RESEARCH CENTER
Paris

FIELD

Activity Report 2016

Section New Results

Edition: 2017-08-25

ALPAGE Project-Team

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.

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.

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

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,

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

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 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.

ALPINES Project-Team

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.

ANGE Project-Team

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

ANTIQUE Project-Team

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 $T\infty(f)† = T1(f†)$.

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.

AOSTE Project-Team

6. New Results

6.1. CCSL as a Logical Clock Calculus Algebra: expressiveness and analysis techniques

Participants: Robert de Simone, Julien Deantoni, Frédéric Mallet, Dongdong An.

CCSL is a simple, half-declarative and half-imperative language describing relations and constraints between sequences of events considered as Logical Clocks. The usage of CCSL for specification of embedded systems is powerful in that it defers the precise setting of physical timing until later implementation design phases (which may vary according to circumstances), see 3.2 .

Early this year we established the universal recursive expressivity of CCSL, by encoding the dynamics of Petri Nets with inhibitor arcs in our framework (still unpublished). Those results were presented by Robert de Simone in a keynote talk at Memocode 2016. This result prompts the use of non-automatic methods for establishing actual schedules as solutions of CCSL specifications seen as schedulability constraints. Steps in that direction were made in [37].

We also considered the extension of CCSL towards stochastic modeling of potential input clocks as my emerge from the Cyber-Physical world (mixing probabilistic modeling of external events with discrete transformations by discrete cyber digital controlers). This work was initiated in [28], and should be further extended in the ongoing PhD thesis of Dongdong An.

Finally, we have also investigated to decide on specific schedules (e.g. periodic schedules) valid for a subset of CCSL. We have established a sufficient static condition for the existence of such a periodic schedule as well as a practical implementation to build such a solution [39] based on a SMT solver.

6.2. Industrial design flow for Embedded System Engineering

Participants: Julien Deantoni, Frédéric Mallet, Marie Agnes Peraldi Frati, Robert de Simone, Hui Zhao, Ales Mishchenko.

As part of the PIA LEOC Clarity collaborative project we considered the inytroduction of formal methods into a high-level model-based design environment for embedded systems, named CAPELLA (https://polarsys.org/capella/). CAPELLA is part of the Polarsys Eclipse project. It originates from Thales, and is currently being deployed in real operational divisions in a number of companies.

Our activities consisted in demonstrating how the theoretical models of Logical Time and derives Models of Computation could be used to give precise semantics and provide simulation benefits, when applied to the modeling paradigms used in CAPELLA and advanced in Clarity. In particular we focused on the connection between timing/performance properties and other kinds of non-functional properties, including model variability.

This year we focused on two mains tasks:

First, we clarified and extended the notion of Modes and States in the Capella system engineering language. Specifically, a specific diagram has been introduced to deal with the system modes. The notion of mode is then used to specify different configurations of the system, mainly in terms of the active functions, their data dependencies, their deployment on the logical and physical architecture as well as the scenario to be verified in this specific mode. In consequence, the behavioural semantics of the mode diagram strongly interacts with the behavioral semantics of the other diagrams. The execution semantics was given by promoting our contributions in GEMOC and BCOoL (see 6.3).

Second, Capella proposes a consistent multi-view approach accross different engineering domains. At some step in the refinement process, these different views are extracted to a domain specific tool (like Simulink for instance). It is then required 1) to verify that the manipulation done in the domaine specific tool respect the original semantics expected by the architect, and 2) to understand the impact of the decisions made in domain specific tools on the interaction with the other views. To do so we provided a generic aproach to confront the race to the behavioral semantics we formally defined in Capella. We are curretly working on a theoretical approach to improve the overal performance of such approach.

While BCOoL and Gemoc only considers discrete models, the PhD thesis of Hui Zhao, which started in March 2016, explores a possible extension that specifically targets Cyber-Physical Systems where we different timed models combined, including both discrete and dense timed models. In this thesis, we also explore the impact of such an heterogeneous modeling framework to guarantee security and safety properties of the combined models. This is done in collaboration with Ludovic Apvrille (who is co-advisor of the thesis) from Telecom ParisTech.

6.3. Coordination of heterogeneous Models of Computation as Domain-Specific Languages

Participants: Matias Vara Larsen, Julien Deantoni, Frédéric Mallet.

Our work this on coordination of heterogeneous languages produced two major results. The first one is the development of BCOoL (Behavioral Coordination Operator Language. BCOoL is a language dedicated to the specification of coordination patterns between heterogeneous languages. It comes with a tool chain allowing the generation of the coordination given a BCOoL operator and specific models. Our second result is the development of an heterogeneous execution engine, integrated to Gemoc studio, to run conjointly different models. Both works re extensively reported in Matias Vara Larsen PhD thesis [19].

6.4. SoC multiview (meta)modeling for performance, power, and thermal aspects

Participants: Amani Khecharem, Robert de Simone, Emilien Kofman, Julien Deantoni.

In the framework of the ANR HOPE project we progressed the definition of multiview metamodels for the design of Systems-on-Chip) (SoC systems integrating performance, power and thermal aspects. The main concern was to stress regularity and commonalitbetween those views, each developed on "domains" defined as partitions of the original block diagram (clock domains, voltage domains, floorplans,...), and with finite state machine controlers setting the levels of these domains; links between distinct views are originally provided by laws of physics, but then usually identified with discrete allowed values(such as OPP, Operating Performance Points, providing the available frequency-voltage levels for processor clocks).

The corresponding methodology, named MuArch, was reported as Ameni Khacharem PhD document [16].

6.5. MoCs and novel architectures

Participants: Amine Oueslati, Robert de Simone, Albert Savary, Emilien Kofman.

In the context of the FUI Clistine project we considered the links betwen formal Models of Computation and parallel programming models (MPI mainly). The objective is to figure to what level an abstraction of MPI processes as concuurent communicating processes can help for the AAA design process being applied to theselection of adequate MPI communications. This topic reflects the ongoing PhD thesis of Amine Oueslati, and the engineering work of Albert Savary in the first semester.

6.6. Solving AAA constraints analytically

Participants: Emilien Kofman, Dumitru Potop Butucaru, Robert de Simone, Amine Oueslati.

We experimented on the use of SMT solvers to compute efficient mappings (bothe schedules and placement allocations) for concurrent embedded applications onto specific embedded architectures of big.LITTLE features (where alocation and migration of tasks can follow concern for low-power consumption). In fact, the work consisted greatly in a study of how the various models could be encoded to scale up, allowing the solvesr to provide results in reasonalble time. The results have been presented [41], [31], and will soon appear as E. Kofman PhD thesis.

6.7. Coupling SystemC and FMI for co-simulation of Cyber-Physical Systems

Participants: Stefano Centomo, Julien Deantoni, Robert de Simone.

In collaboration with Professor Davide Quaglia, from the University of Verona, we are studying the proper joint modeling of interactions between different domains involved in a cyber-physical system (CPS), and specifically between the cyber and physical parts. In our first work, realized in the context of Stefano Centomo master internship, we investigated how an event based hardware description language can be used in an emerging industry standard for co-simulation (FMI/FMU developed orginally in a Modelica framework). Preliminary results were published [26], and we hope to start a PhD as follow-up of these results.

6.8. Behavioural Semantics if Open pNets

Participants: Eric Madelaine, Ludovic Henrio, Siqi Li, Min Zhang.

We have extended our preliminary work on Parameterised Networks of Automata (pNets), by looking at the behavioural semantics and at bisimulation equivalences for open pNet systems. These can be used to encode operators of various process algebras, construct of distributed or reactive system programming languages, or even parallel algorithmic skeletons, and generic distributed algorithms. As a first step, we studied the properties of a strong bisimulation equivalence based on logical hypotheses about the behaviour of process variables in the open systems. This has been published in [22], [33] and an extended version as an Inria research report [43]. We are now implementing algorithms for computing the symbolic behavioural semantics of open pNets, and checking strong bisimulation, using a SAT engine for reasoning on the hypotheses.

In order to understand better this behavioural semantics, we also have defined another version with a denotational flavour, namely using a "Universal Theory of Processes (UTP)" style. There we express the communication actions of pNets using traces of interaction events, and we were able to prove axiomatic properties of some simple (open) pNets. This was published in [32]. In the long term, it could be interesting to study the relations between the FH-bisimulation and the UTP semantics, relating both behavioural, denotational and algebraic semantics of pNets.

6.9. Behavioural semantics for GCM components

Participants: Ludovic Henrio, Oleksandra Kulankhina, Eric Madelaine.

With Ludovic Henrio (Comred/I3S) and Rabea Ameur-Boulifa (Labsoc/Telecom-Paristech), we have pursued our research on the Behavioural semantics, in terms of pNets, of the core concepts of Grid Component Model (GCM). The results are currently submitted for publication as a journal paper, under revision.

6.10. Performance analysis and optimisation of an HPC scientific application

Participants: Luis Agustin Nieto, Sid Touati.

In the context of the internatinal Internship of Luis Agustin Nieto we conducted a large-scale experiment of source code optimization for HPC application. This work is meant to identify potential approaches that may be automatized in the future. The current use case was an application named CONVIV. CONVIV is a computer code implementing the VMFCI Method to solve the stationary Schrödinger equation for a set of distinguishable degrees of freedom (https://svn.oca.eu/trac/conviv). It is used in Chemistry for computing the energy levels of molecules.

This application is very computer-intensive (many hours of computation on a high performance grid computer). We have been given its source code (fortran with OpenMP), and we have been asked to analyse its performance and to optimise its execution time.

We did an extensive set of experiments for this application on many computers, and mainly on the `cicada.unice.fr` shared grid computer used for scientific parallel computing at UNS). We varied many parameters in our experiments:

- The number of threads was 2, 4, 6, 8, 16 threads. We also analysed the sequential code version.
- The thread affinity strategies for scheduling were: none (linux scheduler), scatter, compact.
- We repeated each experience 35 times to analyse performance stability.
- We used 2 compilers (gfortran, ifort) with -O3.
- We did a precise performance profiling using the Intel Vtune tool.

During our experiments we observed that, even with all the parameters above kept fixed, repeating the executions 35 times shows grat variability between best and worst execution times (more than double in some cases). The critical-path functions remained the same for each configuration choice, including in particular specific matrix computation functions.

After investigation and experiments, we succeed in getting a spectacular performance improvement by applying the following optimisations:

- Replace one of the matrix computation function by an MKL one (highly optimised and tuned function done by Intel).
- Use the compact thread scheduling strategy (OpenMP parameter).
- By using gfortran compiler with -O3, we reduced the execution time from 18400 seconds to 820 seconds (speedup=22).
- By using the ifort compiler with -03, we reduced the execution time from 21000 seconds to 620 seconds (speedup=33).

6.11. Formal translation validation of multi-processor real-time schedules

Participants: Keryan Didier, Dumitru Potop-Butucaru.

This research direction is mainly represented by the PhD thesis of Keryan Didier, and takes place in the framework of the ITEA3 ASSUME project. The technical focus of the ASSUME project is on formal compiler verification and on correct real-time implementation for parallel applications. The objective of this PhD thesis is to formally prove the correctness of (part of) the automatic code generation technology of Lopht, considering the respect of non-functional requirements, and in particular real-time requirements such as release dates, deadlines and periods.

During this first year of work we have:

1. Simplified the allocation and scheduling algorithms of Lopht to facilitate proof while still being able to handle the industrial use case. The resulting algorithms consider all the aspects pertaining to functional specification and non-functional requirements, but make simplifying assumptions on the execution platform (by not taking into account memory access interferences during parallel execution).
2. Developed a formally proved translation validation tool to determine the correctness of schedules produced by the algorithms at point (1). The tool is developed and proved in Coq. Coq code extraction is used to produce OCaML code that integrates in the allocation and scheduling flow.
3. Evaluated the tool on a large-scale industrial use case from Airbus (6000 Scade nodes). We demonstrated the tool to our project partners and during the ASSUME project evaluation. This evaluation showed that our scheduling and formally proved validation tools scale up to the size of large applications.

The main limitation of the current work is that it does not take into account the interferences due to concurrent memory accesses. This gives the main research direction for the next year.

We are currently writing a paper on this subject.

6.12. Lopht back-end for TTEthernet-based distributed systems

Participants: Raul Gorcitz, Dumitru Potop-Butucaru.

The global objective of this activity is a large-scale, ongoing effort to assess the possibility of automatically synthesizing full real-time implementations, including the so-called "bus frame" (the network configuration) on complex industrial platforms and for complex functional and non-functional specifications. We worked this year in the context of the post-doctoral position of Raul Gorcitz, funded by the ITEA3 ASSUME project, but also in the framework of our collaboration with CNES and Airbus DS.

The chosen platform was an industry-level evaluation platform using several Single-Board Computers (SBCs) running the VxWorks 653 OS, and connected through a Time-Triggered Ethernet (TTE) network. This platform was provided by CNES, as typical target for embedded applications. TTE is a standardized commmercial communication network, on top of a switched Ethernet basis, ommmercialized by TTTech. TTE adds support for realtime and fault tolerant communications, allows multiple communications of mixed criticalities to share a single physical medium. This is ensured by means of dedicated hardware using a set of configuration files describing the system architecture and behavior. These configurations are synthesized by the proprietary TTEplan tool starting from a global network description file.

The main scientific difficulty was the formal modeling of the behavior of the TTE network, followed by the extension of scheduling algorithms to consider such a network. While preliminary results were obtained and published last year, we completed and demonstrated this work to our industrial partners, and we are currently writing a second paper on the subject.

6.13. Uniprocessor Real-Time Scheduling

Participants: Mehdi Mezouak, Yves Sorel, Walid Talaboulma.

In the context of the master internship of Mehdi Mezouak, we thoroughly tested the offline time triggered scheduler implemented on an ARM Cortex M4 last year. We remind that this scheduler, intended for safety critical applications, uses a scheduling table containing the instants when the scheduler will be called through interruptions triggered by a timer. This table is generated by a uniprocessor offline schedulability analysis which accounts accurately for the scheduler cost itself, and for the cost of all preemptions the data dependent tasks are subjected to. This approach allows accounting for preemptions induced by the cost of other preemptions. We implemented a time measurement system on a LPC4080 microcontroller board of NXP which includes the ARM Cortex M4 and several timers, to determine on the one hand the actual cost of the scheduler and the cost of one preemption, and on the other hand start, resume and completion times of every task of the task sets. For the ARM Cortex M4 with a 120Mhz clock we obtained 142 cycles (2.3 μs) for the scheduler cost and 54 cycles (0.9 μs) for the cost of one preemption. We used these values for schedulability analyses we applied to various task sets. We improved the graphical tools proposed last year to draw the timing diagrams obtained during the schedulability analysis and during the real-time execution of the task set in order to compare them. For example, thanks to these measurement system and tools, we showed that this scheduler, based on a non periodic timer rather than the usual periodic one, allows the periodic execution of tasks without any jitter.

6.14. Multiprocessor Real-Time Scheduling

Participants: Mehdi Mezouak, Salah Eddine Saidi, Yves Sorel.

Always in the context of the master internship of Mehdi Mezouak, we studied the extension to multiprocessor of our offline time triggered scheduler. Since we chose the partitioned multiprocessor scheduling approach rather than the global one which is not suited to safety critical applications due to the prohibitive cost of task migrations, the uniprocessor schedulability analysis is easily extended. Indeed, the main modification consists, for every processor, in accounting for the cost of inter-processor communications and synchronizations due to data dependences when a producer task is allocated to a processor which is different from the one the corresponding consumer task is allocated to. Therefore, new scheduler calls are added to the scheduling table corresponding to instants when awaited data are available, i.e. produced and then transfered. Of course, there are as many scheduling tables, and thus schedulers, as there are processors, and these scheduling tables are supposed to share a unique global time. The implementation of this global time raises a complex problem since it is not possible to dispatch a unique physical clock to all the processors. Among various solutions, we chose to use a physical clock rather than a logical one like in the Lamport's timestamp approach since we are interested in safety critical real-time. In addition, we chose the Berkeley's algorithm based on a master-slave approach where the clock server is maintained by one of the processor of the multiprocessor. This algorithm is more robust to failures than other algorithms based on an external clock server. Finally, using the measurement system mentioned previously, we measured accurately the cost of inter-processor communications according to the number of transfered data, in the case of an ethernet network that we experimented last year to connect several LPC4080 microcontroller boards.
During the second year of the PhD thesis of Salah Eddine Saidi, we continued to study the parallelization on multi-core of FMI-based co-simulation of numerical models, that is increasingly used for the design of Cyber-Physical Systems. Such model developed according to the FMI standard is defined by a number of C functions, called "operations", for computing its variables (inputs, outputs, state) and data dependences between these variables. Each model has an associated integration step and exchanges data with the other models according to its communication step which can be larger or equal to its integration step. These models are represented by a dataflow graph of operations [35] that is compliant with the conditioned repetitive dataflow model of our AAA methodology for functional specification. Our work mainly focused on two aspects. First, we proposed a graph transformation algorithm in order to allow handling multi-rate co-simulation, i.e. where connected models have different communication steps. This algorithm is based on the concept of graph unfolding similarly to the unrolling algorithm of our AAA methodology. The new graph is represented over the hyper-step which is equal to the least common multiple of the communication steps of all the models. Each operation is repeated in the graph according to the ratio between the hyper-step and its communication step. Then, rather than adding edges connecting all the repetitions of dependent operations, specific rules are used to define the repetitions that have to be connected by edges. These rules ensure correct data exchange between the operations as requested in the context of simulation. Second, some FMI functions called to compute model variables may not be "thread-safe", i.e. they cannot be executed in parallel as they may share some resource (e.g. variables). Consequently, if two or more operations belonging to the same model are executed on different cores, a mechanism that ensures these operations are executed in strictly disjoint time intervals must be set up. We proposed an acyclic orientation heuristic to solve this problem. This heuristic adds non directed edges between the operations that belong to the same model, and then assigns directions to these edges with the aim of minimizing the critical path of the resulting graph and subject to the constraint that no cycle is generated in the graph.

6.15. Probabilistic Solutions for Hard Real-Time Systems

Participants: Adriana Gogonel, Dorin Maxim, Antoine Bertout, Tomasz Kloda, Irina Asavoae, Mihail Asavoae, Cristian Maxim, Walid Talaboulma, Slim Ben-Amor, Robert Davis, Liliana Cucu.

The probabilistic solutions for hard real-time systems are built under the hypothesis that worst case values and worst case execution scenarios have extremely low probability of appearance. While continuing the estimation of bounds for the worst case execution times of a program [34], [25], we have proposed the first utilisation of probabilistic description for mixed-criticality systems [42]. Our result is exploiting the heavy tails of the execution times of a program to propose efficient scheduling solutions. Moreover since the feasibility intervals [21] for a probabilistic real-time system is not formally identified, we have formulated the first feasibility reasoning for such systems [47] under fixed-priority assignment policies [20]. Another important problem

for probabilistic real-time systems concerns the feasibility in presence of precedence constraints, often used by our industry partners. The introduction of precedence constraints requires the comparison of probabilistic arrivals and we showed that existing measures are not correct in this context and we proposed and proved correct new measures [24].

ARAMIS Project-Team

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].

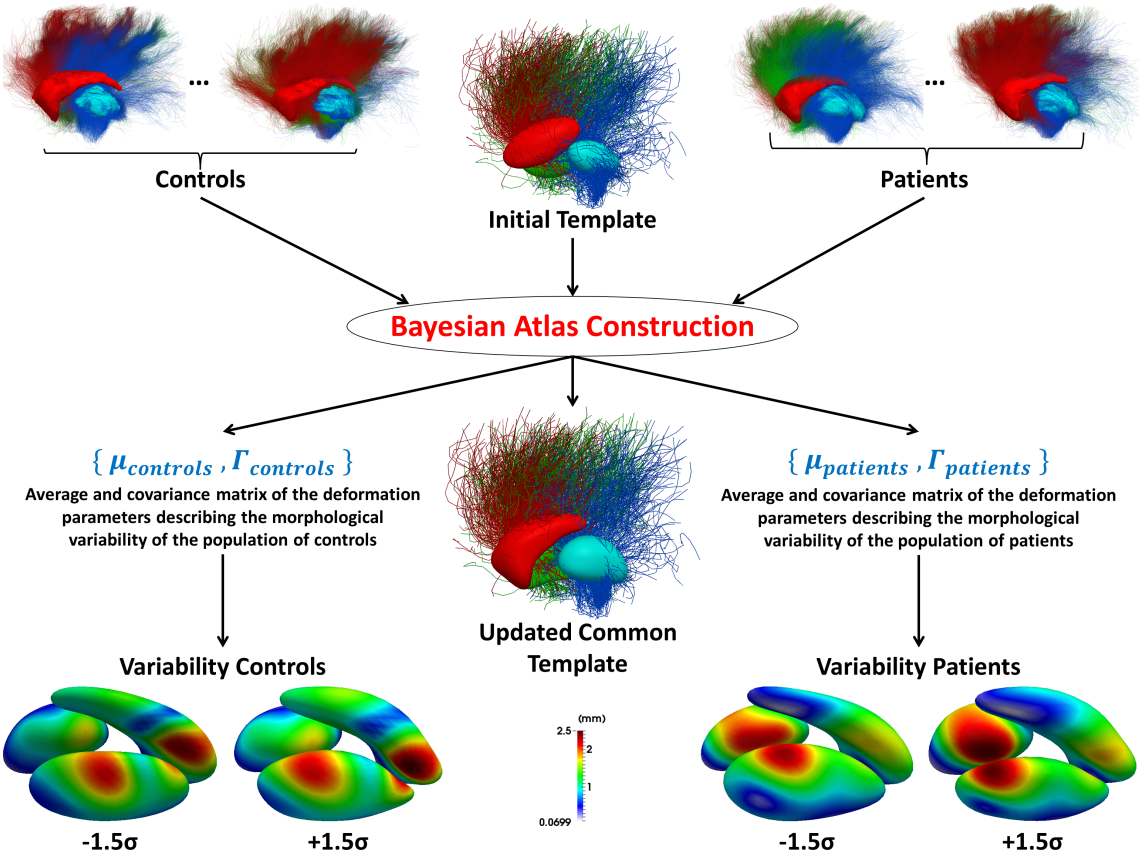

Figure 1. Bayesian framework for atlas construction of multi-object shape complexes comprised of both surface and curve meshes.

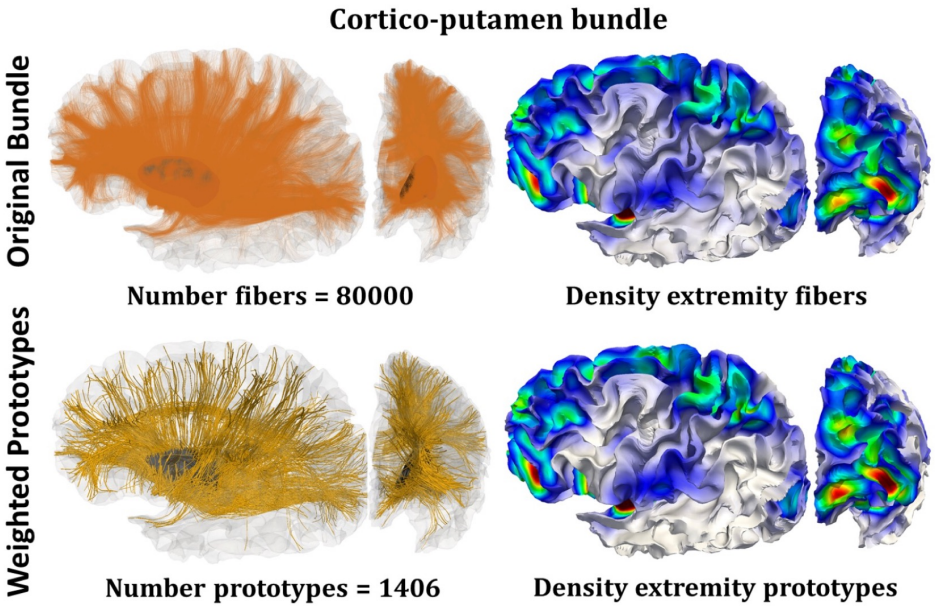

Figure 2. Weighted prototype approximations of a fiber bundle. As it is possible to notice, our approximation alters neither the global shape of the bundle nor the densities of the endpoints onto the cortical surface.

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.

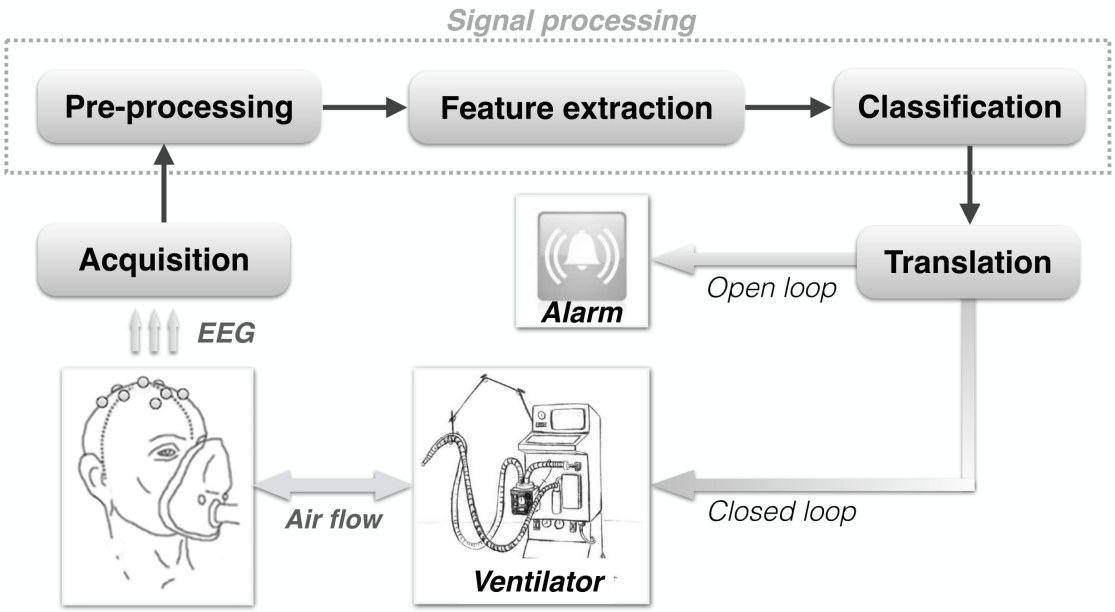

Figure 3. Scheme of a brain-ventilator interface.

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,

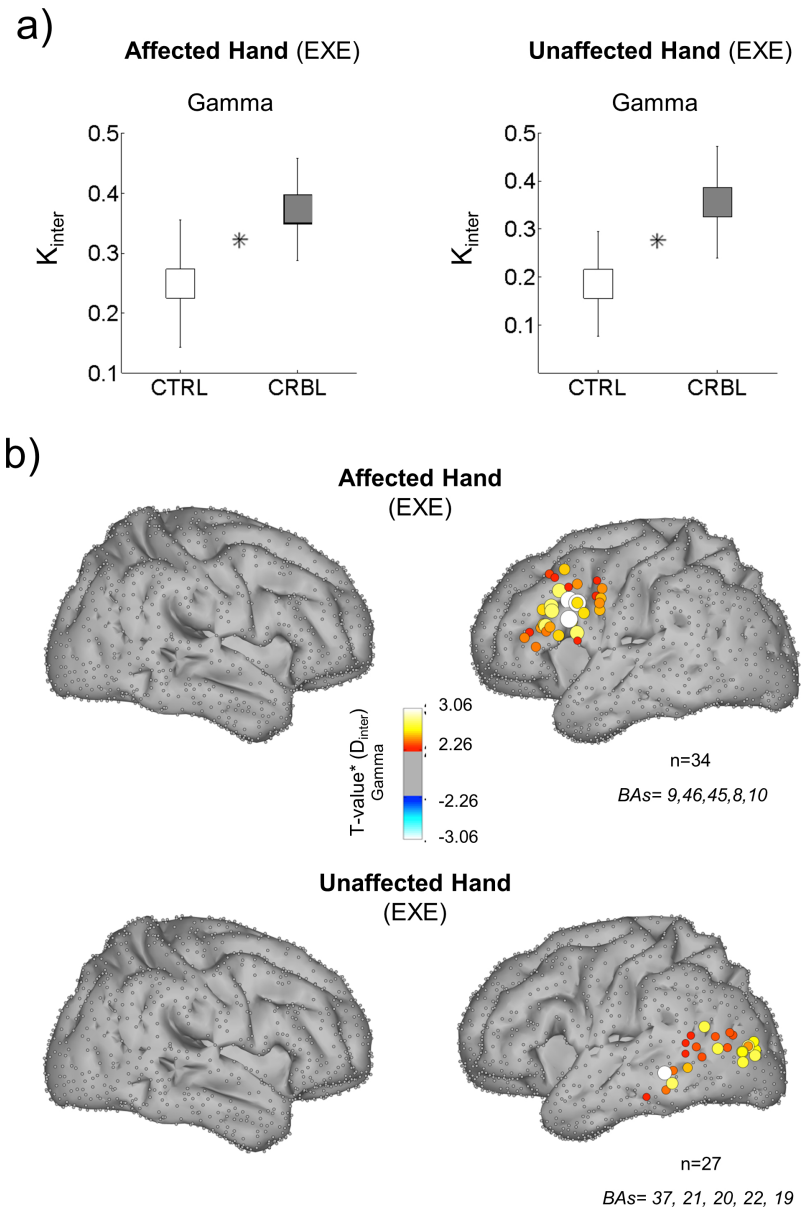

Figure 4. Gamma-band inter-hemispheric density (Kinter) and statistical contrasts of node degrees for brain networks during movement execution Panel a) Averaged Kinter values in the Gamma band for the CTRL and CRBL groups for the affected and unaffected hand conditions in the EXE phase. Panel b) T-value 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].

CASCADE Project-Team

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

CLIME Project-Team

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.

DYOGENE Project-Team

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.

EVA Project-Team

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.

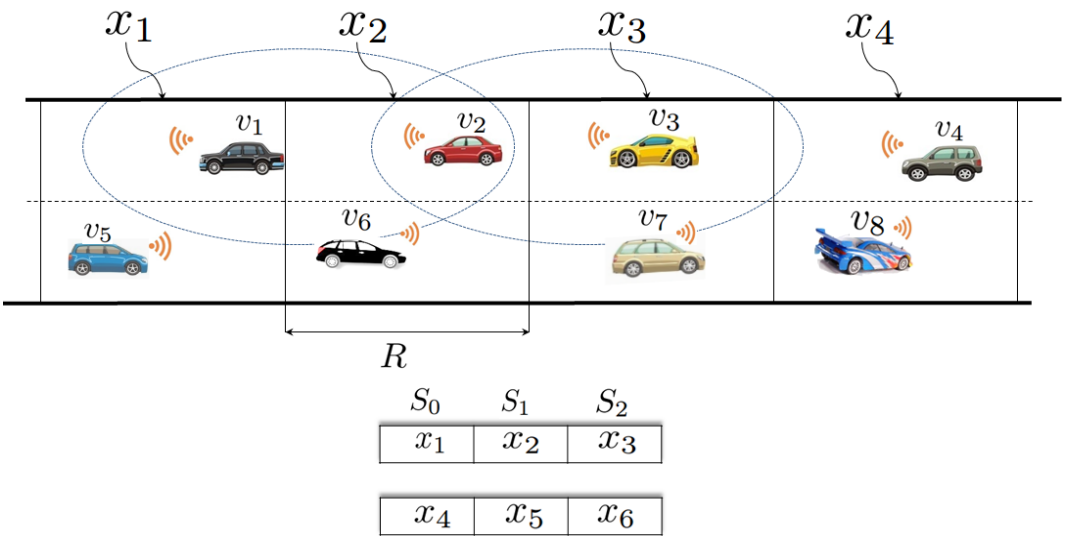

Figure 1. TDMA slots scheduling principle.

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).

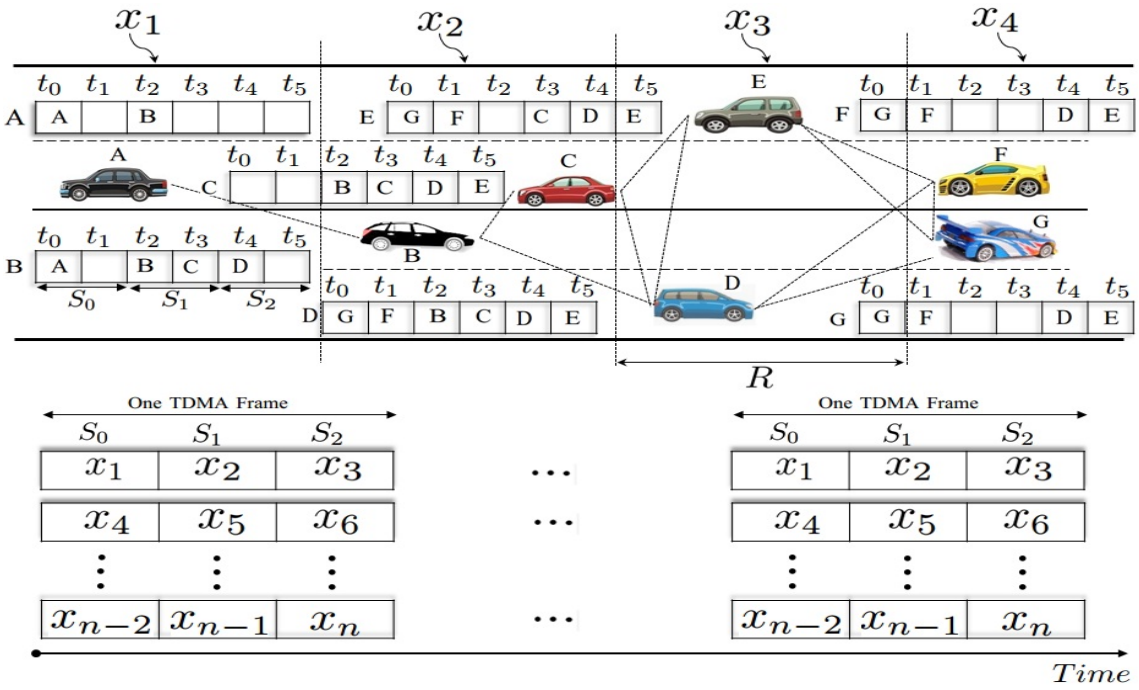

Figure 2. VANET network using DTMAC scheduling scheme.

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.

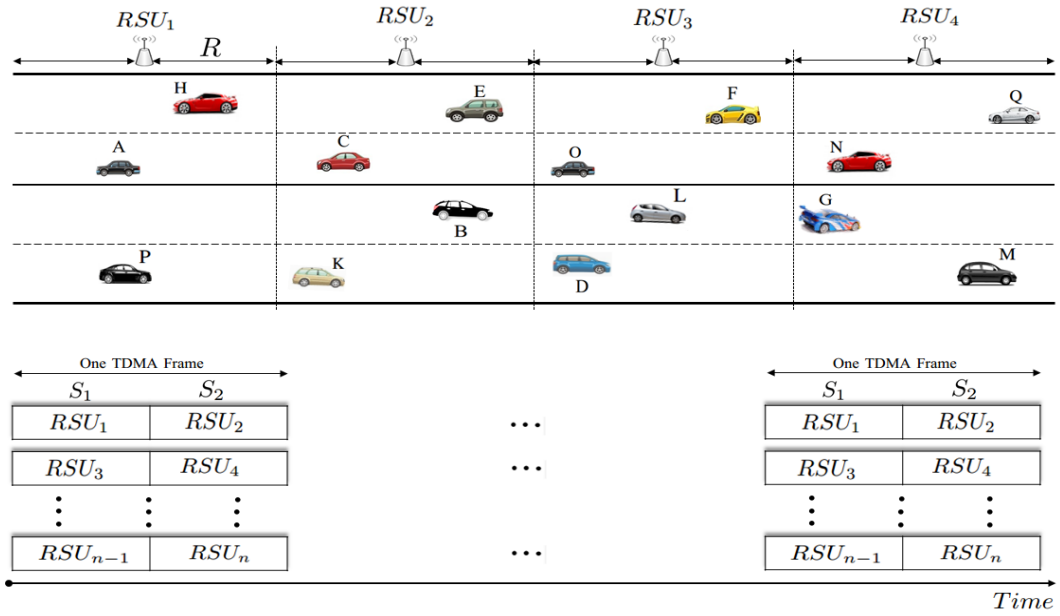

Figure 3. TDMA slots scheduling mechanism of CTMAC

This work was done in collaboration with Eric Renault, Telecom Sud Paris.

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.

GALLIUM Project-Team

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].

GANG Project-Team

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.

MAMBA Project-Team

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 .

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

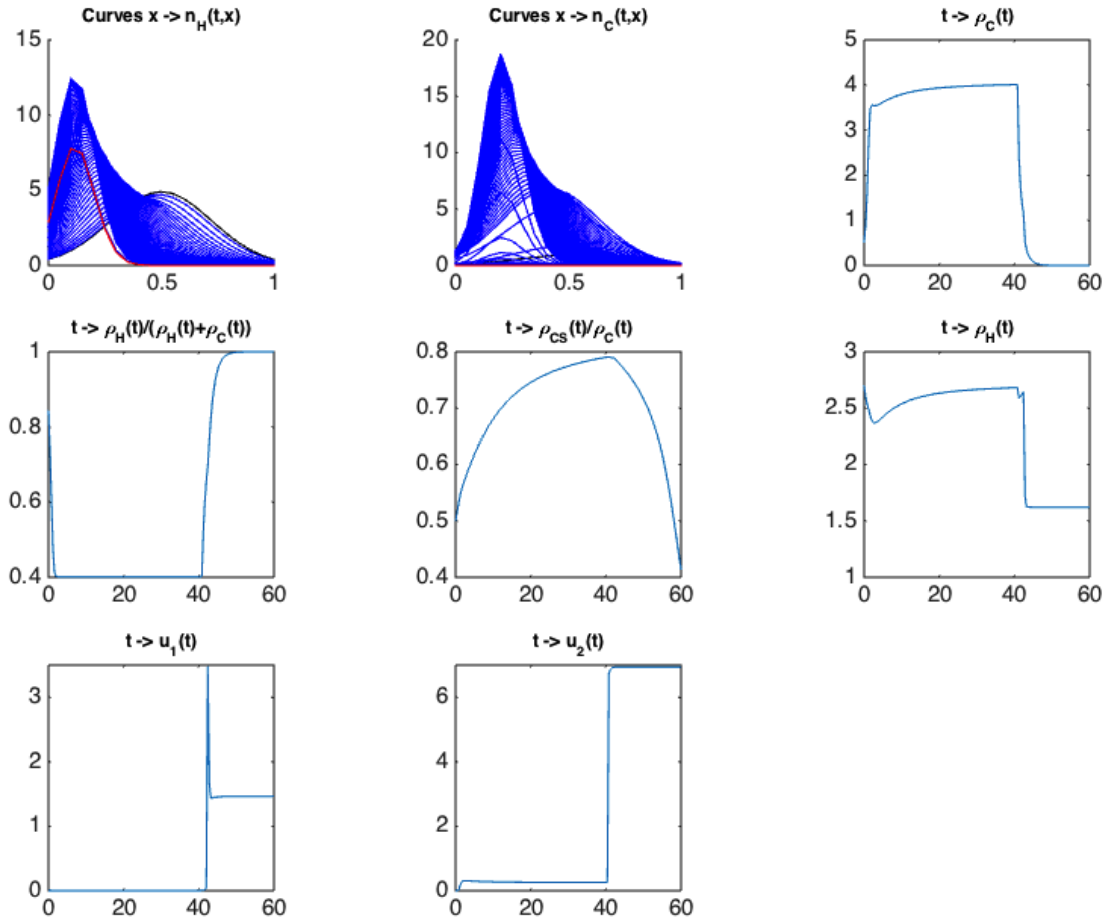

Figure 1. Simulation of the optimal control problem in time horizon $T = 60$. Top, left and middle: time evolution of $x \mapsto n_H(t,x)$, number of healthy cells with drug resistance expression phenotype x, and of $x \mapsto n_C(t,x)$, number of cancer cells with the same phenotype x. The initial conditions 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.

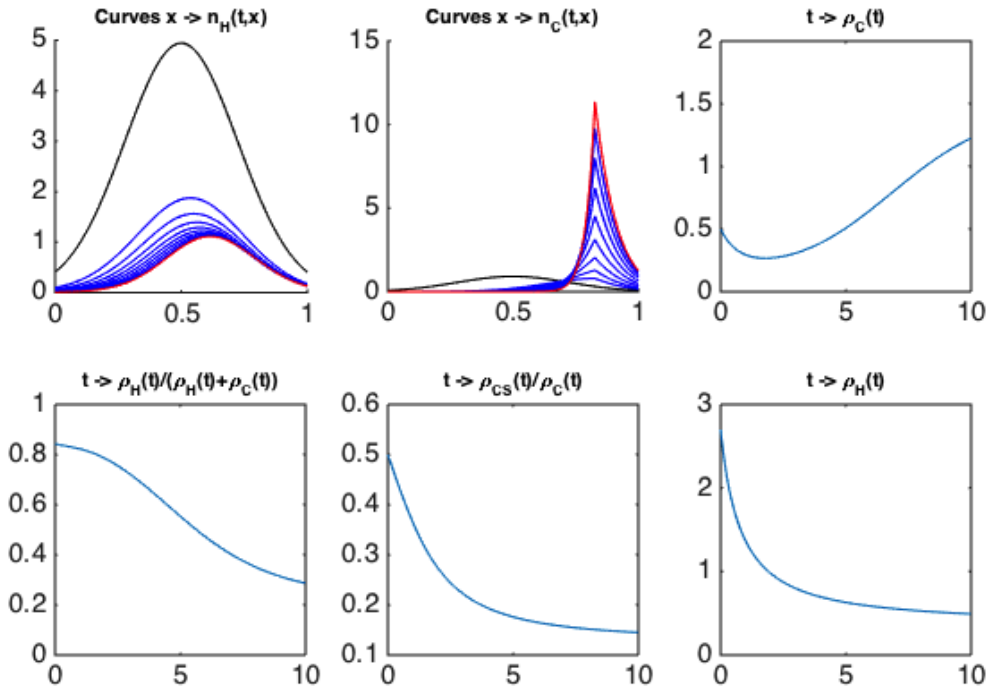

Figure 2. Comparative evolution with constant high drug doses. Catastrophic deleterious effects of the treatment on the concentration of the drug resistance phenotype x in the cancer cell population (top middle), and on the cell population numbers $\rho_C, \rho_H, \rho_{CS}$. Simulation with $u_1(t) = \mathrm{Cst} = 3.5$, $u_2(t) = \mathrm{Cst} = 2$, in time horizon $T = 10$.

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).

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

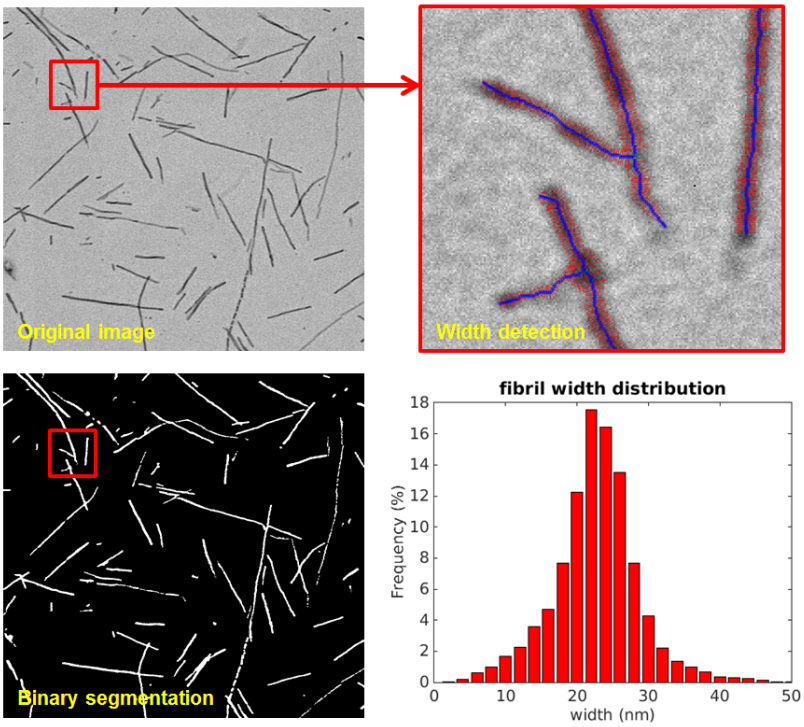

Figure 3. Fibril analysis yielding from original data via segmentation and analysis to the fibril length distribution.

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], 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)

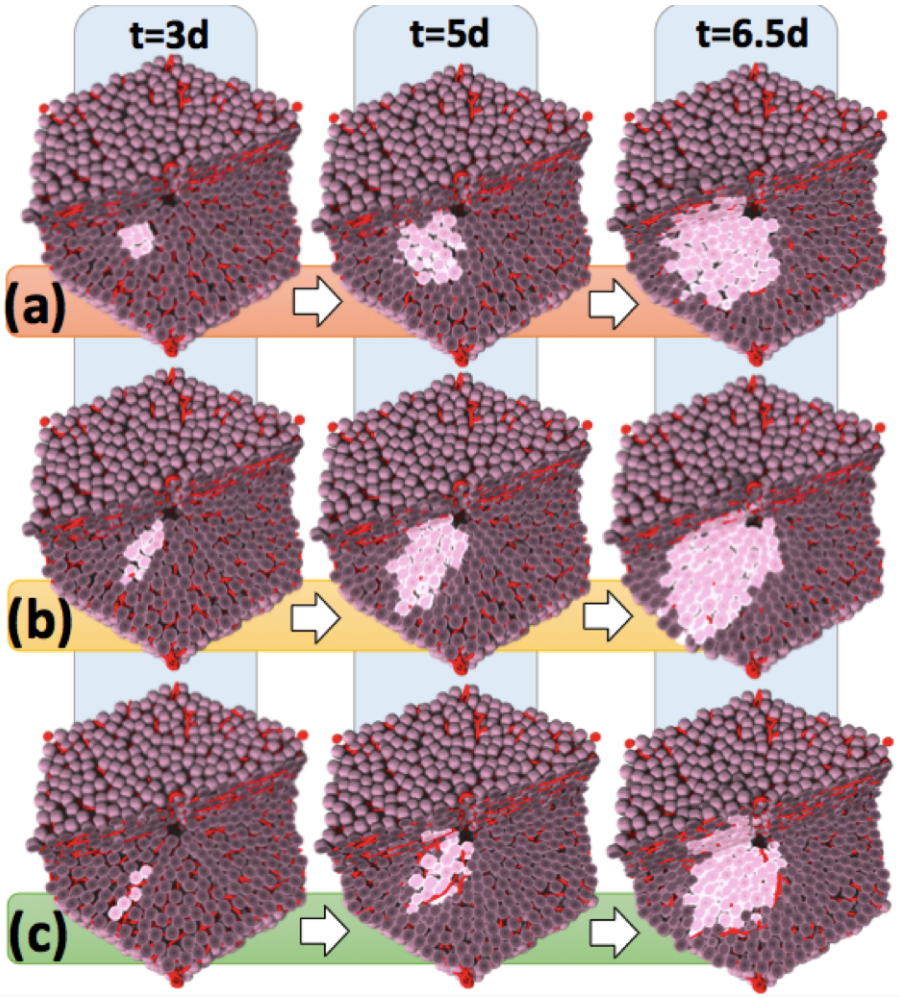

Figure 4. Scenarios of tumour growth in a single liver lobule in (a) absence of hepatocyte-sinusoidal-alignment (HSA), (b) presence of HSA, and (c) presence of HSA with elevated tangential friction impeding hepatocyte movement perpendicular to the columns formed along the sinusoids [35]. The images represent 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.

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.

MATHERIALS Project-Team

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.

MATHRISK Project-Team

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.

MIMOVE Team

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 cooperation 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:

- *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],

[0] http://sensormeasurement.appspot.com/?p=ontologies

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 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.

[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/
[0] https://play.google.com/apps/testing/edu.sarathi.metroCognition

MOKAPLAN Project-Team

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.

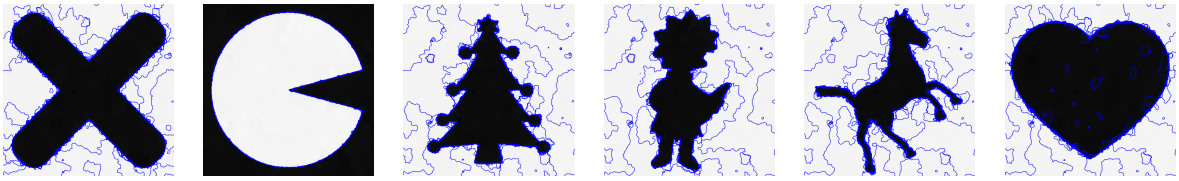

Figure 12. Level lines of denoised images with low regularization (TV denoising)

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.

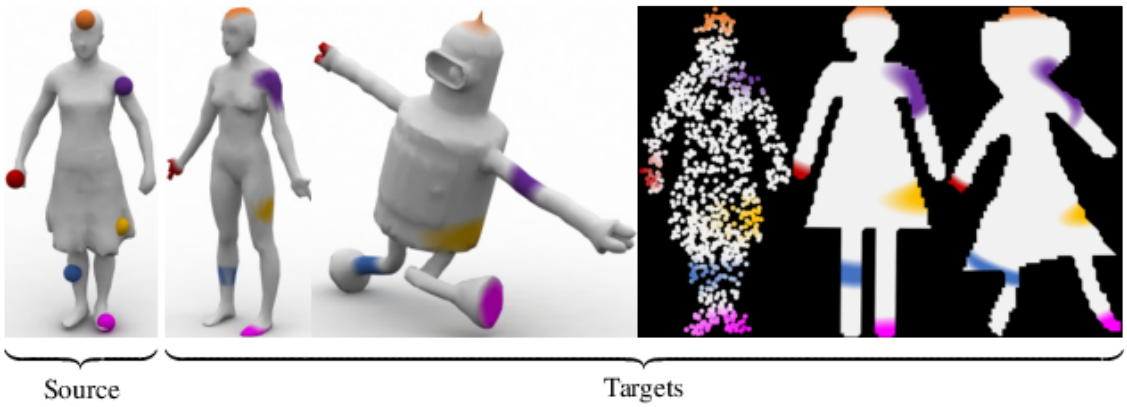

Figure 13. Example of matching induced between an input 3-D shape (on the left) and 3-D or 2-D shapes using the transport coupling computed using our entropy-regularized Gromov-Wasserstein problem. From [18].

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, 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.

MUSE Team

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.

MUTANT Project-Team

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 (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.

MYCENAE Project-Team

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.

PARKAS Project-Team

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).

PI.R2 Project-Team

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 $\mu\tilde{\mu}$-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 Gae¨tan Gilbert (PhD student of Nicolas Tabareau and Matthieu Sozeau at the E´cole 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.

POLSYS Project-Team

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.

PROSECCO Project-Team

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.

QUANTIC Project-Team

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].

RAP Project-Team

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].

REGAL Project-Team

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.

REO Project-Team

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.

RITS Project-Team

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].

SECRET Project-Team

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.

[0] http://competitions.cr.yp.to/caesar.html

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 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.

SERENA Team

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.

SIERRA Project-Team

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.

TAPDANCE Team (section vide)

WHISPER Project-Team

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.

WILLOW Project-Team

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 .

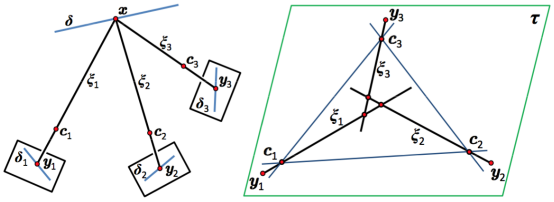

Figure 1. Left: Visual rays associated with three (correct) correspondences. Right: Degenerate epipolar constraints associated with three coplanar, but non-intersecting rays lying in the trifocal plane.

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 .

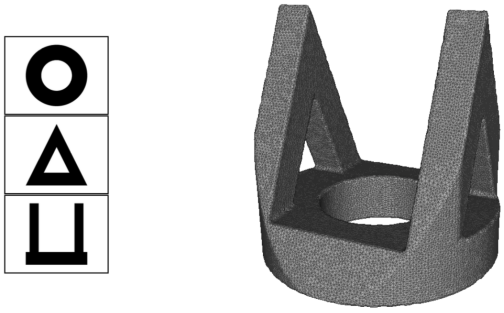

Figure 2. Geometrically consistent silhouettes are feasible projections of a single object.

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 .

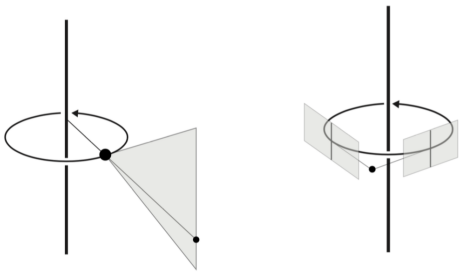

Figure 3. Non-central panoramic (left) and stereo panoramic cameras (right) are examples of non-linear cameras that can be modeled using line congruences.

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 .

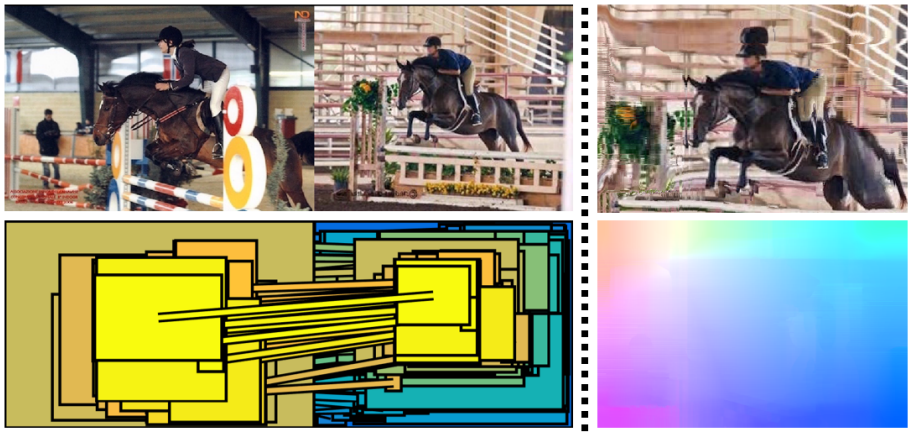

Figure 5. Proposal flow generates a reliable scene flow between similar images by establishing geometrically consistent correspondences between object proposals. (Left) Region-based scene flow by matching object proposals. (Right) Color-coded dense flow field generated from the region matches, and image 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/.

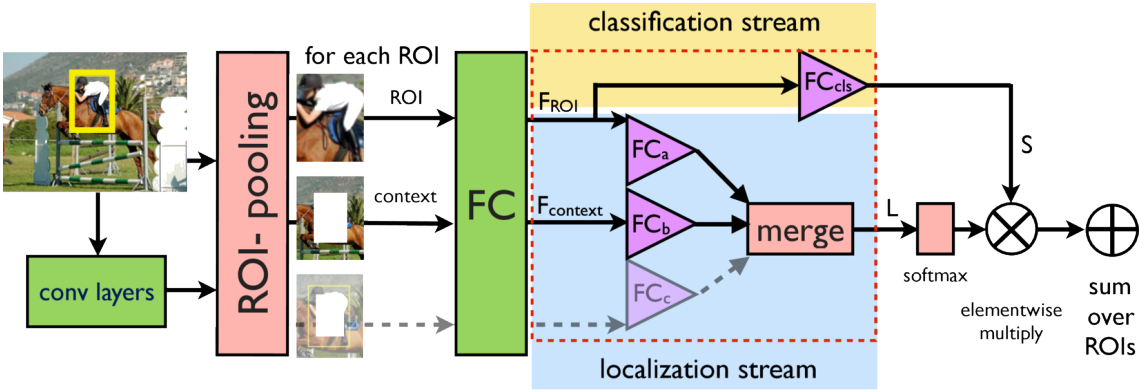

Figure 6. ContextLocNet improves localization by comparing an object score between a proposal and its context.

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.

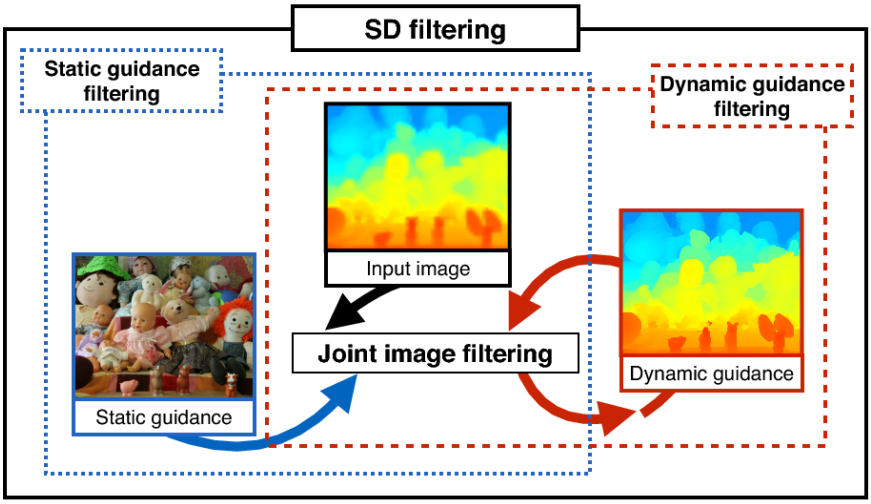

Figure 8. Sketch of joint image filtering and SD filtering: Static guidance filtering convolves an input image with a weight function computed from static guidance, as in the dotted blue box. Dynamic guidance filtering uses weight functions that are repeatedly obtained from regularized input images, 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.

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 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 .

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].

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.

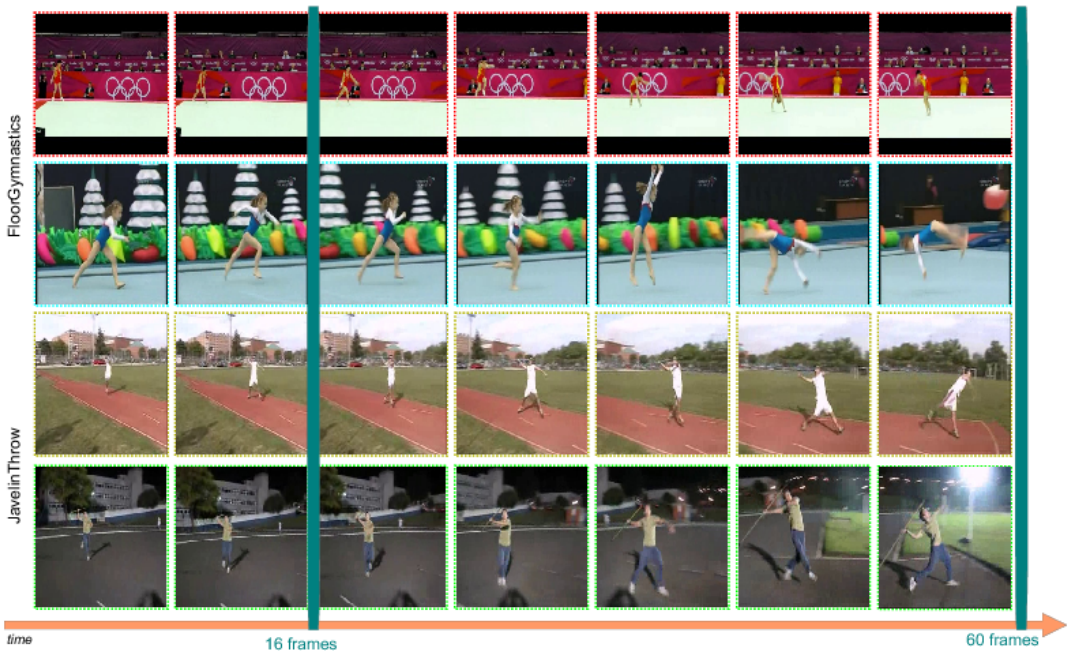

Figure 10. The highest improvement of long-term temporal convolutions in terms of class accuracy is for "JavelinThrow". For 16-frame network, it is mostly confused with "FloorGymnastics" class. We visualize sample videos with 7 frames extracted at every 8 frames. The intuitive explanation is that both 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.

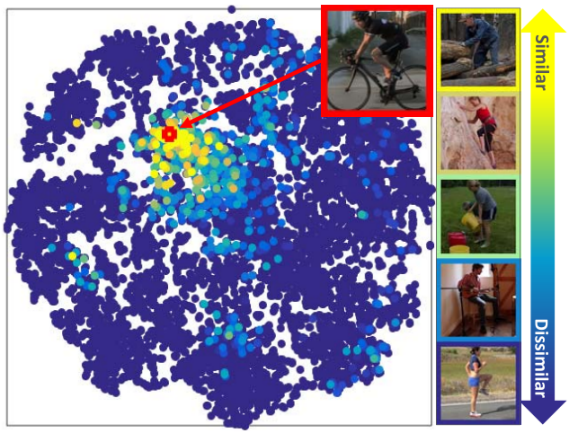

Figure 11. The manifold of our pose embedding visualized using t-SNE. Each point represents a human pose image. To better show correlation between the pose embedding and annotated pose, we color-code pose similarities in annotation between an arbitrary target image (red box) and all the other images. 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.

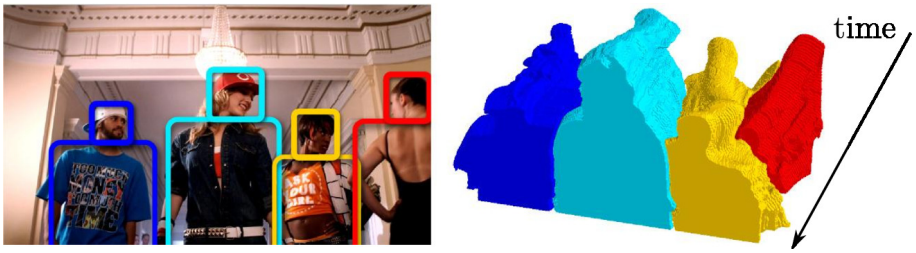

Figure 12. Results of our method applied to multi-person segmentation in a sample video from our database. Given an input video together with the tracks of object bounding boxes (left), our method finds pixel-wise segmentation for each object instance across video frames (right).