Netherlands, October 2016, p. 510 - 526 [*DOI :* 10.1007/978-3-319-46448-0_31], https://hal.inria.fr/hal-01418216.

[16] M. TRAGER, M. HEBERT, J. PONCE.*Consistency of silhouettes and their duals*, in "IEEE Conference on Computer Vision and Pattern Recognition, 2016", Las Vegas, United States, June 2016, https://hal.archives-ouvertes.fr/hal-01287180.

[17] Y. ZHONG, R. ARANDJELOVIĆ, A. ZISSERMAN.*Faces In Places: Compound query retrieval*, in "BMVC - 27th British Machine Vision Conference", York, United Kingdom, September 2016, https://hal.inria.fr/hal-01353886.

Other Publications

[18] A. BABENKO, R. ARANDJELOVIĆ, V. LEMPITSKY.*Pairwise Quantization*, June 2016, working paper or preprint, https://hal.inria.fr/hal-01330582.

[19] J. PONCE, B. STURMFELS, M. TRAGER.*Congruences and Concurrent Lines in Multi-View Geometry*, 2017, Accepted for "Advances in Applied Mathematics", https://hal.inria.fr/hal-01423057.

[20] G. VAROL, I. LAPTEV, C. SCHMID.*Long-term Temporal Convolutions for Action Recognition*, April 2016, working paper or preprint, https://hal.inria.fr/hal-01241518.