RESEARCH CENTER
Paris

FIELD

Activity Report 2018

Section New Results

Edition: 2019-03-07

ANTIQUE Project-Team

6. New Results

6.1. A Theoretical Foundation of Sensitivity in an Abstract Interpretation Framework

Participants: Xavier Rival [correspondant], Sukyoung Ryu, Se-Won Kim.

In [14], we formalize a framework to design static analyses that make use of sensitivity, using the general notion of cardinal power abstraction.

Program analyses often utilize various forms of *sensitivity* such as context sensitivity, call-site sensitivity, and object sensitivity. These techniques all allow for more precise program analyses, that are able to compute more precise program invariants, and to verify stronger properties. Despite the fact that sensitivity techniques are now part of the standard toolkit of static analyses designers and implementers, no comprehensive frameworks allow the description of all common forms of sensitivity. As a consequence, the soundness proofs of static analysis tools involving sensitivity often rely on *ad hoc* formalization, which are not always carried out in an abstract interpretation framework. Moreover, this also means that opportunities to identify similarities between analysis techniques to better improve abstractions or to tune static analysis tools can easily be missed.

In this work, we formalize a framework for the description of *sensitivity in static analysis*. Our framework is based on a powerful abstract domain construction, and utilizes reduced cardinal power to tie basic abstract predicates to the properties analyses are sensitive to. We formalize this abstraction, and the main abstract operations that are needed to turn it into a generic abstract domain construction. We demonstrate that our approach can allow for a more precise description of program states, and that it can also describe a large set of sensitivity techniques, both when sensitivity criteria are static (known before the analysis) or dynamic (inferred as part of the analysis), and sensitive analysis tuning parameters. Last, we show that sensitivity techniques used in state of the art static analysis tools can be described in our framework.

6.2. Memory Abstraction

6.2.1. Abstraction of arrays based on non contiguous partitions

Participants: Jiangchao Liu, Xavier Rival [correspondant].

In [15], we studied the verification of components of embedded programs that utilize arrays to store dynamically chained data-structures. Furthermore, this work constitutes a significant part of Jiangchao Liu's PhD Thesis ([10]).

User-space programs rely on memory allocation primitives when they need to construct dynamic structures such as lists or trees. However, low-level OS kernel services and embedded device drivers typically avoid resorting to an external memory allocator in such cases, and store structure elements in contiguous arrays instead. This programming pattern leads to very complex code, based on data-structures that can be viewed and accessed either as arrays or as chained dynamic structures. The code correctness then depends on intricate invariants mixing both aspects. We propose a static analysis that is able to verify such programs. It relies on the combination of abstractions of the allocator array and of the dynamic structures built inside it. This approach allows to integrate program reasoning steps inherent in the array and in the chained structure into a single abstract interpretation. We report on the successful verification of several embedded OS kernel services and drivers.

6.2.2. Semantic-Directed Clumping of Disjunctive Abstract States

Participants: Huisong Li, Francois Berenger, Bor-Yuh Evan Chang, Xavier Rival [correspondant].

In [29], we studied the semantic directed clumping of disjunctive abstract states. Furthermore, this work constitutes a significant part of Huisong Li's PhD Thesis ([9]).

To infer complex structural invariants, Shape analyses rely on expressive families of logical properties. Many such analyses manipulate abstract memory states that consist of separating conjunctions of basic predicates describing atomic blocks or summaries. Moreover, they use finite disjunctions of abstract memory states in order to account for dissimilar shapes. Disjunctions should be kept small for the sake of scalability, though precision often requires to keep additional case splits. In this context, deciding when and how to merge case splits and to replace them with summaries is critical both for the precision and for the efficiency. Existing techniques use sets of syntactic rules, which are tedious to design and prone to failure. In this paper, we design a semantic criterion to clump abstract states based on their silhouette which applies not only to the conservative union of disjuncts, but also to the weakening of separating conjunction of memory predicates into inductive summaries. Our approach allows to define union and widening operators that aim at preserving the case splits that are required for the analysis to succeed. We implement this approach in the MemCAD analyzer, and evaluate it on real-world C codes from existing libraries, including programs dealing with doubly linked lists, red-black trees and AVL-trees.

6.3. Static Analysis of JavaScript Code

6.3.1. Weakly Sensitive Analysis for Unbounded Iteration over JavaScript Objects

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

In [28], we studied composite object abstraction for the analysis JavaScript.

JavaScript framework libraries like jQuery are widely use, but complicate program analyses. Indeed, they encode clean high-level constructions such as class inheritance via dynamic object copies and transformations that are harder to reason about. One common pattern used in them consists of loops that copy or transform part or all of the fields of an object. Such loops are challenging to analyze precisely, due to weak updates and as unrolling techniques do not always apply. In this work, we observe that precise field correspondence relations are required for client analyses (e.g., for call-graph construction), and propose abstractions of objects and program executions that allow to reason separately about the effect of distinct iterations without resorting to full unrolling. We formalize and implement an analysis based on this technique. We assess the performance and precision on the computation of call-graph information on examples from jQuery tutorials.

6.4. Communication-closed asynchronous protocols

Participants: Andrei Damien, Cezara Drăgoi [correspondant], Alexandru Militaru, Josef Widder.

Fault-tolerant distributed systems are implemented over asynchronous networks, so that they use algorithms for asynchronous models with faults. Due to asynchronous communication and the occurrence of faults (e.g., process crashes or the network dropping messages) the implementations are hard to understand and analyze. In contrast, synchronous computation models simplify design and reasoning. In this paper, we bridge the gap between these two worlds. For a class of asynchronous protocols, we introduce a procedure that, given an asynchronous protocol, soundly computes its round-based synchronous counterpart. This class is defined by properties of the sequential code. We computed the synchronous counterpart of known consensus and leader election protocols, such as, Paxos, and Chandra and Toueg's consensus. Using Verifast we checked the sequential properties required by the rewriting. We verified the round-based synchronous counter-part of Multi-Paxos, and other algorithms, using existing deductive verification methods for synchronous protocols.

6.5. Borel Kernels and their Approximation, Categorically

Participants: Fredrik Dahlqvist, Alexandra Silva, Vicent Danos [correspondant], Ilias Garnier.

In [12] is introduced a categorical framework to study the exact and approximate semantics of probabilistic programs. We construct a dagger symmetric monoidal category of Borel kernels where the dagger-structure is given by Bayesian inversion. We show functorial bridges between this category and categories of Banach lattices which formalize the move from kernel-based semantics to predicate transformer (backward) or state transformer (forward) semantics. These bridges are related by natural transformations, and we show in particular that the Radon-Nikodym and Riesz representation theorems-two pillars of probability theory-define natural transformations. With the mathematical infrastructure in place, we present a generic and endogenous approach to approximating kernels on standard Borel spaces which exploits the involutive structure of our category of kernels. The approximation can be formulated in several equivalent ways by using the func-torial bridges and natural transformations described above. Finally, we show that for sensible discretization schemes, every Borel kernel can be approximated by kernels on finite spaces, and that these approximations converge for a natural choice of topology. We illustrate the theory by showing two examples of how approximation can effectively be used in practice: Bayesian inference and the Kleene * operation of ProbNetKAT.

6.6. Static analysis of rule-based models

Thanks to rule-based modeling languages, we can assemble large sets of mechanistic protein-protein interactions within integrated models. Our goal would be to understand how the behavior 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. To this extend, static analysis can be used to derive various abstractions of the semantics, each of them offering new perspectives on the models.

6.6.1. Trace approximation

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

In [13], we propose an abstract interpretation of the behavior of each protein, in isolation. Given a model written in Kappa, this abstraction computes for each kind of proteins a transition system that describes which conformations this protein may take and how a protein may 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 behavior of each protein of the model.

6.6.2. Detection of polymer formation

Participants: Pierre Boutillier, Aurélie Faure de Pebeyre, Jérôme Feret [correspondant].

Rule-based languages, such as Kappa and BNGL, allow for the description of very combinatorial models of interactions between proteins. A huge (when not infinite) number of different kinds of bio-molecular compounds may arise due to proteins with multiple binding and phosphorylation sites. Knowing beforehand whether a model may involve an infinite number of different kinds of bio-molecular compounds is crucial for the modeler. On the first hand, having an infinite number of kinds of bio-molecular compounds is sometimes a hint for modeling flaws: forgetting to specify the conflicts among binding rules is a common mistake. On the second hand, it impacts the choice of the semantics for the models (among stochastic, differential, hybrid).

In [22], we introduce a data-structure to abstract the potential unbounded polymers that may be formed in a rule-based model. This data-structure is a graph, the nodes and the edges of which are labeled with patterns. By construction, every potentially unbounded polymer is associated to at least one cycle in that graph. This data-structure has two main advantages. Firstly, as opposed to site-graphs, one can reason about cycles without enumerating them (by the means of Tarjan's algorithm for detecting strongly connected components). Secondly, this data-structures may be combined easily with information coming from additional reachability analysis: the edges that are labeled with an overlap that is proved unreachable in the model may be safely discarded.

6.6.3. The static analyzer KaSa

Participants: Pierre Boutillier, Ferdinanda Camporesi, Jean Coquet, Jérôme Feret [correspondant], Kim Quyên Lý, Nathalie Théret, Pierre Vignet.

KaSa is a static analyzer for Kappa models. Its goal is two-fold. Firstly, KaSa assists the modeler by warning about potential issues in the model. Secondly, KaSa may provide useful properties to check that what is implemented is what the modeler has in mind and to provide a quick overview of the model for the people who have not written it. The cornerstone of KaSa is a fix-point engine which detects some patterns that may never occur whatever the evolution of the system may be. From this, many useful information may be collected KaSa warns about rules that may never be applied, about potential irreversible transformations of proteins (that may not be reverted even thanks to an arbitrary number of computation steps) and about the potential formation of unbounded molecular compounds. Lastly, KaSa detects potential influences (activation/inhibition relation) between rules.

In [21], we illustrate the main features of KaSa on a model of the extracellular activation of the transforming growth factor, TGF-b.

6.7. The Kappa platform for rule-based modeling

Participants: Pierre Boutillier, Mutaamba Maasha, Xing Li, Héctor Medina-Abarca, Jean Krivine, Jérôme Feret [correspondant], Ioana Cristescu, Angus Forbes, Walter Fontana.

In [11], we present an overview of the Kappa platform, an integrated suite of analysis and visualization techniques for building and interactively exploring rule-based models. The main components of the platform are the Kappa Simulator, the Kappa Static Analyzer and the Kappa Story Extractor. In addition to these components, we describe the Kappa User Interface, which includes a range of interactive visualization tools for rule-based models needed to make sense of the complexity of biological systems. We argue that, in this approach, modeling is akin to programming and can likewise benefit from an integrated development environment. Our platform is a step in this direction.

We discuss details about the computation and rendering of static, dynamic, and causal views of a model, which include the contact map (CM), snapshots at different resolutions, the dynamic influence network (DIN) and causal compression. We provide use cases illustrating how these concepts generate insight. Specifically, we show how the CM and snapshots provide information about systems capable of polymerization, such as Wnt signaling. A well-understood model of the KaiABC oscillator, translated into Kappa from the literature, is deployed to demonstrate the DIN and its use in understanding systems dynamics. Finally, we discuss how pathways might be discovered or recovered from a rule-based model by means of causal compression, as exemplified for early events in EGF signaling.

The Kappa platform is available via the project website at kappa-language.org. All components of the platform are open source and freely available through the authors' code repositories.

6.8. Conservative approximation of systems of differential equations

We design a tools-kit to reason and abstract the solutions of the systems of differential equations that are described in high-level languages. Our abstractions are conservative in the sense that they provided sound lower and upper bounds for the value of some observables of the system. Our approach consists, firstly, in inferring structural equalities about combinations of variables and structural inequalities about the value of variable derivatives thanks to symbolic reasoning at the level of the languages and, then, in using these numerical constraints to infer two differential equations for the variables of interest — one for the lower bound and one for the upper bound.

We focus on the systems of equations that are described in Kappa. Our goal is to provide a unifying framework that can deal with heterogeneous kinds of abstractions, including truncation, time- and concentration-scale separations, flow-based reduction, symmetries-based reduction.

6.8.1. Approximation of models of polymers

Participants: Ken Chanseau Saint-Germain, Jérôme Feret [correspondant].

We propose a systematic approach to approximate the behavior of models of polymers synthesis/degradation, described in Kappa. Our abstraction consists in focusing on the behavior of all the patterns of size less than a given parameter. We infer symbolic equalities and inequalities which intentionally may be understood as algebraic constructions over patterns, and extensionally as sound properties about the concentration of the bio-molecular species that contain these patterns. Then, we derive a system of equations describing the time evolution of a lower and an upper bounds for the concentration of each pattern of interest.

This work has been presented at VEMDP 2018 (Verification of Engineered Molecular Devices and Programs), in Oxford, 19th July 2018, and at the days "BIOS-IA" of the working group BIOSS, at Pasteur Institute, Paris, 18th December 2018.

6.8.2. Approximation based on time- and/or concentration-scale separation

Participants: Andreea Beica, Jérôme Feret [correspondant].

In [20], we have designed and tested an approximation method for ODE models of biochemical reaction systems, in which the guarantees are our major requirement. Borrowing from tropical analysis techniques, we look at the dominance relations among terms of each species' ODE. These dominance relations can be exploited to simplify the original model, by neglecting the dominated terms. As the dominant subsystems can change during the system's dynamics, depending on which species dominate the others, several possible modes exist. Thus, simpler models consisting of only the dominant subsystems can be assembled into hybrid, piece-wise smooth models, which approximate the behavior of the initial system. By combining the detection of dominated terms with symbolic bounds propagation, we show how to approximate the original model by an assembly of simpler models, consisting in ordinary differential equations that provide time-dependent lower and upper bounds for the concentrations of the initial models species. The utility of our method is twofold. On the one hand, it provides a reduction heuristics that performs without any prior knowledge of the initial system's behavior (i.e., no simulation of the initial system is needed in order to reduce it). On the other hand, our method provides sound interval bounds for each species, and hence can serve to evaluate the faithfulness of tropicalization reduction heuristics for ODE models of biochemical reduction systems. The method is tested on several case studies.

6.9. Sources, propagation and consequences of stochasticity in cellular growth

Participants: Philipp Thomas, Guillaume Terradot, Vicent Danos [correspondant], Andrea Weiße.

Growth impacts a range of phenotypic responses. Identifying the sources of growth variation and their propagation across the cellular machinery can thus unravel mechanisms that underpin cell decisions.

In [17], we present a stochastic cell model linking gene expression, metabolism and replication to predict growth dynamics in single bacterial cells. Alongside we provide a theory to analyze stochastic chemical reactions coupled with cell divisions, enabling efficient parameter estimation, sensitivity analysis and hypothesis testing. The cell model recovers population-averaged data on growth-dependence of bacterial physiology and how growth variations in single cells change across conditions. We identify processes responsible for this variation and reconstruct the propagation of initial fluctuations to growth and other processes. Finally, we study drug-nutrient interactions and find that antibiotics can both enhance and suppress growth heterogeneity. Our results provide a predictive framework to integrate heterogeneous data and draw testable predictions with implications for antibiotic tolerance, evolutionary and synthetic biology.

6.10. Survival of the Fattest: Evolutionary Trade-offs in Cellular Resource Storage

Participants: Guillaume Terradot, Andreea Beica, Andrea Weiße, Vicent Danos [correspondant].

Cells derive resources from their environments and use them to fuel the bio-synthetic processes that determine cell growth. Depending on how responsive the bio-synthetic processes are to the availability of intracellular resources, cells can build up different levels of resource storage.

In [16], 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.

6.11. A Genetic Circuit Compiler: Generating Combinatorial Genetic Circuits with Web Semantics and Inference

Participants: William Waites, Goksel Misirli, Matteo Cavaliere, Vicent Danos [correspondant].

A central strategy of synthetic biology is to understand the basic processes of living creatures through engineering organisms using the same building blocks. Biological machines described in terms of parts can be studied by computer simulation in any of several languages or robotically assembled in vitro. In [19] we present a language, the Genetic Circuit Description Language (GCDL) and a compiler, the Genetic Circuit Compiler (GCC). This language describes genetic circuits at a level of granularity appropriate both for automated assembly in the laboratory and deriving simulation code. The GCDL follows Semantic Web practice and the compiler makes novel use of the logical inference facilities that are therefore available. We present the GCDL and compiler structure as a study of a tool for generating κ-language simulations from semantic descriptions of genetic circuits.

6.12. An Information-Theoretic Measure for Patterning in Epithelial Tissues

Participants: William Waites, Matteo Cavaliere, Élise Cachat, Vicent Danos [correspondant], Jamie A. Davies.

In [18], we present path entropy, an information-theoretic measure that captures the notion of patterning due to phase separation in organic tissues. Recent work has demonstrated, both in silico and in vitro, that phase separation in epithelia can arise simply from the forces at play between cells with differing mechanical properties. These qualitative results give rise to numerous questions about how the degree of patterning relates to model parameters or underlying biophysical properties. Answering these questions requires a consistent and meaningful way of quantifying degree of patterning that we observe. We define a resolution-independent measure that is better suited than image-processing techniques for comparing cellular structures. We show how this measure can be usefully applied in a selection of scenarios from biological experiment and computer simulation, and argue for the establishment of a tissue-graph library to assist with parameter estimation for synthetic morphology.

AOSTE2 Team

7. New Results

7.1. Uniprocessor Mixed-Criticality Real-Time Scheduling

Participants: Liliana Cucu, Robert Davis, Mehdi Mezouak, Yves Sorel.

In the context of the FUI CEOS project 9.1.1.1 , last year we tranformed the PX4 autopilot free software program in a graph of tasks. In this project our main goal is to perform a real-time schedulability analysis on this program in order to prove that the autopilot will meet all its deadlines when it will operate in the multirotor drone the project is intended to built. The tasks will be executed on a Pixhawk electronic board based on an ARM Cortex M4 microprocessor running on the NuttX OS.

We start by determinating the period and measuring the average execution time of each task which is less than the worst case execution time (WCET). Then, using these periods and these measured execution times we perform an online schedulability analysis using a rate monotonic policy (RM) that shown the set of tasks is not schedulable. Consequently, we informed the partners of the CEOS project that the present version of PX4 is not real-time.

Presently, we are transforming the original set of tasks into a set of real-time tasks. To achieve this goal, we associate to every task a periodic high resolution timer corresponding to the period of the task. Each timer generates an interruption when it expires and the task is put in the ready task queue. The scheduler of NuttX will choose in this queue the task to be executed. In order to validate this transformation we operated the multirotor drone in a simulation tool composed of Gazebo for the geometrical environment of the drone and of the Ground Control Station for setting and controlling the drone. We performed two kinds of simulations, a software in the loop simulation (SitL) which simulates the Pixhawk board, the sensors and the actuators, and a hardware in the loop simulation (HitL) which simulates only the sensors and the actuators, whereas the PX4 program runs on the Pixhawk board. We tested the set of real-time tasks in SitL and we are presently testing them in HitL.

Since we can easily change the period of every task, we plan to modify the periods to make the set of real-time tasks schedulable using an online RM schedulability analysis.

In order to manage high criticality real-time tasks we plan to use an offline scheduler whose scheduling table is generated by an offline schedulability analysis tool that is developped in the team. We plan to modify NuttX in order to support such scheduler.

Finally, in order to complete the real-time schedulability analysis of PX4, we estimate the worst case execution time (WCET) of each task. This problem is complex due to the multiple possible paths in a task as well as the different data it consumes. Moreover, the processor and/or the microcontroller itself may have some features like memory contentions, bus accesses, caches, pipelines, speculative branchings that increase the difficulty to determine WCETs. All these variabilities lead us to introduce statistical reasoning in characterizing the timing behavior (WCET, schedulability analyses) of mixed-criticality real-time applications. The isolated execution times of the programs have indicated large variations indicating expected larger variability in real execution scenarios. In order to decrease the pessimism of the statistical bounds, we are adapting our models to move towards multi-variate approaches.

7.2. Multiprocessor Real-Time Scheduling

Participants: Slim Ben Amor, Evariste Ntaryamira, Salah Eddine Saidi, Yves Sorel, Walid Talaboulma.

The last part of the PhD thesis of Salah Eddine Saidi, was dedicated to the parallelization of FMI-based co-simulation under real-time constraints. More precisely we address HiL (Hardware in the Loop) co-simulation where a part of the co-simulation is replaced by its real counterpart which is physically available. The real and simulated parts have to exchange data during the execution of the co-simulation under real-time constraints. In other words, the inputs (resp. ouputs) of the real part are sampled periodically, sending (resp. receiving) data to (resp. from) the simulated part. Every periodic data exchange defines a set of real-time constraints to be satisfied by the simulated part. We proposed a method for defining these real-time constraints and propagating them to all the data dependent functions that specify the co-simulation (simulated part). Starting from these constraints we have to schedule the FMI-based co-simulation on a multi-core. We propose an ILP-based algorithm as well as a heuristic that allow the execution of the co-simulation on a multi-core processor while ensuring the previously defined real-time constraints are respected [6]. The proposed heuristic is a list scheduling heuristic. It builds the multi-core schedule iteratively. At each iteration, a list of candidate functions is constructed. The heuristic computes the priority for each candidate function on every core and selects the core for the which the priority is maximized. The priority of a function is a dynamic priority as its computation depends on the partial scheduling solution that has already been computed.

All works achieved by Salah Eddine Saidi on the parallelization of FMI-based co-simulation of numerical models were presented in his PhD thesis defense and manuscript [1].

Avionics applications are based on the specification of "data chains". Every data chain is a sequence of periodic real-time communicating tasks that are processing the data from sensors up to actuators. Such data chain determines an order in which the tasks propagate data but not in which they are executed. Indeed, inter-task communication and scheduling are independent. We focus on the latency computation, considered as the time elapsed from getting the data from an input and processing it to an output of a data chain. We propose a method for the worst-case latency computation of data chains composed of periodic tasks and executed by a partitioned fixed-priority preemptive scheduler upon a multiprocessor platform [5].

The PhD thesis of Slim Ben Amor is dedicated to the study of multiprocessor scheduling of real-time systems in presence of precedence constraints. This year we have proposed new models [10] for dependent real-time task with probabilistic worst-case execution time (WCET) that are scheduled using a partitioned reasoning. We explore existing solutions from [15] as the closest problem to our dependent task scheduling on multiprocessor and we study their extension to probabilistic models. We conclude that the probabilistic extension would be very difficult with heavy computation since the deterministic solution is based on the resolution of complex ILP optimization problem. Then, we decide to build a new solution to the deterministic problem that should be simple to extend to probabilistic problem. The proposed solution [11] consists of calculating the response time of each sub-tasks in a given DAG task taking in consideration preemptions caused by higher priority sub-tasks executed on the same processor. Then, we evaluate the global response time of the whole graph layer by layer, which allows deciding the schedulability of the entire system.

During the third year of Walid Talaboulma PhD thesis, we continued exploring solutions to make the WCET (Worst Case Execution Time) estimation as independent as possible with respect to the memory accesses. WCET analysis done on a unicore processor (in isolation) is not sufficient when we run our tasks on a multicore processors, the problem of Co-runner interference arises due to contention in shared hardware. Our solution is based on the generation of programs memory access profile, that we obtain by running tasks on a cycle accurate System Simulator, with a precise cycle accurate model of DDRAM memory controller and a full model of memory hierarchy including caches and main memory devices, and we log every memory event that occurs inside the simulation. Our solution does not necessarily require modifications of software layer, or recompilation of task code. We use those profiles to account for co runners interference and add it to WCET value obtained in isolation, and by updating our schedule, we can also insert idle times at correct scheduling events to decrease the interference.

The PhD thesis of Evariste Ntaryamira is dedicated to the study of multiprocessor real-time systems while ensuring the data freshness. This year we have underlined the difficulty of this scheduling problem [13], [8] while proposing a model to include both time and data constraints. We explore existing solutions from [16]

as the closest problem to our data-dependent scheduling problem. The case study associated to this thesis is jointly prepared with the members of the RITS Inria team.

7.3. Safe Parallelization of Hard Real-Time Avionics Software

Participants: Keryan Didier, Dumitru Potop Butucaru.

This work took place in the framework of the ITEA3 ASSUME project, which funds the PhD thesis of Keryan Didier, and in close collaboration with Inria PARKAS, Airbus, Safran Aircraft Engines, and Kalray.

The key difficulty of real-time scheduling is that timing analysis and resource allocation depend on each other. An exhaustive search for the optimal solution not being possible for complexity reasons, heuristic approaches are used to break this dependency cycle. Two such approaches are typical in real-time systems design. The first approach uses unsafe timing characterizations for the tasks (e.g., measurements) to build the system, and then checks the respect of real-time requirements through a global timing analysis. The second approach uses a formal model of the hardware platform enabling timing characterizations that are safe for all possible resource allocations (worst-case bounds).

So far, the practicality of the second approach had never been established. Automated real-time parallelization flows still relied on simplified hypotheses ignoring much of the timing behavior of concurrent tasks, communication and synchronization code. And even with such unsafe hypotheses, few studies and tools considered the—harmonic—multiperiodic task graphs of real-world control applications, and the problem of statically managing all their computational, memory, synchronization and communication resources.

This year, we presented the first demonstration of the feasibility of the second approach, showing good practical results for classes of real-world applications and multiprocessor execution platforms whose timing predictability allows keeping pessimism under control. This requires something that is missing in previous work: *the tight orchestration of* **all** *implementation phases*: WCET analysis, resource allocation, generation of *glue code* ensuring the sequencing of tasks on cores and the synchronization and memory coherency between the cores, compilation and linking of the resulting C code. This orchestration is conducted on very detailed timing model that considers both the tasks and the generated glue code, and which includes resource access interferences due to multi-core execution. While orchestration is our main contribution, it should not be understood as a mere combination of existing tools and algorithms. The whole point of our approach is to carefully coordinate every analysis, mapping and code generation phase to enable predictable execution and to keep pessimism under control. To this end, we contributed application normalization phase to facilitate timing analysis, an original code generation algorithm designed to provide mapping-independent worst-case execution time bounds, and new real-time scheduling algorithms capable of orchestrating memory allocation and scheduling.

Our flow scales to an avionics application comprising more than 5000 unique nodes, targeting the Kalray MPPA 256 many-core platform, selected for its timing predictability. First results are presented in the report [9].

7.4. Real-time Platform Modeling

Participants: Fatma Jebali, Dumitru Potop Butucaru.

This work took place in the framework of the ITEA3 ASSUME project, which funds the post-doc of Fatma Jebali.

One key difficulty in embedded systems design is the existence of multiple models of the same hardware system, developed separately, at different abstraction levels, and used in various phases of the design flow. In the design of real-time embedded systems, we can identify, among other:

- Cycle-accurate system models used to perform fine-grain hardware simulation, mostly during HW and driver design phases. These models provide an exact functional and temporal representation of system execution.

- Microarchitectural models used for pipeline simulation during WCET (*Worst-Case Execution Time*) analysis [19], [20], [18]. These models are used to compute safe over-approximations of the duration of a sequential piece of code, i.e., one function running without interruption on a processor core). To provide precise results, these models preserve much of the microarchitectural detail of processor pipelines and memory hierarchy (e.g. cache states, data transfer latencies).

Both simulation models usually have cyclic activation patterns, but establishing semantic consistency between them is challenging for several reasons. First, the activation pattern, which is the logical time base of the simulation, depends on the abstraction level. In cycle-accurate models, simulation cycles correspond to hardware clock ticks, whereas in WCET analysis models they correspond to changes in the program counter of the sequential program. Second, data abstractions are different in the two simulation models. Cycle-accurate simulators are often also *bit-accurate*, *i.e.* provide exactly the same results as the actual hardware. By comparison, pipeline simulators in WCET analysis abstract away most data types and related operators, typically retaining only Booleans, which can be exploited at analysis time. Last, but not least, the simulators are usually pieces of C/C++ code manually written by different teams or obtained through complex translation processes from high-level Architecture Description Languages (ADLs) that may not have a clear semantics. Formally relating such pieces of code is difficult.

This year we proposed a method to ensure the semantic consistency between the two HW models we consider, focusing on time abstraction issues. Our method relies on *desynchronization* theory [25], which defines sufficient properties ensuring that a synchronous model can be seen as an asynchronous Kahn Process Network (KPN). When a synchronous HW model satisfies these properties, any scheduling of its computations that is compatible with data dependencies will produce the same result (a property known as scheduling-independence). We showed how to control scheduling through changes of the logical time base of the model prior to code generation using a synchronous language compiler. In particular, a careful choice of the logical time base allows us to produce, from the same model, either a cycle-accurate simulator, or the one needed for WCET analysis. In conjunction with some data abstraction, this logical time manipulation allows the synthesis of semantically consistent simulators from a single model.

Furthermore, we can ensure by construction that synchronous models satisfy the properties required by desynchronization theory. To this end, we introduced a new hardware modelling language, named xMAStime, allowing the compositional modeling of systems satisfying the required properties. Results were presented at the ACSD'18 conference [4].

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):

- Advanced primitives for privacy in the cloud
- Efficient functional encryption
- Several predicate-encryption schemes
- New primitives for efficient anonymous authentication
- Analyses of currently deployed zero-knowledge SNARKs

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, Daniel Kästner [AbsInt GmbH], Michael Schmidt [AbsInt GmbH], Bernhard Schommer [AbsInt GmbH].

In the context of our work on compiler verification (see section 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 ARM, PowerPC, RISC-V and x86 architectures [9]. This compiler comprises a back-end part, translating the Cminor intermediate language to PowerPC assembly and reusable for source languages other than C [8], and a front-end translating the CompCert C subset of C to Cminor. The compiler is mostly written within the specification language of the Coq proof assistant, from which Coq's extraction facility generates executable OCaml code. The compiler comes with a 100000-line machine-checked Coq proof of semantic preservation establishing that the generated assembly code executes exactly as prescribed by the semantics of the source C program.

This year, we improved the CompCert C compiler in several directions:

- A new built-in function, `__builtin_ais_annot` makes it easy to transfer annotations (also known as flow facts) written at the source code level in AbsInt's aiS annotation language all the way down to the level of the generated machine code. The aiT static analyzer for Worst-Case Execution Times, which operates at the machine code level, can then take advantage of these annotations to produce better WCET estimates.
- In preparation for a qualification with respect to industry standards for avionics software, conformance with the ISO C 1999 and ISO C 2011 standards was improved, with the addition of many diagnostics required by the standards.
- Performance of the generated code was slightly improved via changes to the heuristics for function inlining and for instruction selection.
- The semantic modeling of external function calls was made more precise, reflecting the fact that these functions can destroy some registers and some stack locations.

We released three versions of CompCert incorporating these improvements: version 3.2 in January 2018, version 3.3 in May 2018, and version 3.4 in September 2018.

Two papers on CompCert were presented at conferences. The first paper, with Daniel Kästner as lead author, was presented at the 2018 ERTS congress [22]. It describes the use of CompCert to compile software for nuclear power plant equipment developed by MTU Friedrichshafen, and the required certification of CompCert according to the IEC 60880 regulations for the nuclear industry. The second paper, with Bernhard Schommer as lead author, was presented at the 2018 WCET workshop [23]. It describes the `__builtin_ais_annot` source-level annotation mechanism mentioned above and its uses to help WCET analysis.

7.1.2. Verified code generation in the polyhedral model

Participants: Nathanaël Courant, Xavier Leroy.

The polyhedral model is a high-level intermediate representation for loop nests iterating over arrays and matrices, as found in numerical code. It supports a great many loop optimizations (fusion, splitting, interchange, blocking, etc) in a uniform, mathematically-elegant manner.

Nathanaël Courant, as part of his MPRI Master's internship and under Xavier Leroy's supervision, developed a Coq formalization of the polyhedral model. He then implemented and proved correct in Coq a code generator that produces efficient sequential code from an optimized polyhedral representation. Code generation is a delicate part of polyhedral compilation, involving complex, error-prone algorithms. Nathanaël Courant's verified code generator includes the major algorithms from Cédric Bastoul's reference paper [31]. The Coq specifications and proofs are available at https://github.com/Ekdohibs/PolyGen.

7.1.3. Testing compiler optimizations

Participant: Gergö Barany.

Compilers should be correct, but they should ideally also generate machine code that is as efficient as possible. Gergö Barany continued work on testing the quality of the generated code.

In a differential testing approach, one generates random C programs, compiles them with different compilers, then compares the generated code using a custom binary analysis tool. This tool finds missed optimizations by comparing criteria such as the number of instructions, the number of reads from the stack (for comparing the quality of register spilling), or the numbers of various other classes of instructions affected by optimizations of interest.

The system has found previously unreported missing optimizations in the GCC, Clang, and CompCert compilers. An article [19] was presented at the 27th International Conference on Compiler Construction (CC 2018), where it was honored with the Best Paper Award.

7.1.4. A verified model of register aliasing in CompCert

Participants: Gergö Barany, Xavier Leroy.

Some CPU architectures such as ARM feature register aliasing: Each of its 64-bit floating-point registers can also be accessed as two separate 32-bit halves. Modifying a superregister changes (invalidates) the data stored in subregisters and vice versa, but this behavior was not yet modeled in CompCert's semantics.

We continued work on re-engineering much of CompCert's semantic model of the register file and of the call stack. Rather than simple mappings of locations to values, the register file and the stack are now modeled more realistically as blocks of memory containing bytes that represent fragments of values. In this way, we can verify a semantic model in which a 64-bit register or stack slot may contain either a single 64-bit value or a pair of two unrelated 32-bit values. This ongoing work was presented at the workshop on Syntax and Semantics of Low-Level Languages (LOLA 2018) [25].

7.2. Language design and type systems

7.2.1. Refactoring with ornaments in ML

Participants: Thomas Williams, Lucas Baudin, Didier Rémy.

Thomas Williams, Lucas Baudin, and Didier Rémy have been working on refactoring and other transformations of ML programs based on mixed ornamentation and disornamentation. Ornaments have been introduced as a way of describing changes in data type definitions that can reorganize or add pieces of data. After a new data structure has been described as an ornament of an older one, the functions that operate on the bare structure can be partially or sometimes totally lifted into functions that operate on the ornamented structure.

Williams and Rémy improved the formalisation of the lifting framework: using ornament inference, an ML program is first elaborated 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 to an ML program. Williams and Rémy studied the semantics of this intermediate language and used it to prove the correctness of the lifting, using logical relations techniques. This work has been presented at POPL 2018 [12]. More technical details appear in a research report [43].

Lucas Baudin and Dider Rémy also studied the inverse transformation, disornamentation, which allows removing pieces of information from a data structure and adjusting the code accordingly. They showed that the framework of ornamentation can also be used to allow mixed ornamentation and disornamentation transformations. They also designed a new patch language to describe in a more robust manner how the code must be modified during such transformations. This enables a new class of applications, such as maintaining two views of a data structure in sync. For example, the location information in an abstract syntax tree, which is used to report error messages but obfuscates the code, can be projected away, leading to a simpler version of the code, which can then be modified and often automatically reornamented into the richer version of the code with locations. Disonamentation has been presented by Lucas Baudin at the ML 2018 workshop. Ornamentation, including mixed disornamentation, has also been presented at the MSFP 2018 workshop in Oxford.

A small prototype with ornamentation has been written by Thomas Williams and extended with disornamentation by Lucas Baudin. Thomas Williams has also started developing a new version of the prototype that will handle most of the OCaml language.

7.3. Shared-memory concurrency

7.3.1. The Linux Kernel Memory Model

Participants: Luc Maranget, Jade Alglave [University College London & ARM Ltd], Paul Mckenney [IBM Corporation], Andrea Parri [Sant'Anna School of Advanced Studies, Pisa, Italy], Alan Stern [Harvard University].

Modern multi-core and multi-processor computers do not follow the intuitive "sequential consistency" model that would define a concurrent execution as the interleaving of the executions of its constituent threads and that would command instantaneous writes to the shared memory. This situation is due both to in-core optimisations such as speculative and out-of-order execution of instructions, and to the presence of sophisticated (and cooperating) caching devices between processors and memory. Luc Maranget is taking part in an international research effort to define the semantics of the computers of the multi-core era, and more generally of shared-memory parallel devices or languages, with a clear initial focus on devices.

This year saw a publication on languages in an international conference. A multi-year effort to define a weak memory model for the Linux Kernel has yielded a scholarly paper [18] presented at the *Architectural Support for Programming Languages and Operating Systems* (ASPLOS) conference in March 2018. The article describes a formal model, the *Linux Kernel Memory Model* (LKMM), which defines how Linux kernel programs are supposed to behave. The model, a CAT model, can be simulated using the **herd** simulator, allowing programmers to experiment and develop intuitions. The model was tested against hardware and refined in consultation with Linux maintainers. Finally, the ASPLOS paper formalizes the *fundamental law of the Read-Copy-Update synchronization mechanism* and proves that one of its implementations satisfies this law. It is worth noting that the LKMM is now part of the Linux kernel source (in the `tools/`) section). Luc Maranget and his co-authors are the official maintainers of this document.

7.3.2. The ARMv8 and RISC-V memory model

Participants: Will Deacon [ARM Ltd], Luc Maranget, Jade Alglave [University College London & ARM Ltd].

Jade Alglave and Luc Maranget are working on a mixed-size version of the ARMv8 memory model. This model builds on the `aarch64.cat` model authored last year by Will Deacon (ARM Ltd). This ongoing work is subject to IP restrictions which we hope to lift next year.

Luc Maranget is an individual member of the memory model group of the RISC-V consortium (https://riscv.org/). Version V2.3 of the User-Level ISA Specification is now complete and should be released soon. This version features the first occurrence of a detailed memory model expressed in English, as well as its transliteration in CAT authored by Luc Maranget.

7.3.3. Work on diy

Participant: Luc Maranget.

This year, new synchronisation primitives were added to the Linux kernel memory model; ARMv8 atomic instructions were added; and more.

A more significant improvement is the introduction of *mixed-size* accesses. The tools can now handle a new view of memory, where memory is made up of elementary cells (typically *bytes*) that can be read or written as groups of contiguous cells (typically up to *quadwords* of 8 bytes). This preliminary work paves the way to the simulation of more elaborate memory models.

7.3.4. Unifying axiomatic and operational weak memory models

Participants: Jean-Marie Madiot, Jade Alglave [University College London & ARM Ltd], Simon Castellan [Imperial College London].

Modern multi-processors optimize the running speed of programs using a variety of techniques, including caching, instruction reordering, and branch speculation. While those techniques are perfectly invisible to sequential programs, such is not the case for concurrent programs that execute several threads and share memory: threads do not share at every point in time a single consistent view of memory. A *weak memory model* offers only weak consistency guarantees when reasoning about the permitted behaviors of a program. Until now, there have been two kinds of such models, based on different mathematical foundations: axiomatic models and operational models.

Axiomatic models explicitly represent the dependencies between the program and memory actions. These models are convenient for causal reasoning about programs. They are also well-suited to the simulation and testing of *hardware* microprocessors.

Operational models represent program states directly, thus can be used to reason on programs: program logics become applicable, and the reasoning behind nondeterministic behavior is much clearer. This makes them preferable for reasoning about *software*.

Jean-Marie Madiot has been collaborating with weak memory model expert Jade Alglave and concurrent game semantics researcher Simon Castellan in order to unify these styles, in a way that attempts to combine the best of both approaches. The first results are a formalisation of TSO-style architectures using partial-order techniques similar to the ones used in game semantics, and a proof of a stronger-than-state-of-art "data-race freedom" theorem: well-synchronised programs can assume a strong memory model. These results have been submitted for publication.

This is a first step towards tractable verification of concurrent programs, combining software verification using concurrent program logics, in the top layer, and hardware testing using weak memory models, in the bottom layer. Our hope is to leave no unverified gap between software and hardware, even (and especially) in the presence of concurrency.

7.3.5. Granularity control for parallel programs

Participants: Umut Acar, Vitaly Aksenov, Arthur Charguéraud, Adrien Guatto [Université Paris Diderot], Mike Rainey, Filip Sieczkowski [University of Wrocław].

This year, the DeepSea team continued their work on granularity control techniques for parallel programs.

A first line of research is based on the use of programmer-supplied asymptotic complexity functions, combined with runtime measurements. This work first appeared at PPoPP 2018 [16] in the form of a brief announcement, and was subsequently accepted for publication at PPoPP 2019 as a full paper.

A second line of research, known as *heartbeat scheduling*, is based on instrumenting the runtime system so that parallel function calls are initially executed as normal function calls, by pushing a frame on the stack, and subsequently can be promoted and become independent threads. This research has been presented at PLDI 2018 [14].

7.3.6. Theory and analysis of concurrent algorithms

Participant: Vitaly Aksenov.

Vitaly Aksenov, in collaboration with Petr Kuznetsov (Télécom ParisTech) and Anatoly Shalyto (ITMO University), proved that no wait-free linearizable implementation of a stack using read, write, compare & swap and fetch & add operations can be help-free. This proof corrects a mistake in an earlier proof by Censor-Hillel et al. The result was published at the the International Conference on Networked Systems (NETYS 2018) [17].

Vitaly Aksenov, in collaboration with Dan Alistarh (IST Austria) and Petr Kuznetsov (Télécom ParisTech), worked on performance prediction for coarse-grained locking. They describe a simple model that can be used to predict the throughput of coarse-grained lock-based algorithms. They show that their model works well for CLH locks, and thus can be expected to work for other popular lock designs such as TTAS or MCS. This work appeared as a brief announcement at PODC 2018 [16].

The aforementioned results by Vitaly Aksenov are also covered in his Ph.D. manuscript [11].

7.4. The OCaml language and system

7.4.1. The OCaml system

Participants: Damien Doligez, Armaël Guéneau, Xavier Leroy, Luc Maranget, David Allsop [University of Cambridge], Florian Angeletti, Frédéric Bour [Facebook], Stephen Dolan [University of Cambridge], Alain Frisch [Lexifi], Jacques Garrigue [University of Nagoya], Sébastien Hinderer, Nicolás Ojeda Bär [Lexifi], Thomas Refis [Jane Street], Gabriel Scherer [team Parsifal], Mark Shinwell [Jane Street], Leo White [Jane Street], Jeremy Yallop [University of Cambridge].

This year, we released three versions of the OCaml system: versions 4.06.1 and 4.07.1 are minor releases that fix 7 and 8 issues, respectively; version 4.07.0 is a major release that introduces many improvements in usability and performance, and fixes about 40 issues. The main novelties are:

- The standard library modules were reorganized to appear as sub-modules of a new `Stdlib` module. The purpose of this reorganization is to facilitate the addition of new standard library modules while minimize risks of conflicts with user modules of the same name.
- Modules `Float` (floating-point operations) and `Seq` (sequences) were added to the standard library, taking advantage of the new organization mentioned above.
- Since 4.01, it has been possible to select a variant constructor or record field from a sub-module that is not opened in the current scope, if type information is available at the point of use. This now also works for GADT constructors.
- The GC now handles the accumulation of custom blocks in the minor heap better. This solves some memory-usage issues observed in code which allocates a large amount of small custom blocks, typically small bigarrays.

7.4.2. Package management infrastructure

Participant: Damien Doligez.

This year, Damien Doligez has worked on the `opamcheck` tool, which is designed to check the compatibility of different versions of OCaml on the whole code base of `opam`, OCaml's package manager. As a by-product of this work, he has proposed numerous fixes to the `opam` package repository and to its dependency graph.

7.4.3. Work on the compiler's test suite and build system

Participant: Sébastien Hinderer.

In 2018, Sébastien Hinderer has worked on the OCaml compiler's test suite. More precisely, he has finished porting over 800 tests in the compiler's test suite so that they can be run by the tool `ocamltest`, developed by Sébastien earlier. To achieve this, it has been necessary to extend both `ocamltest` and the domain-specific language that is used to describe how tests should be executed.

In addition, Sébastien has fixed and properly documented the procedure that is used to bootstrap the OCaml compiler. Being able to compile the compiler using itself is an important feature: it is crucial, for instance, when the compiler is released. In addition to fixing the bootstrap procedure, Sébastien has introduced a way to test this procedure through continuous integration, which guarantees that it will not be broken again in the future.

Finally, Sébastien has continued to improve and refactor the compiler's build system, and, most importantly, has replaced the hand-written configuration script by an `autoconf`-generated one, which will be part of the upcoming 4.08 release of OCaml. This represents an important step towards the ability to produce cross-compilers for OCaml, which has been a long-standing issue for the whole OCaml community.

7.4.4. Optimizing OCaml for satisfiability problems

Participants: Sylvain Conchon [LRI, Univ. Paris-Saclay], Albin Coquereau [ENSTA-ParisTech], Mohamed Iguernlala [OCamlPro], Fabrice Le fessant [OCamlPro], Michel Mauny.

This work aims at improving the performance of the Alt-Ergo SMT solver, which is implemented in OCaml. For safety reasons, and to ease reasoning about its algorithms, the implementation of Alt-Ergo uses a functional programming style and persistent data structures, which are sometimes less efficient than imperative style and mutable data. Moreover, some efficient algorithms, such as CDCL SAT solvers, are naturally expressed in an imperative style.

Following our previous work on optimizing Alt-Ergo's built-in SAT solver, some efforts were needed to enable the comparison of our solver with other SMT solvers. We developed an OCaml library for parsing and type-checking SMT-LIB2. Since Alt-Ergo natively uses a polymorphic typing discipline, and since the community needs such advanced features, we proposed an extension of the SMT-LIB2 syntax where functions may be polymorphic.

The resulting new version of Alt-Ergo was presented at the 2018 SMT Workshop in Oxford [33]. Comparisons of Alt-Ergo with other SMT solvers, mainly developed in C++, took place during the competition that is associated with the workshop. They showed that Alt-Ergo's performance is similar to that of its competitors.

Albin Coquereau's Ph.D. defense is planned for Spring 2019.

7.4.5. Improvements in Menhir

Participant: François Pottier.

In 2018, the OCaml parser of the OCaml compiler was migrated from `ocamlyacc` to Menhir, at last. François Pottier took this opportunity to partially clean up the parser, reducing redundancy by taking advantage of Menhir's features. In the future, we hope to continue to work on the OCaml parser by improving the quality of its syntax error messages.

This cleanup work was also an occasion to revisit Menhir's grammar description language: François Pottier designed and implemented a new input syntax for Menhir, which seems slightly more powerful and elegant than the previous syntax.

7.5. Software specification and verification

7.5.1. Formal reasoning about asymptotic complexity

Participants: Armaël Guéneau, Arthur Charguéraud [team Camus], François Pottier.

For a couple years, Armaël Guéneau, Arthur Charguéraud, François Pottier have been investigating the use of Separation Logic, extended with Time Credits, as an approach to the formal verification of the time complexity of OCaml programs. In particular, Armaël has developed in Coq a theory and a set of tactics that allow working with asymptotic complexity bounds. He has presented the main aspects of this work at the conference ESOP 2018 [21]. Furthermore, a key part of the machinery for working with asymptotic complexity bounds has been released as a standalone, reusable Coq library, procrastination. Armaël presented this library at the Coq Workshop in July 2018 [29].

In 2018, Armaël has worked on a more ambitious case study, namely a recent incremental cycle detection algorithm, whose amortized complexity analysis is nontrivial. A machine-checked proof has been completed; a paper is in preparation.

7.5.2. Time Credits and Time Receipts in Iris

Participants: Glen Mével, Jacques-Henri Jourdan [CNRS], François Pottier.

From March to August 2018, Glen Mével did an M2 internship at Gallium, where he was co-advised by Jacques-Henri Jourdan (CNRS) and François Pottier. Glen extended the program logic Iris with time credits and time receipts.

Time credits are a well-understood concept, and have been used in several papers already by Armaël Guéneau, Arthur Charguéraud, and François Pottier. However, because Iris is implemented and proved sound inside Coq, extending Iris with time credits requires a nontrivial proof, which Glen carried out, based on a program transformation which inserts "tick" instructions into the code. As an application of time credits, Glen verified inside Iris the correctness of Okasaki's notion of "debits", which allows reasoning about the time complexity of programs that use thunks.

Time receipts are a new concept, which (we showed) allows proving that certain undesirable events, such as integer overflows, cannot occur until a very long time has elapsed. Glen extended Iris with time receipts and proved the soundness of this extension. As an application of time credits and receipts together, Jacques-Henri Jourdan updated Charguéraud and Pottier's earlier verification of the Union-Find data structure [3] and proved that integer ranks cannot realistically overflow, even if they are stored using only $\log W$ bits, where W is the number of bits in a machine word.

This work has been first submitted to POPL 2019, then (after significant revision) re-submitted to ESOP 2019.

7.5.3. Verified Interval Maps

Participant: François Pottier.

In the setting of ANR project Vocal, which aims to build a library of verified data structures for OCaml, François Pottier carried out a formal reconstruction of "interval maps". An interval map, a data structure proposed by Bonichon and Cuoq in 2010, represents a set of possible heaps, that is, a set of mappings of integer addresses to abstract values. Interval maps are used in the Frama-C program analysis tool. François Pottier re-implemented this data structure in Coq and carried out a formal verification of its main operations. This work, which represents about 4 months of work, remains unpublished at this time. It would be desirable to publish it and to envision its integration in Frama-C; this however requires further effort.

7.5.4. Chunked Sequences

Participants: Émilie Guermeur, Arthur Charguéraud, François Pottier.

In June and July 2018, Émilie Guermeur, an undergraduate student at Carnegie Mellon University (Pittsburgh, USA) did a 6-week internship, co-advised by Arthur Charguéraud and François Pottier. She wrote a full-fledged OCaml implementation of "chunked sequences", a data structure which offers an efficient representation of sequences of elements. This data structure exists in two forms, a persistent form and an ephemeral (mutable) form; efficient conversion operations are offered. François Pottier subsequently implemented a test harness, based on afl-fuzz, which allowed us to submit Émilie's code to intensive testing and detect and fix a few bugs. This work is not yet published; we intend to pursue it in 2019, to publish the library and perhaps to verify it.

7.5.5. TLA+

Participants: Damien Doligez, Leslie Lamport [Microsoft Research], Ioannis Filippidis, 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 [36]. This requires building tools to help write TLA+ specifications and mechanically check proofs.

Since October 2018, Ioannis Filippidis has been working on extending the TLAPS tool to deal with proofs of temporal properties. Under some well-defined circumstances, an occurrence of the ENABLED operator applied to a formula f can be replaced by a version of f where the primed variables are replaced by new existentially-quantified variables. The result is a first-order formula that can be sent to one of TLAPS's first-order back-ends. This rewriting of ENABLED suffices to prove a large class of liveness properties. Ioannis has started implementing this in TLAPS.

OURAGAN Team

7. New Results

7.1. On $SL(3,\mathbb{C})$-representations of the Whitehead link group

In [9], we describe a family of representations in $SL(3,\mathbb{C})$ of the fundamental group π of the Whitehead link complement. These representations are obtained by considering pairs of regular order three elements in $SL(3,\mathbb{C})$ and can be seen as factorising through a quotient of π defined by a certain exceptional Dehn surgery on the Whitehead link. Our main result is that these representations form an algebraic component of the $SL(3,\mathbb{C})$-character variety of π.

7.2. A simplified approach to rigorous degree 2 elimination in discrete logarithm algorithms

In [10], we revisit the ZigZag strategy of Granger, Kleinjung and Zumbrägel. In particular, we provide a new algorithm and proof for the so-called degree 2 elimination step. This allows us to provide a stronger theorem concerning discrete logarithm computations in small characteristic fields $F_{q^{k_0 k}}$ with k close to q and k_0 a small integer. As in the aforementioned paper, we rely on the existence of two polynomials h_0 and h_1 of degree 2 providing a convenient representation of the finite field $F_{q^{k_0 k}}$.

7.3. Computing Chebyshev knot diagrams

A Chebyshev curve $\mathcal{C}(a,b,c,\phi)$ has a parametrization of the form $x(t) = T_a(t); \quad y(t) = T_b(t); z(t) = T_c(t+\phi)$, where a, b, c are integers, $T_n(t)$ is the Chebyshev polynomial of degree n and $\phi \in \mathbb{R}$. When $\mathcal{C}(a,b,c,\phi)$ is nonsingular, it defines a polynomial knot. In [12], we determine all possible knot diagrams when ϕ varies. Let a, b, c be integers, a is odd, $(a,b) = 1$, we show that one can list all possible knots $\mathcal{C}(a,b,c,\phi)$ in $O(n^2)$ bit operations, with $n = abc$.

7.4. Programmable projective measurement with linear optics

In [8] present a scheme for a universal device which can be programmed by quantum states to perform a chosen projective measurement, and its implementation in linear optics. In particular, our scheme takes a single input system (the input register), and M-1 systems all in a state psi (the program registers), whose role is to encode the measurement direction, and approximates the projective measurement with respect to the state psi on the input system. Importantly the scheme is entirely independent of the measurement basis choice psi. This is done optimally in M, if we demand the input state psi always returns the appropriate outcome, and limits to the ideal projective measurement with M. The size of the linear optical circuit we propose scales as M log M , and requires O(M log M) classical side processing. Our scheme can be viewed as an extension of the swap test to the instance where one state is supplied many times.

7.5. Updating key size estimations for pairings

Recent progress on NFS imposed a new estimation of the security of pairings. In [6], we study the best attacks against some of the most popular pairings. It allows us to propose new pairing-friendly curves of 128 bits and 192 bits of security.

7.6. How to Securely Compute with Noisy Leakage in Quasilinear Complexity

Since their introduction in the late 90's, side-channel attacks have been considered as a major threat against cryptographic implementations. This threat has raised the need for formal leakage models in which the security of implementations can be proved. At Eurocrypt 2013, Prouff and Rivain introduced the noisy leakage model which has been argued to soundly capture the physical reality of power and electromagnetic leakages. In their work, they also provide the first formal security proof for a masking scheme in the noisy leakage model. However their work has two important limitations: (i) the security proof relies on the existence of a leak-free component, (ii) the tolerated amount of information in the leakage (aka leakage rate) is of $O(1/n)$ where n is the number of shares in the underlying masking scheme. The first limitation was nicely tackled by Duc, Dziembowski and Faust one year later (Eurocrypt 2014). Their main contribution was to show a security reduction from the noisy leakage model to the conceptually simpler random- probing model. They were then able to prove the security of the well-known Ishai-Sahai-Wagner scheme (Crypto 2003) in the noisy leakage model. The second limitation was addressed last year in a paper by Andrychowicz, Dziembowski and Faust (Eurocrypt 2016). The proposed construction achieves security in the strong adaptive probing model with a leakage rate of $O(1/logn)$ at the cost of a $O(n^2logn)$ complexity. we argue that their result can be translated into the noisy leakage model with a leakage rate of $O(1)$ by using secret sharing based on algebraic geometric codes. They further argue that the efficiency of their construction can be improved by a linear factor using packed secret sharing but no details are provided.

In [14], we show how to compute in the presence of noisy leakage with a leakage rate up to $\widetilde{O}(1)$ in complexity $\widetilde{O}(n)$. They use a polynomial encoding allowing quasilinear multiplication based on the fast Number Theoretic Transform (NTT). They first show that the scheme is secure in the random-probing model with leakage rate $O(1/logn)$. Using the reduction by Duc et al. this result can be translated in the noisy leakage model with a $O(1/|F|^2logn)$ leakage rate. However, as in the work of Andrychowicz et al. , our construction also requires $|F| = O(n)$. In order to bypass this issue, we refine the granularity of our computation by considering the noisy leakage model on logical instructions that work on constant-size machine words. we provide a generic security reduction from the noisy leakage model at the logical-instruction level to the random-probing model at the arithmetic level. This reduction allows to prove the security of the construction in the noisy leakage model with leakage rate $\widetilde{O}(1)$).

7.7. A New Public-Key Cryptosystem via Mersenne Numbers

In [13], we propose a new public-key cryptosystem whose security is based on the computational intractability of the following problem: Given a Mersenne number $p = 2^n - 1$ where n is a prime, a positive integer h , and two n -bit integers T,R , find two n -bit integers F,G each of Hamming weight at most h such that T = F ·R+G modulo p , under the promise that they exist.

7.8. Workspace, Joint space and Singularities of a family of Delta-Like Robot

In [11], we describe the workspace, the joint space and the singularities of a family of delta-like parallel robots by using algebraic tools. The different func- tions of SIROPA library are introduced, which is used to induce an estimation about the complexity in representing the singularities in the workspace and the joint space. A Gröbner based elimination is used to compute the singularities of the manipulator and a Cylindrical Algebraic Decomposition algorithm is used to study the workspace and the joint space. From these algebraic objects, they propose some certified three-dimensional plotting describing the shape of works- pace and of the joint space which will help the engineers or researchers to decide the most suited configuration of the manipulator they should use for a given task. Also, the different parameters associated with the complexity of the serial and parallel singularities are tabulated, which further enhance the selection of the different configuration of the manipulator by comparing the complexity of the singularity equations.

7.9. Certified Non-conservative Tests for the Structural Stability of Discrete Multidimensional Systems

In [7], we present new computer algebra based methods for testing the structural stability of n-D discrete linear systems (with $n \geq 2$). More precisely, they show that the standard characterization of the structural stability of a multivariate rational transfer function (namely, the denominator of the transfer function does not have solutions in the unit polydisc of $\mathbb{C}^n$) is equivalent to the fact that a certain system of polynomials does not have real solutions. We then use state-of-the-art computer algebra algorithms to check this last condition, and thus the structural stability of multidimensional systems.

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 modeling 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 continued work on our verified Lustre compiler. We developed a new semantic model for the modular reset feature provided by the Scade language and required for the compilation of hierarchical state machines. This work was presented at the SCOPES workshop in Germany in May [17]. Work continues on connecting this semantic model to the intermediate compilation target.

We completed work on generalizing the compiler to treat clocked arguments. This involved changes to our intermediate Obc language and the addition of a pass to add some (necessary) variable initializations in an efficient way. This work was accepted for presentation at the Journées Francophones des Langages Applicatifs in 2019.

6.2. Julia Subtyping Reconstructed

Participant: Francesco Zappa Nardelli.

Julia is a programming language recently designed at MIT to support the needs of the scientific community. Julia occupies a unique position in the design landscape, it is a dynamic language with no type system, yet it has a surprisingly rich set of types and type annotations used to specify multimethod dispatch. The types that can be expressed in function signatures include parametric union types, covariant tuple types, parametric user-defined types with single inheritance, invariant type application, and finally types and values can be reified to appear in signatures. With Vitek started a research project to study the design and the pragmatic use of the Julia language. At first we focused on the Julia subtyping algorithm. We studied the empirical evidence that users appeal to all the features provided by Julia and we report on a formalisation and implementation of the subtyping algorithm. This has been published in [15]. We are pursuing this line of research studying of the algorithm advances of Julia can be integrated into other programming languages.

6.3. Comparing Designs for Gradual Types

Participant: Francesco Zappa Nardelli.

The enduring popularity of dynamically typed languages has given rise to a cottage industry of static type systems, often called gradual type systems, that let developers annotate legacy code piecemeal. Type soundness for a program which mixes typed and untyped code does not ensure the absence of errors at runtime, rather it means that some errors will caught at type checking time, while other will be caught as the program executes. After a decade of research it is clear that the combination of mutable state, self references and subtyping presents interesting challenges to designers of gradual type systems. We have reviewed the state of the art in gradual typing for objects, and introduced a class-based object calculus with a static type system, dynamic method dispatch, transparent wrappers and dynamic class generation that we use to model key features of several gradual type systems by translation to it, and discuss the implications of the respective designs. This has been published in [18].

6.4. Fast and reliable unwinding via DWARF tables

Participants: Theophile Bastian, Francesco Zappa Nardelli.

DWARF is a widely-used debugging data format. DWARF is obviously relied upon by debuggers, but it plays an unexpected role in the runtime of high-level programming languages and in the implementation of program analysis tools. The debug information itself can be pervaded by subtle bugs, making the whole infrastructure unreliable. In this project we are investigating techniques and tools to perform validation and synthesis of the DWARF stack unwinding tables, to speedup DWARF-based unwinding, as well as exploring adventurous projects that can be built on top of reliable DWARF information.

At the time of writing, we have a tool that can validate DWARF unwind tables generated by mainstream compilers; the approach is effective, we found a problem in Clang table generation and several in GLIBC inline-assembly snippets. We also designed and implemented a tool that can synthesise DWARF unwind tables from binary that lacks them (e.g. because the compiler did not generate them - immediate applications: JITs assembly, inline assembly, ...). Additionally we have designed and implemented a ahead-of-time compiler of DWARF unwind tables to assembly, and an ad-hoc unwinder integrated with the defacto standard unwinder libuwind. It can speed up unwinding by a factor between 25x and 60x (depending on application), with a 2.5x size overhead for unwind information.

Discussion is in progress to get these tools included in mainstream tool (e.g. the GNU profiler Perf).

6.5. Sundials/ML: OCaml interface to Sundials Numeric Solvers

Participants: Timothy Bourke, Marc Pouzet.

This year we made major updates to the Sundials/ML OCaml interface to support v3.1.x of the Sundials Suite of numerical solvers.

This release adds support for the new generic matrix and linear solver interfaces. Major work was required to add these new modules, update the existing solver interfaces, and ensure backwards compatibility with Sundials to v2.7.0 (which is still the version installed by Debian stable). We also improved our treatment of integer types used in indexing, refactored the Dls and Sls matrix modules, improved our generation of performance stats (by adding confidence intervals), made the configure script more robust, and untangle the mass-solver and Jacobian interfaces of the ARKODE solver.

6.6. Zélus

Participants: Timothy Bourke, Marc Pouzet.

This year, we made a major revision of the language and compiler, called now the version 2. The language now deal with higher order functions. All the static analyses, type inference, causality inference and the initialization analysis has been extended. The code generation has also been improved, in particular the interface with the numeric solver. Several larger examples have been written.

A paper that present the overall approach followed in ZELUS has been published [12].

6.7. Deterministic Concurrency: A Clock-Synchronised Shared Memory Approach

Participant: Marc Pouzet.

Synchronous programming (SP) provides deterministic concurrency. So far, however, communication has been constrained to a set of primitive clock-synchronised shared memory (scm) data types, such as data-flow registers, streams and signals with restricted read and write accesses that limit modularity and behavioural abstractions. In the paper [23], we propose an extension to the SP theory which retains the advantages of deterministic concurrency, but allows communication to occur at higher levels of abstraction than currently supported by SP data types. Our approach is as follows. To avoid data races, each csm type publishes a policy interface for specifying the admissibility and precedence of its access methods. Each instance of the csm type has to be policy-coherent, meaning it must behave deterministically under its own policy—a natural requirement if the goal is to build deterministic sys- tems that use these types. In a policy-constructive system, all access methods can be scheduled in a policy-conformant way for all the types without deadlocking. In this paper, we show that a policy-constructive program exhibits deterministic concurrency in the sense that all policy-conformant interleavings produce the same input-output behaviour. Policies are conservative and support the csm types existing in current SP languages. This work is a follower of a old work we did in 2009, published at LCTES about scheduling policies.

6.8. Compiling synchronous languages for multi-processor implementations

Participants: Guillaume Iooss, Albert Cohen, Timothy Bourke, Marc Pouzet.

This work was performed with industrial partners in the context of the ASSUME project.

We have continued to improve our front-end tools for a use case provided by Airbus. This tool now generates three kinds of monolithic Lustre program, which are taken as an input of the Lopht tool (AOSTE team), which in turn generates an executable for the Kalray MPPA. In particular, one of the code generators is based on the hyper-period-expansion transformation, which unrolls the computation and generates a single step function running at the slowest period. This transformation allows managing the multi-periodic aspect of the application at the source level. Together with the work of the AOSTE team and Airbus, it allows us to execute the full application on a MPPA (TRL-5 Airbus certification level).

We have also improved the front-end tools for the use case provided by Safran. These tools ware integrated into the Heptagon compiler. Many improvements to the parser and many convenient program transformations (tuple and array destruction, equation clustering, ...) were implemented in the Heptagon compiler in order to treat this use case and enable the Lopht tool to extract the best performance. In particular, we have investigated the impact of inlining on the degree of parallelism exposed by the use-case application.

In addition to the work described above, we have defined a language extension for 1-synchronous clocks, strictly periodic clocks with a single activation. We show that we can derive a scheduling problem from the clock constraints in a program. However, solving these constraints by using an interesting cost functions (such as WCET load balancing across the different phases of a period) with an ILP does not scale for the two use cases. Thus, we used the fact that we do not need the optimal solution to fall back on heuristics, which finds a good solution within acceptable bounds. We have also investigated the effect of a non-determinism operator on the scheduling constraints, which gives extra freedom for choosing a schedule.

In collaboration (this year) with Dumitru Potop-Butucaru and Keryan Didier (Inria, AOSTE team); Jean Souyris and Vincent Bregeon (Airbus); Philippe Baufreton and Jean-Marie Courtelle (Safran).

In collaboration with ANSYS, a compilation technique has been designed for compiling SCADE to multi-core [24].

PI.R2 Project-Team

6. New Results

6.1. Effects in proof theory and programming

Participants: Hugo Herbelin, Yann Régis-Gianas, Alexis Saurin, Exequiel Rivas Gadda.

6.1.1. Interfaces for computational effects

Exequiel Rivas studied the relation between interfaces for computational effects in programming languages: arrows, idioms and monads. Building on previous results of Lindley, Yallop and Wadler, a categorical account was developed by means of monoidal adjunctions. This work was presented in MSFP 2018 [40] and later in SYCO I. Together with Ruben Pieters and Tom Schrijvers, a journal version of the article is currently being prepared that includes this work and previous work on non-monadic handlers. It will be submitted to the Journal of Functional Programming.

6.1.2. Monads with merging

In collaboration with Mauro Jaskelioff, Exequiel Rivas developed monads with merge-like operators. These operators are based on two well-known algebraic theories for concurrency: classic process algebras and the more recent concurrent monoids. This resulted in an article submitted to FoSSaCS.

6.1.3. Relative effects: coherence for skew structures

In joint work with Mauro Jaskelioff, Tarmo Uustalu and Niccolò Veltri, Exequiel Rivas developed coherence theorems in the setting of categories with skew structures: skew monoidal categories, skew near-rig categories, skew semigroup categories. These skew structures are motivated by the study of relative effects in programming languages, where the primary example are relative monads. The results are formalised in the programming language Agda. A journal article is currently being written.

6.1.4. Effectful proving

Hugo Herbelin started a program of reconstruction of different levels of computational strength of logic by means of translation to a core logic of polarised linear connectives.

6.1.5. On the computational strength of choice axioms

With the goal of transferring the effectful computational contents of the dependent choice to other forms of choice or bar induction axioms, Hugo Herbelin worked at clarifying the folklore regarding the strengths of various forms of choice and of bar induction.

In collaboration with Boban Velickovic, Alexis Saurin advised the LMFI master internship of Ikram Cherigui on classical realisability and forcing in set theory.

6.1.6. Effectful systems in Coq

In collaboration with Thomas Letan (Agence Nationale pour la Sécurité des Systèmes Informatiques), Pierre Chifflier (ANSSI) and Guillaume Hiet (Centrale Supélec), Yann Régis-Gianas developed a new approach to model and verify effectful systems in Coq. This work has been presented at FM 2018 [38].

6.2. Reasoning and programming with infinite data

Participants: Yann Régis-Gianas, Alexis Saurin, Abhishek De, Luc Pellissier, Xavier Onfroy.

This theme is part of the ANR project Rapido (see the National Initiatives section) which goes until end of september 2019.

6.2.1. Proof theory of infinitary and circular proofs

In collaboration with David Baelde, Amina Doumane, Guilhem Jaber and Denis Kuperberg, Alexis Saurin extended the proof theory of infinite and circular proofs for fixed-point logics in various directions by relaxing the validity condition necessary to distinguish sound proofs from invalid ones. The original validity condition considered by Baelde, Doumane and Saurin in CSL 2016 rules out lots of proofs which are computationally and semantically sound and does not account for the cut-axiom interaction in sequent proofs.

In the setting of sequent calculus, Saurin introduced together with Baelde, Doumane and Jaber a relaxed validity condition to allow infinite branches to be supported by threads bouncing on axioms and cuts. This allows for a much more flexible criterion, inspired from Girard's geometry of interaction. The most general form of this criterion does not ensure productivity due to a discrepancy between the sequential nature of proofs in sequent calculus and the parallel nature of threads. Several directions of research have therefore been investigated from that point:

- In sequent calculus, Baelde, Doumane and Saurin provided a slight restriction of the full bouncing validity which grants productivity and validity of the cut-elimination process. This restriction still strictly extends previous notions of validity and is actually expressive enough to be undecidable as proved together with Kuperberg. Decidability can be recovered by constraining the shapes of bounces. Doumane and Saurin were able in the fall 2018 to generalise the CSL proof technique to be applicable to bouncing threads. Those results are currently being written targetting a submission early 2019.
- In the setting of natural deduction, Saurin and Jaber introduced a validity criterion aiming at ensuring productivity of a circular λ-calculus with inductive and coinductive types.
- In the fall 2018, Abhishek De started his PhD under Saurin's supervision. The first part of his PhD work is dedicated to lifting the proof theory of circular and infinitary proofs to the setting of proof nets, in which the bouncing criterion will be much more convenient to work with since the discrepancy between sequent proofs and parallel threads will be dealt with.

6.2.2. Brotherston-Simpson's conjecture: Finitising circular proofs

An important and most active research topic on circular proofs is the comparison of circular proof systems with usual proof systems with induction and co-induction rules à la Park. This can be viewed as comparing the proof-theoretical power of usual induction reasoning with that of Fermat's infinite descent method. Berardi and Tatsuta, as well as Simpson, obtained in 2017 important results in this direction for logics with inductive predicates à la Martin-Löf. Those frameworks, however, are weaker than those of fixpoint logic which can express and mix least and greatest fixpoints by interleaving μ and ν statements. New results on this topics followed in 2018.

In a work with Nollet and Tasson, Saurin published in CSL 2018 a new validity condition which is quite straightfoward to check (it can be checked at the level of elementary cycles of the circular proofs, while the other criteria need to check a condition on every infinite branch) and still capture all circular proofs obtained from $\mu MALL$ finite proofs [46]. The condition for cycling in those proofs is more constrained than that of Baelde, Doumane and Saurin, but the proof contains more information which can be used to exctract inductive invariants. With this validity condition which can be useful for proof search for circular proofs, they obtained partial finitisation results and are currently aiming at solving the most general Brotherston-Simpson's conjecture.

6.2.3. Streams and classical logic

Luc Pellissier started a post-doc in december 2018 funded by the RAPIDO project and started working with Alexis Saurin on the stream interpretation of $\Lambda\mu$-calculi by investigating the connection between $\Lambda\mu$-calculus and the parsimonious λ-calculus.

6.2.4. Formalising circular proofs and their validity condition

During the spring and summer 2018, Saurin started with Xavier Onfroy a formalisation of circular proofs in Coq. Until now, Onfroy formalised parity-automata and their meta-theory as a first step to capture the decidability condition of circular proofs. Preliminary formalisations of circular proofs have been considered by Onfroy but shall still be pursued in order to fit into the picture.

6.3. Effective higher-dimensional algebra

Participants: Antoine Allioux, Pierre-Louis Curien, Eric Finster, Yves Guiraud, Cédric Ho Thanh, Matthieu Sozeau.

6.3.1. Rewriting methods in algebra

Yves Guiraud has written with Philippe Malbos (Univ. Lyon 1) 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 has appeared in Mathematical Structures in Computer Science [32].

Yves Guiraud has completed a four-year collaboration with Eric Hoffbeck (Univ. Paris 13) and Philippe Malbos (Univ. Lyon 1), whose aim was to develop a theory of rewriting in associative algebras, with a view towards applications in homological algebra. They adapted the known notion of polygraph [71] to higher-dimensional associative algebras, and used these objects to develop a rewriting theory on associative algebras that generalises the two major tools for computations in algebras: Gröbner bases [70] and Poincaré-Birkhoff-Witt bases [107]. Then, they transposed the construction of [14], based on an extension of Squier's theorem [110] in higher dimensions, to compute small polygraphic resolutions of associative algebras from convergent presentations. Finally, this construction has been related to the Koszul homological property, yielding necessary or sufficient conditions for an algebra to be Koszul. The resulting work will appear in Mathematische Zeitschrift [31].

Yves Guiraud has written his "Habilitation à diriger des recherches" manuscript, as a survey on rewriting methods in algebra based on Squier theory [13]. The defense is planned for Spring 2019.

Yves Guiraud works with Dimitri Ara (Univ. Aix-Marseille), Albert Burroni, Philippe Malbos (Univ. Lyon 1), François Métayer (Univ. Nanterre) and Samuel Mimram (École Polytechnique) on a reference book on the theory of polygraphs and higher-dimensional categories, and their applications in rewriting theory and homotopical algebra.

Yves Guiraud works with Marcelo Fiore (Univ. Cambridge) on the theoretical foundations of higher-dimensional algebra, in order to develop a common setting to develop rewriting methods for various algebraic structures at the same time. Practically, they aim at a definition of polygraphic resolutions of monoids in monoidal categories, based on the recent notion of n-oid in an n-oidal category. This theory will subsume the known cases of monoids and associative algebras, and encompass a wide range of objects, such as Lawvere theories (for term rewriting), operads (for Gröbner bases) or higher-order theories (for the λ-calculus).

Opetopes are a formalisation of higher many-to-one operations leading to one of the approaches for defining weak ω-categories. Opetopes were originally defined by Baez and Dolan. A reformulation (leading to a more carefully crafted definition) has been later provided by Batanin, Joyal, Kock and Mascari, based on the notion of polynomial functor. Pierre-Louis Curien, Cédric Ho Thanh and Samuel Mimram have developped (in several variants) a type-theoretical treatment of opetopes and finite opetopic sets, and have shown that the models of their type theory are indeed the opetopic sets as defined mathematically by the above authors. This work is being submitted to an international conference. Also, Cédric Ho Thanh has given a direct precise proof of the equivalence between many-to-one polygraphs and opetopic sets, thus establishing a connection with the theory of polygraphs [57].

6.3.2. Garside methods in algebra and rewriting

Building on [9], Yves Guiraud is currently finishing with Matthieu Picantin (Univ. Paris 7) a work that generalises already known constructions such as the bar resolution, several resolutions defined by Dehornoy and Lafont [79], and the main results of Gaussent, Guiraud and Malbos on coherent presentations of Artin monoids [10], to monoids with a Garside family. This allows an extension of the field of application of the rewriting methods to other geometrically interesting classes of monoids, such as the dual braid monoids.

Still with Matthieu Picantin, Yves Guiraud develops an improvement of the classical Knuth-Bendix completion procedure, called the KGB (for Knuth-Bendix-Garside) completion procedure. The original algorithm tries to compute, from an arbitrary terminating rewriting system, a finite convergent presentation, by adding relations to solve confluence issues. Unfortunately, this algorithm fails on standard examples, like most Artin monoids with their usual presentations. The KGB procedure uses the theory of Tietze transformations, together with Garside theory, to also add new generators to the presentation, trying to reach the convergent Garside presentation identified in [9]. The KGB completion procedure is partially implemented in the prototype Rewr, developed by Yves Guiraud and Samuel Mimram.

6.3.3. Foundations and formalisation of higher algebra

Antoine Allioux (PhD started in February), Eric Finster, Yves Guiraud and Matthieu Sozeau are exploring the development of higher algebra in type theory. To formalise higher algebra, one needs a new source of coherent structure in type theory. Finster has developed an internalisation of polynomial monads (of which opetopes and ∞-categories are instances) in type theory, which ought to provide such a coherent algebraic structure, inspired by the work of Kock et al [96]. Antoine Allioux is focusing on building an equivalence of types between categories seen as polynomial monads and the standard univalent categories in Homotopy Type Theory [22]. Another result that should follow is the ability to define simplicial types in Homotopy Type Theory, a long standing open problem in the field. An article on this subject is in preparation. Once armed with such a definition mechanism for higher algebraic structures and their algebras, it should be possible to internalise results from higher rewriting theory in type theory, which was the initial goal of this project.

6.3.4. Type Theory and Higher Topos Theory

Eric Finster explored the connections between intensional type theory and the theory of higher topoi, as developed in the works on Joyal and Lurie [103]. In particular, in collaboration with Mathieu Anel, André Joyal and Georg Biedermann, he gave a proof of a new result about the generation of left exact modalities in higher topoi, which has a corresponding internalisation in Homotopy Type Theory. Applications of this result to the Goodwillie Calculus, an advanced technique in abstract homotopy theory, resulted in the article [28].

6.4. Incrementality

Participants: Thibaut Girka, Yann Régis-Gianas.

6.4.1. Incrementality in proof languages

In collaboration with Paolo Giarrusso, Philipp Shuster 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 mechanised proof for this transformation as well as a prototype language featuring efficient derivatives for functional programs. A paper has been submitted to ESOP 2019.

In collaboration with Olivier Martinot (Paris Diderot), Yann Régis-Gianas studied a new technique to implement incrementalised operations on lists. A paper is to be submitted to ICFP 2019.

6.4.2. Difference languages

Kostia Chardonnet and Yann Régis-Gianas started the formalisation of difference languages for Java, using the framework developed by Thibaut Girka. In particular, Kostia Chardonnet implemented a mechanised small step operational semantics for a large subset of Java. A paper is in preparation.

6.5. Metatheory and development of Coq

Participants: Hugo Herbelin, Pierre Letouzey, Yann Régis-Gianas, Matthieu Sozeau, Gaëtan Gilbert, Cyprien Mangin, Théo Winterhalter, Théo Zimmermann, Thierry Martinez.

6.5.1. Homotopy type theory

Hugo Herbelin developed the syntax for a variant of Cohen, Coquand, Huber and Mörtberg's Cubical Type Theory where equality on types is defined to be equivalence of types, thus satisfying univalence by construction.

6.5.2. Proof irrelevance and Homotopy Type Theory

Gaëtan Gilbert (PhD student of N. Tabareau, Gallinette and M. Sozeau) continued developing the theory and implementation of *strict* propositions in the calculus of inductive constructions. In collaboration with Jesper Cockx (Chalmers), they developed this notion in full in an article at POPL 19 [30]. Strict propositions enjoy definitional proof-irrelevance and are compatible with both Univalence and Uniqueness of Identity Proofs, providing a foundation for further research in both directions: dealing with strict structures in homotopy type theory, and improving the support for programming with dependent types and proofs. They have shown in particular how to translate inductive types that can be seen as strict propositions into recursively defined types, providing a fix to the "singleton elimination" criterion used in Coq to treat the interaction of propositions (in Prop) and informative objects (in Type). Together with Pierre Letouzey, Matthieu Sozeau is pursuing an adaptation of the Prop sort informed by this new result. In particular, Pierre Letouzey is now experimenting with alternative ways to handle the accessibility arguments of Coq general fixpoints during extraction. Historically, the elimination of these arguments was a consequence of the accessibility inductive type being in Prop. But this can actually be seen as a more general dead-code elimination method. This leverages the need for accessibility to be in sort Prop, and hence opens new prospects concerning the Prop universe and the proof irrelevance.

6.5.3. Extensionality and Intensionality in Type Theory

Théo Winterhalter, Nicolas Tabareau and Matthieu Sozeau studied and formalised a complete translation from Extensional to Intensional Type Theory in Coq, now published at CPP 2019 [43]. They show that, contrary to the original paper proof of Oury, the target intensional type theory only needs to be extended with the Uniqueness of Identity Proofs principle and Functional Extensionality, settling concretely and formally a question that was studied semantically and up-to now only on paper by Hofmann and Altenkirch [61]. The translation was formalised using the Template-Coq framework and gives rise to an executable translation from partial terms of ETT into terms of Coq annotated with transports of equalities. This provides a simple way to justify the consistency of type theories extending the definitional equality relation by provable propositional equalities, and shows the equivalence of 2-level type theory [62] and the Homotopy Type System proposed by Voevodsky.

6.5.4. Dependent pattern-matching and recursion

Cyprien Mangin and Matthieu Sozeau have continued work on the Equations plugin of Coq, Equations now provides means to define nested, mutual and well-founded recursive definitions, together with a definitional compilation of dependent-pattern matching avoiding the use of axioms. In recent work, Matthieu Sozeau uncovered a new way to deal with dependent pattern-matching on inductive families avoiding more uses of the K axiom, inspired by the work of Cockx [74], that integrates well with the simplification engine developed for Equations. An article describing this work is in revision [58].

Thierry Martinez continued the implementation of a dependent pattern-matching compilation algorithm in Coq based on the PhD thesis work of Pierre Boutillier and on the internship work of Meven Bertrand. The algorithm based on small inversion and generalisation is the object of a paper to be submitted to the TYPES post-proceedings.

6.5.5. Explicit Cumulativity

Pierre Letouzey continued exploring with the help of Matthieu Sozeau a version of Coq's logic (CIC) where the cumulativity rule is 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 would 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. A prototype version of Coq is now available, but only a fragment of the standard library has been adapted to explicit cumulativity. In particular, the equalities of equalities currently need some amending, and this process is quite cumbersome.

6.5.6. Cumulativity for Inductive Types

Together with Amin Timany, Matthieu Sozeau developed the Calculus of Cumulative Inductive Constructions which extends the cumulativity relation of universes to universe polymorphic inductive types. This work was presented at FSCD 2018 [42]. The development of the model of this calculus suggested a refinement of the implementation which was integrated in Coq 8.8, providing a more flexible subtyping relation on inductive types in Coq. Notably, this work shrinks the gap to emulate the so-called "template" polymorphism of Coq with cumulative universe polymorphism. Cumulative Inductive Types also provide an apropriate basis to formalise the notions of small and large categories in type theory, avoiding the introduction of coercions. In particular, it provides a way to define a well-behaved category of types and functions and constructions on it, like the Yoneda embedding, which would not be expressible without cumulativity. Finally, Cumulative Inductive Types allow the definition of syntactic models of type theories with cumulativity inside Coq, as pioneered by Boulier *et al* [69].

6.5.7. Mathematical notations in Coq

Hugo Herbelin developed new extensions of the system of mathematical notation of Coq: support for autonomous auxiliary grammars, support for binders over arbitrary patterns, support for generic notations for applications.

6.5.8. Software engineering aspects of the development of Coq

Théo Zimmermann has studied software engineering and open collaboration aspects of the development of Coq.

Following the migration of the Coq bug tracker from Bugzilla to GitHub which he conducted in 2017, he analyzed data (extracted through the GitHub API), in collaboration with Annalí Casanueva Artís from the Paris School of Economics. The results show an increased number of bugs by core developers and an increased diversity of the people commenting bug reports. These results validate *a posteriori* the usefulness of such a switch. A paper [60] has been written and has been presented at the EAQSE workshop (without proceedings). The current objective is to publish the paper in the MSR 2019 conference.

Following discussions dating back from the end of 2017, he has founded the coq-community GitHub organisation in July 2018. This is a project for a collaborative, community-driven effort for the long-term maintenance and advertisement of Coq packages. Already 10 pre-existing Coq projects (plugins and libraries) have been moved to this organisation since then (seven of them are former Coq contribs that were fixed from time to time by the Coq developers themselves – mostly by Hugo Herbelin). The organisation also hosts a "manifesto" repository for general discussion, documentation and advice to developers (including already a few reusable templates for Coq projects), and a docker-coq project to provide reusable Docker images with Coq. The next objectives are to get started on the collaborative documentation (starting with a work by Pierre Castéran from LaBRI) and to create an editorial committee. Théo Zimmermann and Yann Régis-Gianas are preparing an article of the model proposed by the various existing *-community GitHub organisations (including the elm-community organisation from which coq-community was inspired, and ocaml-community which was influenced by coq-community itself).

In addition, Théo Zimmermann has coordinated efforts to improve the documentation of Coq, has documented the release process that he had put in place with Maxime Dénès, and has developed a GitHub / GitLab bot (in OCaml) that is used to automatise many useful functions for the Coq development (continuous integration and backporting of pull requests in particular). The goal is to make this bot modular and reusable for other projects.

6.5.9. Coordination of the development of Coq

The amount of contributions to the Coq system increased significantly in the recent years (around 50 pull-requests are reviewed, discussed and merged each month, approximately). Hugo Herbelin, Matthieu Sozeau and Théo Zimmermann, helped by members from Gallinette (Nantes) and Marelle (Sophia-Antipolis), devoted an important part of their time to coordinate the development, to review propositions of extensions of Coq from external and/or young contributors, and to propose themselves extensions (see the corresponding paragraphs).

6.6. Formalisation and verification

Participants: Pierre-Louis Curien, Kailiang Ji, Pierre Letouzey, Jean-Jacques Lévy, Cyprien Mangin, Daniel de Rauglaudre, Matthieu Sozeau.

6.6.1. Proofs and surfaces

Following ideas of J. Richter-Gebert, Pierre-Louis Curien, together with Jovana Obradović (former PhD student of the team and now postdoc in Prague), joined a project with Zoran Petrić and other Serbian colleagues on formalising proofs of incidence theorems (arising by repeated use of Menelaus theorem) by means of a cyclic sequent calculus, by which is meant that a (proof of a) sequent $\vdash \Gamma$ stands for the conjunction of all (proofs of) traditional sequents $\Gamma \smallsetminus \psi \vdash \psi$. We have designed a proof system, showed its soundness, and experimented it on an extended set of examples from elementary projective geometry. A paper is being written.

6.6.2. Hofstadter nested recursive functions and Coq

Pierre Letouzey continued this year the study of a family of nested recursive functions proposed by D. Hofstadter in his book “Gödel Escher Bach”. This is a generalisation of the earlier work [20], bringing a large number of new insights as well as many new conjectures. Most of the work is already certified in Coq, with generalised and/or nicer proofs, see https://www.irif.fr/~letouzey/hofstadter_g/. Many interactions with Fibonacci numbers or similar recursive sequence have been found. Pierre Letouzey even stumbled upon a Rauzy fractal during this investigation, which is still ongoing.

6.6.3. Real Numbers in Coq

The present Coq library of real numbers is made of 17 axioms. Daniel de Rauglaudre has been studying the possibility of making an implementation with one only axiom: the Limited Principle of Omniscience (LPO) which says that we can differentiate an infinite sequence of 0s from an infinite sequence holding something else than 0 (it seems obvious but it cannot be proved in constructive logic). This axiom had been already used in the formal proof of Puiseux’ theorem done some years ago (only axiom of this proof too).

Real numbers are defined by an infinite sequence of digits and the operations of addition and multiplication by algorithms using LPO.

It was tested in OCaml, the axiom being replaced by a function having a limit corresponding to the precision of the computation and it seems to work. But the proof in Coq that this implementation is a field stumbles on difficulties about the associativity of addition which is more complicated than expected. Several tracks have been experimented with Hugo Herbelin’s help.

6.6.4. Proofs of algorithms on graphs

Jean-Jacques Lévy and Chen Ran (a PhD student at the Institute of Software, Beijing) pursue their work about formal proofs of graph algorithms. Their goal is to provide proofs of algorithms checked by computer and 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 (Kosaraju - 1978 and Tarjan - 1972). Their proofs use the multi-sorted first-order logic with inductive predicates of the Why3 system (research-team Toccata, Saclay). They also widely use the numerous automatic provers interfaced with Why3. A very minor part of these proofs is also achieved in Coq. The difficulty of this approach is to combine automatic provers and the intuitive design. Another point is to define the good level of abstraction in order to avoid too many implementation features while keeping an effective presentation.

In 2017, the same proofs were fully completed in Coq-ssreflect with the Mathematical Components library by Cohen and Théry (research-team Marelle, Sophia-Antipolis), and in Isabelle-HOL by Merz (research-team VeriDis, Nancy), both proofs with the assistance of J.-J. Lévy. These proofs are between a factor 3 to 8 in length with respect to the initial Why3 proofs, but more importantly they look less human readable, mainly because of the absence of automatic deduction and several technicalities about termination. On the way, this collaboration led to a new, better presentation of the Why3 proof.

Part of this work (Tarjan 1972) was presented at JFLA 2017, a more comprehensive version was presented at the VSTTE 2017 conference in Heidelberg. Scripts of proofs can be found at http://jeanjacqueslevy.net/why3, where other proofs of graph algorithms are also present: acyclicity test, articulation points, biconnected components. A proof of Tarjan's planarity test is also under design. A paper entitled "Formal Proofs of Tarjan's Algorithm in Why3, Coq and Isabelle" is under submission to a conference.

6.6.5. Certified compilation and meta-programming

Matthieu Sozeau participates to the CertiCoq project (https://www.cs.princeton.edu/~appel/certicoq) whose aim is to verify a compiler from Coq's Gallina language down to CompCert C-light which provides itself a certified compilation path to assembly language. Matthieu Sozeau focused on the front-end part of CertiCoq, providing formal proofs of the first two phases of the compiler. The first phase translates from Coq syntax to a more amenable representation for metatheoretical study, and the second phase performs extraction to an untyped lambda-calculus with datatypes and mutual (co-)fixpoints. These two phases are of general use and are now integrated and developed in the MetaCoq project. The CertiCoq team expects to release a first version of the compiler in the beginning of 2019, along with an article describing it.

MetaCoq is a project led by Matthieu Sozeau, in collaboration with Simon Boulier and Nicolas Tabareau in Nantes, Abhishek Anand and Gregory Malecha (BedRock Systems, Inc) and Yannick Forster in Saarbrucken. The project was born from the extension of the Template-Coq reification plugin of G. Malecha, which now contains:

- A specification of the typing rules of Coq and its basic metatheoretical properties (weakening, substitution). This specification is not entirely complete yet, as the positivity and guard-checking of definitions is missing. Cyprien Mangin has formalised the regular tree structure used by the guard checker, and a simple positivity check for inductive types. Its integration is ongoing.
- A (partial) proof of the correctness and completeness of a reference type-checker with respect to these rules.
- An implementation of the extraction phase of Coq, which is used in the CertiCoq project. The proof of "syntactic" correctness of this phase, that is the preservation of weak call-by-value reduction by extraction is ongoing.
- A monad giving the ability to program arbitrary plugins in Coq itself, in the style of MTac.

. The foundation of this project was published at ITP 2018 [37], and a journal article is in preparation.

In collaboration with Jan-Oliver Kaiser (MPI-SWS), Beta Ziliani (CONICET/FAMAF), Robbert Krebbers (ICIS) and Derek Dreyer (MPI-SWS), Yann Régis-Gianas participates in the Mtac2 project, a metaprogramming language for Coq. The new version of this language has been presented at ICFP 2018 [34]. It includes in particular in a depedently-typed variant of the LCF tactic typing discipline.

In collaboration with Xavier Denis (Paris Diderot), Yann Régis-Gianas is implementing a compiler for Mtac2.

6.6.6. *Equivalences for free!*

Nicolas Tabareau (Inria Nantes), Eric Tanter (U. Chile in Santiago) and Matthieu Sozeau developed a new parametricity translation for justifying the transport of programs and proofs by equivalences in type theory [36]. Inspired by the Univalence axiom, they show that every construction of type theory (minus inductive families indexed by universes) respect type equivalence, and provide a modified parametricity translation that can be used to construct the proof of invariance by equivalence of any term. This translation is engineered so that transports do not appear during this inference, allowing an easy implementation of a transfer metaprogram in type theory using type class inference. Using this metaprogram, one can automatically transport libraries of implementations and their proofs from one type to an equivalent one, including cases where dependent types are used. While the translation ultimately relies on the univalence axiom to treat universes, its use can be avoided in many cases, providing an effective translation that can be evaluated inside type theory.

6.6.7. *Detecting K-Synchronisability Violations*

Ahmed Bouajjani, Constantin Enea, Kailiang Ji and Shaz Qadeer introduced a bounded analysis that explores a special type of computations, called k-synchronous, for analyzing message passing programs. They gave a procedure for deciding k-*synchronisability* of a program, i.e., whether every computation is equivalent (has the same happens-before relation) to one of its k-synchronous computations. They also showed that reachability over k-synchronous computations and checking k-synchronisability are both PSPACE-complete. Furthermore, they introduced a class of programs called *flow-bounded* for which the problem of deciding whether there exists a $k > 0$ for which the program is k-synchronisable, is decidable. The k-*synchronisability* violation detection algorithm was implemented in Spin model checker. This work was published at CAV 2018 [48].

POLSYS Project-Team

6. New Results

6.1. Fundamental algorithms and structured polynomial systems

6.1.1. Towards Mixed Gröbner Basis Algorithms: the Multihomogeneous and Sparse Case

One of the biggest open problems in computational algebra is the design of efficient algorithms for Gröbner basis computations that take into account the sparsity of the input polynomials. We can perform such computations in the case of unmixed polynomial systems, that is systems with polynomials having the same support, using the approach of Faugère, Spaenlehauer, and Svartz [ISSAC'14]. In [15] we present two algorithms for sparse Gröbner bases computations for mixed systems. The first one computes with mixed sparse systems and exploits the supports of the polynomials. Under regularity assumptions, it performs no reductions to zero. For mixed, square, and 0-dimensional multihomogeneous polynomial systems, we present a dedicated, and potentially more efficient, algorithm that exploits different algebraic properties that performs no reduction to zero. We give an explicit bound for the maximal degree appearing in the computations.

6.1.2. Bilinear Systems with Two Supports: Koszul Resultant Matrices, Eigenvalues, and Eigenvectors

A fundamental problem in computational algebraic geometry is the computation of the resultant. A central question is when and how to compute it as the determinant of a matrix whose elements are the coefficients of the input polynomials up-to sign. This problem is well understood for unmixed multihomogeneous systems, that is for systems consisting of multihomogeneous polynomials with the same support. However, little is known for mixed systems, that is for systems consisting of polynomials with different supports. In [14] we consider the computation of the multihomogeneous resultant of bilinear systems involving two different supports. We present a constructive approach that expresses the resultant as the exact determinant of a *Koszul resultant matrix*, that is a matrix constructed from maps in the Koszul complex. We exploit the resultant matrix to propose an algorithm to solve such systems. In the process we extend the classical eigenvalues and eigenvectors criterion to a more general setting. Our extension of the eigenvalues criterion applies to a general class of matrices, including the Sylvester-type and the Koszul-type ones.

6.1.3. A Polynomial-Division-Based Algorithm for Computing Linear Recurrence Relations

Sparse polynomial interpolation, sparse linear system solving or modular rational reconstruction are fundamental problems in Computer Algebra. They come down to computing linear recurrence relations of a sequence with the Berlekamp–Massey algorithm. Likewise, sparse multivariate polynomial interpolation and multidimensional cyclic code decoding require guessing linear recurrence relations of a multivariate sequence.

Several algorithms solve this problem. The so-called Berlekamp–Massey–Sakata algorithm (1988) uses polynomial additions and shifts by a monomial. The SCALAR-FGLM algorithm (2015) relies on linear algebra operations on a multi-Hankel matrix, a multivariate generalization of a Hankel matrix. The Artinian Gorenstein border basis algorithm (2017) uses a Gram-Schmidt process.

In [16], we propose a new algorithm for computing the Gröbner basis of the ideal of relations of a sequence based solely on multivariate polynomial arithmetic. This algorithm allows us to both revisit the Berlekamp–Massey–Sakata algorithm through the use of polynomial divisions and to completely revise the SCALAR-FGLM algorithm without linear algebra operations.

A key observation in the design of this algorithm is to work on the mirror of the truncated generating series allowing us to use polynomial arithmetic modulo a monomial ideal. It appears to have some similarities with Padé approximants of this mirror polynomial.

Finally, we give a partial solution to the transformation of this algorithm into an adaptive one.

As an addition from the paper published at the ISSAC conferance, in [24], we give an adaptive variant of this algorithm taking into account the shape of the final Gröbner basis gradually as it is discovered. The main advantage of this algorithm is that its complexity in terms of operations and sequence queries only depends on the output Gröbner basis.

All these algorithms have been implemented in MAPLE and we report on our comparisons.

6.1.4. In-depth comparison of the Berlekamp–Massey–Sakata and the Scalar-FGLM algorithms: the adaptive variants

The BERLEKAMP–MASSEY–SAKATA algorithm and the SCALAR-FGLM algorithm both compute the ideal of relations of a multidimensional linear recurrent sequence.

Whenever quering a single sequence element is prohibitive, the bottleneck of these algorithms becomes the computation of all the needed sequence terms. As such, having adaptive variants of these algorithms, reducing the number of sequence queries, becomes mandatory.

A native adaptive variant of the SCALAR-FGLM algorithm was presented by its authors, the so-called ADAPTIVE SCALAR-FGLM algorithm.

In [25], our first contribution is to make the BERLEKAMP–MASSEY–SAKATA algorithm more efficient by making it adaptive to avoid some useless relation testings. This variant allows us to divide by four in dimension 2 and by seven in dimension 3 the number of basic operations performed on some sequence family.

Then, we compare the two adaptive algorithms. We show that their behaviors differ in a way that it is not possible to tweak one of the algorithms in order to mimic exactly the behavior of the other. We detail precisely the differences and the similarities of both algorithms and conclude that in general the ADAPTIVE SCALAR-FGLM algorithm needs fewer queries and performs fewer basic operations than the ADAPTIVE BERLEKAMP–MASSEY–SAKATA algorithm.

We also show that these variants are always more efficient than the original algorithms.

6.1.5. Bit complexity for multi-homogeneous polynomial system solving Application to polynomial minimization

Multi-homogeneous polynomial systems arise in many applications. In [10] we provide bit complexity estimates for solving them which, up to a few extra other factors, are quadratic in the number of solutions and linear in the height of the input system under some genericity assumptions. The assumptions essentially imply that the Jacobian matrix of the system under study has maximal rank at the solution set and that this solution set if finite. The algorithm is probabilistic and a probability analysis is provided. Next, we apply these results to the problem of optimizing a linear map on the real trace of an algebraic set. Under some genericity assumptions, we provide bit complexity estimates for solving this polynomial minimization problem.

6.2. Solving Systems over the Reals and Applications

6.2.1. Univariate real root isolation in an extension field and applications

In [11] we present algorithmic, complexity and implementation results for the problem of isolating the real roots of a univariate polynomial in $B_\alpha \in L[y]$, where $L = \mathbb{Q}(\alpha)$ is a simple algebraic extension of the rational numbers. We revisit two approaches for the problem. In the first approach, using resultant computations, we perform a reduction to a polynomial with integer coefficients and we deduce a bound of $\widetilde{\mathcal{O}}_B(N^8)$ for isolating the real roots of B_α, where N is an upper bound on all the quantities (degree and bitsize) of the input polynomials. The bound becomes $\widetilde{\mathcal{O}}_B(N^7)$ if we use Pan's algorithm for isolating the real roots. In the second approach we isolate the real roots working directly on the polynomial of the input. We compute improved separation bounds for the roots and we prove that they are optimal, under mild assumptions. For isolating the real roots we consider a modified Sturm algorithm, and a modified version of `descartes`' algorithm. For the former we prove a Boolean complexity bound of $\widetilde{\mathcal{O}}_B(N^{12})$ and for the latter a bound of $\widetilde{\mathcal{O}}_B(N^5)$. We present aggregate separation bounds and complexity results for isolating the real roots of all polynomials B_{α_k}, when

α_k runs over all the real conjugates of α. We show that we can isolate the real roots of all polynomials in $\widetilde{\mathcal{O}}_B(N^5)$. Finally, we implemented the algorithms in C as part of the core library of MATHEMATICA and we illustrate their efficiency over various data sets.

6.2.2. On the Maximal Number of Real Embeddings of Spatial Minimally Rigid Graphs

The number of embeddings of minimally rigid graphs in $\mathbb{R}^D$ is (by definition) finite, modulo rigid transformations, for every generic choice of edge lengths. Even though various approaches have been proposed to compute it, the gap between upper and lower bounds is still enormous. Specific values and its asymptotic behavior are major and fascinating open problems in rigidity theory. Our work in [13] considers the maximal number of real embeddings of minimally rigid graphs in $\mathbb{R}^3$. We modify a commonly used parametric semi-algebraic formulation that exploits the Cayley-Menger determinant to minimize the *a priori* number of complex embeddings, where the parameters correspond to edge lengths. To cope with the huge dimension of the parameter space and find specializations of the parameters that maximize the number of real embeddings, we introduce a method based on coupler curves that makes the sampling feasible for spatial minimally rigid graphs. Our methodology results in the first full classification of the number of real embeddings of graphs with 7 vertices in $\mathbb{R}^3$, which was the smallest open case. Building on this and certain 8-vertex graphs, we improve the previously known general lower bound on the maximum number of real embeddings in $\mathbb{R}^3$.

6.2.3. Lower bounds on the number of realizations of rigid graphs

Computing the number of realizations of a minimally rigid graph is a notoriously difficult problem. Towards this goal, for graphs that are minimally rigid in the plane, we take advantage of a recently published algorithm, which is the fastest available method, although its complexity is still exponential. Combining computational results with the theory of constructing new rigid graphs by gluing, in [4] we give a new lower bound on the maximal possible number of (complex) realizations for graphs with a given number of vertices. We extend these ideas to rigid graphs in three dimensions and we derive similar lower bounds, by exploiting data from extensive Gröbner basis computations.

6.2.4. The Complexity of Subdivision for Diameter-Distance Tests

In [1] we present a general framework for analyzing the complexity of subdivision-based algorithms whose tests are based on the sizes of regions and their distance to certain sets (often varieties) intrinsic to the problem under study. We call such tests diameter-distance tests. We illustrate that diameter-distance tests are common in the literature by proving that many interval arithmetic-based tests are, in fact, diameter-distance tests. For this class of algorithms, we provide both non-adaptive bounds for the complexity, based on separation bounds, as well as adaptive bounds, by applying the framework of continuous amortization. Using this structure, we provide the first complexity analysis for the algorithm by Plantinga and Vegeter for approximating real implicit curves and surfaces. We present both adaptive and non-adaptive a priori worst-case bounds on the complexity of this algorithm both in terms of the number of subregions constructed and in terms of the bit complexity for the construction. Finally, we construct families of hypersurfaces to prove that our bounds are tight.

6.2.5. Real root finding for equivariant semi-algebraic systems

Let R be a real closed field. In [19] we consider basic semi-algebraic sets defined by n-variate equations/inequalities of s symmetric polynomials and an equivariant family of polynomials, all of them of degree bounded by $2d < n$. Such a semi-algebraic set is invariant by the action of the symmetric group. We show that such a set is either empty or it contains a point with at most $2d-1$ distinct coordinates. Combining this geometric result with efficient algorithms for real root finding (based on the critical point method), one can decide the emptiness of basic semi-algebraic sets defined by s polynomials of degree d in time $(sn)^{O(d)}$. This improves the state-of-the-art which is exponential in n. When the variables $x_1, ..., x_n$ are quantified and the coefficients of the input system depend on parameters $y_1, ..., y_t$, one also demonstrates that the corresponding one-block quantifier elimination problem can be solved in time $(sn)^{O(dt)}$.

6.2.6. Exact algorithms for semidefinite programs with degenerate feasible set

Let $A_0, ..., A_n n$ be $m \times m$ symmetric matrices with entries in Q, and let $A(x)$ be the linear pencil $A_0 + x_1 A_1 + \cdots + x_n A_n$, where $x = (x1, ..., xn)$ are unknowns. The linear matrix inequality (LMI) $A(x) \succeq 0$ defines the subset of R^n, called spectrahedron, containing all points x such that $A(x)$ has nonnegative eigenvalues. The minimization of linear functions over spectrahedra is called semidefinite programming (SDP). Such problems appear frequently in control theory and real algebra, especially in the context of nonnegativity certificates for multivariate polynomials based on sums of squares. Numerical software for solving SDP are mostly based on the interior point method, assuming some non-degeneracy properties such as the existence of interior points in the admissible set. In [21], we design an exact algorithm based on symbolic homotopy for solving semidefinite programs without assumptions on the feasible set, and we analyze its complexity. Because of the exactness of the output, it cannot compete with numerical routines in practice but we prove that solving such problems can be done in polynomial time if either n or m is fixed.

6.2.7. A lower bound on the positive semidefinite rank of convex bodies

The positive semidefinite rank of a convex body C is the size of its smallest positive semidef-inite formulation. In [3] we show that the positive semidefinite rank of any convex body C is at least $\sqrt{\log d}$ where d is the smallest degree of a polynomial that vanishes on the boundary of the polar of C. This improves on the existing bound which relies on results from quantifier elimination. Our proof relies on the Bézout bound applied to the Karush-Kuhn-Tucker conditions of optimality. We discuss the connection with the algebraic degree of semidefinite programming and show that the bound is tight (up to constant factor) for random spectrahedra of suitable dimension.

6.2.8. On the complexity of computing real radicals of polynomial systems

Let $f = (f_1, ..., f_s)$ be a sequence of polynomials in $Q[X_1, ..., X_n]$ of maximal degree D and $V \subset C^n$ be the algebraic set defined by f and r be its dimension. The real radical $\sqrt[re]{\langle f \rangle}$ associated to f is the largest ideal which defines the real trace of V. In [20] when V is smooth, we show that $\sqrt[re]{\langle f \rangle}$ has a finite set of generators with degrees bounded by V. Moreover, we present a probabilistic algorithm of complexity $(snDn)^{O(1)}$ to compute the minimal primes of $\sqrt[re]{\langle f \rangle}$. When V is not smooth, we give a probabilistic algorithm of complexity $s^{O(1)}(nD)^{O(nr2^r)}$ to compute rational parametrizations for all irreducible components of the real algebraic set $V \cap R^n$. Experiments are given to show the efficiency of our approaches.

6.2.9. Algorithms for Weighted Sums of Squares Decomposition of Non-negative Univariate Polynomials

It is well-known that every non-negative univariate real polynomial can be written as the sum of two polynomial squares with real coefficients. When one allows a weighted sum of finitely many squares instead of a sum of two squares, then one can choose all coefficients in the representation to lie in the field generated by the coefficients of the polynomial. In particular, this allows an effective treatment of polynomials with rational coefficients. In [9], we describe, analyze and compare both from the theoretical and practical points of view, two algorithms computing such a weighted sums of squares decomposition for univariate polynomials with rational coefficients. The first algorithm, due to the third author relies on real root isolation, quadratic approximations of positive polynomials and square-free decomposition but its complexity was not analyzed. We provide bit complexity estimates, both on the runtime and the output size of this algorithm. They are exponential in the degree of the input univariate polynomial and linear in the maximum bitsize of its complexity. This analysis is obtained using quantifier elimination and root isolation bounds. The second algorithm, due to Chevillard, Harrison, Joldes and Lauter, relies on complex root isolation and square-free decomposition and has been introduced for certifying positiveness of poly-nomials in the context of computer arithmetics. Again, its complexity was not analyzed. We provide bit complexity estimates, both on the runtime and the output size of this algorithm, which are polynomial in the degree of the input polynomial and linear in the maximum bitsize of its complexity. This analysis is obtained using Vieta's formula and root isolation bounds. Finally, we report on our implementations of both algorithms and compare them in practice on several

application benchmarks. While the second algorithm is, as expected from the complexity result, more efficient on most of examples, we exhibit families of non-negative polynomials for which the first algorithm is better.

6.2.10. On Exact Polya and Putinar's Representations

We consider the problem of finding exact sums of squares (SOS) decompositions for certain classes of non-negative multivariate polynomials, relying on semidefinite programming (SDP) solvers. In [18] we start by providing a hybrid numeric-symbolic algorithm computing exact rational SOS decompositions for polynomials lying in the interior of the SOS cone. It computes an approximate SOS decomposition for a perturbation of the input polynomial with an arbitrary-precision SDP solver. An exact SOS decomposition is obtained thanks to the perturbation terms. We prove that bit complexity estimates on output size and runtime are both polynomial in the degree of the input polynomial and simply exponential in the number of variables. Next, we apply this algorithm to compute exact Polya and Putinar's representations respectively for positive definite forms and positive polynomials over basic compact semi-algebraic sets. We also compare the implementation of our algorithms with existing methods in computer algebra including cylindrical algebraic decomposition and critical point method.

6.3. Solving Systems in Finite Fields, Applications in Cryptology and Algebraic Number Theory.

6.3.1. Linear Repairing Codes and Side-Channel Attacks

To strengthen the resistance of countermeasures based on secret sharing, several works have suggested to use the scheme introduced by Shamir in 1978, which proposes to use the evaluation of a random d-degree polynomial into $nd + 1$ public points to share the sensitive data. Applying the same principles used against the classical Boolean sharing, all these works have assumed that the most efficient attack strategy was to exploit the minimum number of shares required to rebuild the sensitive value; which is $d + 1$ if the reconstruction is made with Lagrange's interpolation. In [2], we highlight first an important difference between Boolean and Shamir's sharings which implies that, for some signal-to-noise ratio, it is more advantageous for the adversary to observe strictly more than d + 1 shares. We argue that this difference is related to the existence of so-called exact linear repairing codes, which themselves come with reconstruction formulae that need (much) less information (counted in bits) than Lagrange's interpolation. In particular, this result implies that, contrary to what was believed, the choice of the public points in Shamir's sharing has an impact on the countermeasure strength. As another contribution, we exhibit a positive impact of the existence of linear exact repairing schemes; we indeed propose to use them to improve the state-of-the-art multiplication algorithms dedicated to Shamir's sharing. We argue that the improvement can be effective when the multiplication operation in the base field is at least two times smaller than in its sub-fields.

6.3.2. On the Use of Independent Component Analysis to Denoise Side-Channel Measurements

Independent Component Analysis (ICA) is a powerful technique for blind source separation. It has been successfully applied to signal processing problems, such as feature extraction and noise reduction , in many different areas including medical signal processing and telecommunication. In [17], we propose a framework to apply ICA to denoise side-channel measurements and hence to reduce the complexity of key recovery attacks. Based on several case studies, we afterwards demonstrate the overwhelming advantages of ICA with respect to the commonly used preprocessing techniques such as the singular spectrum analysis. Mainly, we target a software masked implementation of an AES and a hardware unprotected one. Our results show a significant Signal-to-Noise Ratio (SNR) gain which translates into a gain in the number of traces needed for a successful side-channel attack. This states the ICA as an important new tool for the security assessment of cryptographic implementations.

PROSECCO Project-Team

7. New Results

7.1. Composition Theorems for CryptoVerif and Application to TLS 1.3

Participant: Bruno Blanchet.

We presented composition theorems for security protocols, to compose a key exchange protocol and a symmetric-key protocol that uses the exchanged key. Our results rely on the computational model of cryptography and are stated in the framework of the tool CryptoVerif. They support key exchange protocols that guarantee injective or non-injective authentication. They also allow random oracles shared between the composed protocols. To our knowledge, they are the first composition theorems for key exchange stated for a computational protocol verification tool, and also the first to allow such flexibility.

As a case study, we applied our composition theorems to a proof of TLS 1.3 Draft-18. This work fills a gap in our previous analysis of TLS 1.3 in CryptoVerif [52]. It appears in [31], [39].

7.2. Mechanised Cryptographic Proof of the WireGuard VPN Protocol

Participants: Benjamin Lipp, Bruno Blanchet, Karthikeyan Bhargavan.

WireGuard is a free and open source Virtual Private Network (VPN) that aims to replace IPsec and OpenVPN. It is based on a new cryptographic protocol derived from the Noise Protocol Framework. We provide the first mechanised cryptographic proof of the protocol underlying WireGuard, using the CryptoVerif proof assistant.

We analyse the entire WireGuard protocol as it is, including transport data messages, in an ACCE-style model. We contribute proofs for correctness, message secrecy, forward secrecy, mutual authentication, session uniqueness, and resistance against key compromise impersonation, identity mis-binding, and replay attacks. We also discusse the strength of the identity hiding provided by WireGuard.

Our work also provides novel theoretical contributions that are reusable beyond WireGuard. First, we extend CryptoVerif to account for the absence of public key validation in popular Diffie-Hellman groups like Curve25519, which is used in many modern protocols including WireGuard. To our knowledge, this is the first mechanised cryptographic proof for any protocol employing such a precise model. Second, we prove several indifferentiability lemmas that are useful to simplify the proofs for sequences of key derivations. This work is under submission.

7.3. Meta-F*: Proof automation with SMT, Tactics, and Metaprograms

Participants: Guido Martinez, Danel Ahman, Victor Dumitrescu, Nick Giannarakis [Princeton University], Chris Hawblitzel [Microsoft Research], Catalin Hritcu, Monal Narasimhamurthy [University of Colorado Boulder], Zoe Paraskevopoulou [Princeton University], Clément Pit-Claudel [MIT], Jonathan Protzenko [Microsoft Research], Tahina Ramananandro [Microsoft Research], Aseem Rastogi [Microsoft Research], Nikhil Swamy [Microsoft Research].

We introduced Meta-F* [69], a tactics and metaprogramming framework for the F* program verifier. The main novelty of Meta-F* is allowing to use tactics and metaprogramming to discharge assertions not solvable by SMT, or to just simplify them into well-behaved SMT fragments. Plus, Meta-F* can be used to generate verified code automatically.

Meta-F* is implemented as an F* effect, which, given the powerful effect system of F*, heavily increases code reuse and even enables the lightweight verification of metaprograms. Metaprograms can be either interpreted, or compiled to efficient native code that can be dynamically loaded into the F* type-checker and can interoperate with interpreted code. Evaluation on realistic case studies shows that Meta-F* provides substantial gains in proof development, efficiency, and robustness.

7.4. When Good Components Go Bad: Formally Secure Compilation Despite Dynamic Compromise

Participants: Carmine Abate, Arthur Azevedo de Amorim [CMU], Roberto Blanco, Ana Nora Evans [University of Virginia], Guglielmo Fachini [Nozomi Networks], Catalin Hritcu, Théo Laurent, Benjamin C. Pierce [University of Pennsylvania], Marco Stronati [Nomadic Labs], Andrew Tolmach [Portland State University].

We proposed a new formal criterion [47] for evaluating secure compilation schemes for unsafe languages, expressing end-to-end security guarantees for software components that may become compromised after encountering undefined behavior—for example, by accessing an array out of bounds.

Our criterion is the first to model dynamic compromise in a system of mutually distrustful components with clearly specified privileges. It articulates how each component should be protected from all the others—in particular, from components that have encountered undefined behavior and become compromised. Each component receives secure compilation guarantees—in particular, its internal invariants are protected from compromised components—up to the point when this component itself becomes compromised, after which we assume an attacker can take complete control and use this component's privileges to attack other components. More precisely, a secure compilation chain must ensure that a dynamically compromised component cannot break the safety properties of the system at the target level any more than an arbitrary attacker-controlled component (with the same interface and privileges, but without undefined behaviors) already could at the source level.

To illustrate the model, we construct a secure compilation chain for a small unsafe language with buffers, procedures, and components, targeting a simple abstract machine with built-in compartmentalization. We give a careful proof (mostly machine-checked in Coq) that this compiler satisfies our secure compilation criterion. Finally, we show that the protection guarantees offered by the compartmentalized abstract machine can be achieved at the machine-code level using either software fault isolation or a tag-based reference monitor.

7.5. The Meaning of Memory Safety

Participants: Arthur Azevedo de Amorim [CMU], Catalin Hritcu, Benjamin C. Pierce [University of Pennsylvania].

We give a rigorous characterization of what it means for a programming language to be memory safe [51], capturing the intuition that memory safety supports local reasoning about state. We formalize this principle in two ways. First, we show how a small memory-safe language validates a noninterference property: a program can neither affect nor be affected by unreachable parts of the state. Second, we extend separation logic, a proof system for heap-manipulating programs, with a memory-safe variant of its frame rule. The new rule is stronger because it applies even when parts of the program are buggy or malicious, but also weaker because it demands a stricter form of separation between parts of the program state. We also consider a number of pragmatically motivated variations on memory safety and the reasoning principles they support. As an application of our characterization, we evaluate the security of a previously proposed dynamic monitor for memory safety of heap-allocated data.

7.6. Recalling a Witness: Foundations and Applications of Monotonic State

Participants: Danel Ahman, Cédric Fournet [Microsoft Research], Catalin Hritcu, Kenji Maillard, Aseem Rastogi [Microsoft Research], Nikhil Swamy [Microsoft Research].

We provide a way to ease the verification of programs whose state evolves monotonically [48]. The main idea is that a property witnessed in a prior state can be soundly recalled in the current state, provided (1) state evolves according to a given preorder, and (2) the property is preserved by this preorder. In many scenarios, such monotonic reasoning yields concise modular proofs, saving the need for explicit program invariants. We distill our approach into the monotonic-state monad, a general yet compact interface for Hoare-style reasoning about monotonic state in a dependently typed language. We prove the soundness of the monotonic-state monad

and use it as a unified foundation for reasoning about monotonic state in the F* verification system. Based on this foundation, we build libraries for various mutable data structures like monotonic references and apply these libraries at scale to the verification of several distributed applications.

7.7. A Monadic Framework for Relational Verification: Applied to Information Security, Program Equivalence, and Optimizations

Participants: Niklas Grimm [Vienna University of Technology], Kenji Maillard, Cédric Fournet [Microsoft Research], Catalin Hritcu, Matteo Maffei [Vienna University of Technology], Jonathan Protzenko [Microsoft Research], Tahina Ramananandro [Microsoft Research], Aseem Rastogi [Microsoft Research], Nikhil Swamy [Microsoft Research], Santiago Zanella-Béguelin [Microsoft Research].

Relational properties describe multiple runs of one or more programs. They characterize many useful notions of security, program refinement, and equivalence for programs with diverse computational effects, and they have received much attention in the recent literature. Rather than developing separate tools for special classes of effects and relational properties, we advocate using a general purpose proof assistant as a unifying framework for the relational verification of effectful programs. The essence of our approach is to model effectful computations using monads and to prove relational properties on their monadic representations, making the most of existing support for reasoning about pure programs [67].

We apply this method in F* and evaluate it by encoding a variety of relational program analyses, including information flow control, program equivalence and refinement at higher order, correctness of program optimizations and game-based cryptographic security. By relying on SMT-based automation, unary weakest preconditions, user-defined effects, and monadic reification, we show that, compared to unary properties, verifying relational properties requires little additional effort from the F* programmer.

7.8. A Formal Treatment of Accountable Proxying over TLS

Participants: Karthikeyan Bhargavan, Ioana Boureanu [University of Surrey], Antoine Delignat-Lavaud [Microsoft Research], Pierre-Alain Fouque [University of Rennes], Cristina Onete [University of Limoges].

Much of Internet traffic nowadays passes through active proxies, whose role is to inspect, filter, cache, or transform data exchanged between two endpoints. To perform their tasks, such proxies modify channel-securing protocols, like TLS, resulting in serious vulnerabilities. Such problems are exacerbated by the fact that middleboxes are often invisible to one or both endpoints, leading to a lack of accountability. A recent protocol, called mcTLS, pioneered accountability for proxies, which are authorized by the endpoints and given limited read/write permissions to application traffic.

Unfortunately, we show that mcTLS is insecure: the protocol modifies the TLS protocol, exposing it to a new class of middlebox-confusion attacks. Such attacks went unnoticed mainly because mcTLS lacked a formal analysis and security proofs. Hence, our second contribution is to formalize the goal of accountable proxying over secure channels. Third, we propose a provably-secure alternative to soon-to-be-standardized mcTLS: a generic and modular protocol-design that carefully composes generic secure channel-establishment protocols, which we prove secure. Finally, we present a proof-of-concept implementation of our design, instantiated with unmodified TLS 1.3 draft 23, and evaluate its overheads [29].

7.9. hacspec: towards verifiable crypto standards

Participants: Karthikeyan Bhargavan, Franziskus Kiefer [Mozilla], Pierre-Yves Strub [Ecole Polytechnique].

We designed and published hacspec, a formal specification language for cryptographic primitives. Specifications (specs) written in hacspec are succinct, easy to read and implement, and lend themselves to formal verification using a variety of existing tools. The syntax of hacspec is similar to the pseudocode used in cryptographic standards but is equipped with a static type system and syntax checking tools that can find errors. Specs written in hacspec are executable and can hence be tested against test vectors taken from standards and specified in a common format. Finally, hacspec is designed to be compilable to other formal specification languages like F*, EasyCrypt, Coq, and cryptol, so that it can be used as the basis for formal proofs of functional correctness and cryptographic security using various verification frameworks.

We published a paper presenting the syntax, design, and tool architecture of hacspec. We demonstrated the use of the language to specify popular cryptographic algorithms, and developed preliminary compilers from hacspec to F* and to EasyCrypt. Our eventual goal is to invite authors of cryptographic standards to write their pseudocode in hacspec and to help the formal verification community develop the language and tools that are needed to promote high-assurance cryptographic sofware backed by mathematical proofs. All our code is released publicly on GitHub.

7.10. Largest-scale user study of secure messaging and API usage

Participants: Francesca Musiani [CNRS], Ksenia Ermoshina [CNRS], Harry Halpin, Iness Ben Guirat [INSAT].

As part of the NEXTLEAP EC project, we engaged in the largest ever user study of secure messaging applications, focusing on typical users as well as "high-risk" users in the Middle East and Ukraine, as well as developers.[41]. This work has been shared with standardization efforts such as the IETF Message Layer Security (MLS) effort in which Inria is participating, as well as W3C standardization of the W3C Web Authentication API. This work helped influence the formal verification of the privacy properties of hardware-based cryptographic authentication, which is a feature needed by many at risk users whose accounts are often the focus of hacks. This work has also led a fundamental inquiry into the social governance of standards and the role of formal verification in the future of standards.[42] As this work is highly interdisciplinary, it has featured collaboration with sociologists at CNRS and interns from INSAT in Tunisia, as well as a lecture series hosted at Centre Pompidou under the direction of Bernard Stiegler and Harry Halpin.

SECRET Project-Team

7. New Results

7.1. Symmetric cryptology

Participants: Xavier Bonnetain, Christina Boura, Anne Canteaut, Pascale Charpin, Daniel Coggia, Sébastien Duval, Gaëtan Leurent, María Naya Plasencia, Léo Perrin, Yann Rotella, André Schrottenloher, Ferdinand Sibleyras.

7.1.1. Block ciphers

Our recent results mainly concern either the analysis or the design of lightweight block ciphers.
Recent results:

- Nonlinear approximations of block ciphers: A. Canteaut, together with C. Beierle and G. Leander have exhibited the relationship between nonlinear invariants for block ciphers and nonlinear approximations. They have shown that, in some cases, the linear hull effect may be formalized in terms of nonlinear invariants. They have also introduced a new framework to study the probability of nonlinear approximations over iterated block ciphers [13], [26]
- Impossible differential cryptanalysis: C. Boura, V. Lallemand and M. Naya-Plasencia have introduced new techniques and complexity analyses for impossible differential cryptanalysis. They also showed that the technique of multiple differentials can be applied to impossible differential attacks [16]
- Construction of lightweight MDS matrices: S. Duval and G. Leurent have exhibited MDS matrices with the lowest known implementation cost. They have been constructed by a search through a space of circuits yielding MDS matrices [20], [11]

7.1.2. 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: 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 [17].
- Cryptanalysis of Goldreich pseudo-random generator: Goldreich's PRG is a theoretical construction which expands a short random string into a long pseudo-random string by applying a simple d-ary predicate to public random sized subsets of the bits of the seed. While the security of Goldreich's PRG has been thoroughly investigated, with a variety of results deriving provable security guarantees against classes of attacks in some parameter regimes and necessary criteria to be satisfied by the underlying predicate, little was known about its concrete security and efficiency. Motivated by the hope of getting practical instantiations of this construction, Y. Rotella and his co-authors initiated a study of the concrete security of Goldreich's PRG, and evaluated its resistance to cryptanalytic attacks. They developped a new guess-and-determine-style attack, and identified new criteria which captured the security guarantees [44].

7.1.3. Authenticated encryption

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 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. In this context, the members of the project-team have obtained some cryptanalytic results on several candidates to the CAESAR competition.

Recent results:

- State-recovery attack on a simplified version of Ketje Jr. [21], [34]
- Cryptanalysis of Morus, one of the finalists of the CAESAR competition [42]

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:

- Differential Equivalence of Sboxes: C. Boura, A. Canteaut and their co-authors have studied two notions of differential equivalence of Sboxes corresponding to the case when the functions have the same difference table, or when their difference tables have the same support [15], [25]. They proved that these two notions do not coincide, and that they are invariant under some classical equivalence relations like EA and CCZ equivalence. They also proposed an algorithm for determining the whole equivalence class of a given function.
- Boomerang Uniformity of Sboxes: The boomerang attack is a cryptanalysis technique against block ciphers which combines two differentials for the upper part and the lower part of the cipher. The Boomerang Connectivity Table (BCT) is a tool introduced by Cid *et al.* at Eurocrypt 2018 for analysing the dependency between these two differentials. C. Boura and A. Canteaut [14] have provided an in-depth analysis of BCT, by studying more closely differentially 4-uniform Sboxes. They have completely characterized the BCT of all differentially 4-uniform permutations of 4 bits and then study these objects for some cryptographically relevant families of Sboxes, as the inverse function and quadratic permutations. These two families are the first examples of differentially 4-uniform Sboxes optimal against boomerang attacks for an even number of variables, answering an open question raised by Cid *et al.*.
- CCZ equivalence of Sboxes: A. Canteaut and L. Perrin have characterized CCZ-equivalence as a property of the zeroes in the Walsh spectrum of an Sbox (or equivalently in their DDT). They used this framework to show how to efficiently upper bound the number of distinct EA-equivalence classes in a given CCZ-equivalence class. More importantly, they proved that CCZ-equivalence can be reduced to the association of EA-equivalence and an operation called twisting. They then revisited several results from the literature on CCZ-equivalence and showed how they can be interpreted in light of this new framework [18], [29]

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

- Links between linear and differential properties of Sboxes: P. Charpin together with J. Peng has established new links between the differential uniformity and the nonlinearity of some Sboxes in the case of two-valued functions and quadratic functions. More precisely, they have exhibited a lower bound on the nonlinearity of monomial permutations depending on their differential uniformity, as well as an upper bound in the case of differentially two-valued functions [19], [55]
- Construction of building-blocks with resistance against fault-attacks at a low implementation overhead [50].

7.1.5. Modes of operation and generic attacks

In order to use a block cipher in practice, and to achieve a given security notion, a mode of operation must be used on top of the block cipher. Modes of operation are usually studied through provable security, and we know that their use is secure as long as the underlying primitive is secure, and we respect some limits on the amount of data processed. The analysis of generic attack helps us understand what happens when the hypotheses of the security proofs do not hold, or the corresponding limits are not respected. Comparing proofs and attacks also shows gaps where our analysis is incomplete, and when improved proof or attacks are required.

Recent results:

- Use of block ciphers operating on small blocks with the CTR mode [53]: the security proof of the CTR mode requires that no more than $2^{n/2}$ blocks are encrypted with the same key, but the known attacks reveal very little information and are considered less problematic than on CBC. However, G. Leurent and F. Sibleyras have exhibited concrete attacks against the CTR mode when processing close to $2^{n/2}$ blocks of data, and have shown that an attacker can actually extract as much information as in the case of CBC encryption.
- Generic attacks against some MAC constructions based on block ciphers [52]: G. Leurent and F. Sibleyras, together with M. Nandi, have studied the security of several recent MAC constructions with provable security beyond the birthday bound, namely `SUM-ECBC`, `PMAC+`, `3kf9`, `GCM-SIV2`, and some variants. They described a new cryptanalysis technique for double-block MACs and they showed how to build a forgery attack with query complexity $\mathcal{O}(2^{3n/4})$, proving that these schemes do not reach full security in the information-theoretic model. Surprisingly, their attack on `LightMAC+` invalidates a recent security proof by Naito. Moreover, they gave the first attack against `SUM-ECBC` and `GCM-SIV2`, with complexity below 2^n.

7.2. Code-based cryptography

Participants: Rodolfo Canto Torres, Thomas Debris, Matthieu Lequesne, Nicolas Sendrier, Jean-Pierre Tillich, Valentin Vasseur.

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 structured codes,
- addressing new functionalities, like identity-based encryption, hashing or symmetric encryption.

Our recent work on code-based cryptography has to be seen in the context of the recently launched NIST competition whose purpose is to standardize quantum-safe public-key primitives. This call concerns all three major cryptographic primitives, namely public-key cryptosytems, key-exchange protocols and digital signature schemes. The most promising techniques today for addressing this issue are code-based cryptography, lattice-based cryptography, mutivariate cryptography, and hash-based cryptography.

Our contributions in this area are two-fold and consist in:

- designing and analysis new code-based solutions;
- cryptanalyzing code-based schemes, especially candidates to the NIST competition.

7.2.1. Design of new code-based solutions

The members of the project-team have submitted several candidates to the NIST competition, including a key-exchange protocol based on quasi-cyclic MDPC codes [41]. Their recent work on MDPC codes is important in this context in order to carefully analyze the properties of this candidate.

Recent results:

- Thwarting the GJS attack: the decryption algorithm of the QC-MDPC cryptosystem is based on an iterative bit-flipping algorithm, which fails with a small probability. These failures have been exploited in 2016 by Guo, Johansson and Stankovski to perform a key-recovery attack. JP Tillich recently analyzed how this attack can be avoided by increasing the key size of the scheme. Most notably, he proved that, under a very reasonable assumption, the error probability after decoding decays almost exponentially with the code-length with just two iterations of bit-flipping. With an additional assumption, it even decays exponentially with an unbounded number of iterations, implying that in this case the increase of the key size equired for resisting to the GJS attack is only moderate [54].
- Design of a new KEM with IND-CCA2 security in a model considering decoding failures [46]: M. Lequesne, N. Sendrier and their co-authors explored the underlying causes of the GJS attack, how it can be improved and how it can be mitigated. They derived a new timing attack performing well even in cases which were more challenging to the GJS attack. They also showed how to construct a new KEM, called ParQ that can reduce the decryption failure rate to a level negligible in the security parameter. They formally proved the IND-CCA2 security of ParQ, in a model that considers decoding failures.
- Design of a new code-based signature scheme [81]: T. Debris, N. Sendrier and JP Tillich recently proposed a "hash-and-sign" code-based signature scheme called `Wave`, which uses a family of ternary generalized (U, U + V) codes. `Wave` achieves existential unforgeability under adaptive-chosen-message attacks in the random oracle model with a tight reduction to two assumptions from coding theory: one is a distinguishing problem that is related to the trapdoor inserted in the scheme, the other one is a multiple-target version of syndrome decoding. This scheme enjoys efficient signature and verification algorithms. For 128-bit security, signature are 8000-bit long and the public-key size is slightly smaller than one megabyte.

7.2.2. Cryptanalysis of code-based schemes

Recent results:

- Cryptanalysis of two public-key cryptosystems based on the rank syndrome decoding problem [41]: JP Tillich and his co-authors proposed an attack on the Rank Syndrome Decoding problem which improves the previously best known algorithm for solving this problem. This attack breaks for some parameters some recently proposed cryptosystems based on LRPC codes or Gabidulin codes, including Loidreau's cryptosystem and the LRPC cryptosystem.
- Cryptanalysis of the NIST submission `RankSign` and of a recently proposed IBE scheme: T. Debris and JP Tillich have presented an algebraic attack against `RankSign` that exploits the fact that the augmented LRPC codes used in this scheme have codewords with a very low weight. This attack shows that all the parameters proposed for this candidate can be broken. They also proved that, for the IBE scheme based on `RankSign`, the problem is deeper than finding a new signature in rank-based cryptography, since they found an attack on the generic problem upon which the security reduction relies [45].

- Cryptanalysis of the EDON-K key encapsulation mechanism submitted to the NIST competition: EDON-K is a candidate to the NIST competition which is inspired by the McEliece scheme but uses another family of codes defined over $\mathbb{F}_{2^{128}}$ instead of $\mathbb{F}_2$ and is not based on the Hamming metric. M. Lequesne and JP Tillich presented an attack making the scheme insecure for the intended use. Indeed, recovering the error in the McEliece scheme corresponding to EDON-K can be viewed as a decoding problem for the rank-metric with a super-code of an LRPC code of very small rank A suitable parity-check matrix for this super-code can then be easily derived from the public key and used to recover the error [51].
- Attack against RLCE [80]: M. Lequesne and JP Tillich, together with A. Couvreur, recently presented a key-recovery attack against the Random Linear Code Encryption (RLCE) scheme recently submitted by Y. Wang to the NIST competition. This attack recovers the secret-key for all the short key-parameters proposed by the author. It uses a polynomial-time algorithm based on a square code distinguisher.

7.3. Quantum Information

Participants: Xavier Bonnetain, Rémi Bricout, André Chailloux, Shouvik Ghorai, Antoine Grospellier, Anirudh Krishna, Anthony Leverrier, Vivien Londe, María Naya Plasencia, Andrea Olivo, Jean-Pierre Tillich, André Schrottenloher.

Our research in quantum information focusses on several axes: quantum codes with the goal of developing better error correction strategies to build large quantum computers, quantum cryptography which exploits the laws of quantum mechanics to derive security guarantees, relativistic cryptography which exploits in addition the fact that no information can travel faster than the speed of light and finally quantum cryptanalysis which investigates how quantum computers could be harnessed to attack classical cryptosystems.

7.3.1. Quantum codes

Protecting quantum information from external noise is an issue of paramount importance for building a quantum computer. It also worthwhile to notice that all quantum error-correcting code schemes proposed up to now suffer from the very same problem that the first (classical) error-correcting codes had: there are constructions of good quantum codes, but for the best of them it is not known how to decode them in polynomial time.

Two PhD students within the project-team work on this topic. First, Antoine Grospellier, co-advised by A. Leverrier and O. Fawzi (Ens Lyon), studies 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:

- Decoding algorithm for quantum expander codes [48], [47], [56] In this work, A. Grospellier, A. Leverrier and O. Fawzi analyze an efficient decoding algorithm for quantum expander codes and prove that it can correct a linear number of random errors with a negligible failure probability. As an application, this shows that this family of codes can be used to obtain quantum fault-tolerance with only a constant overhead in terms of qubits, compared to a polylogarithmic overhead as in previous schemes. This is a crucial step in order to eventually build large universal quantum computers.

7.3.2. Quantum cryptography

Quantum cryptography exploits the laws of quantum physics to establish the security of certain cryptographic primitives. The most studied one is certainly quantum key distribution, which allows two distant parties to establish a secret using an untrusted quantum channel. Our activity in this field is particularly focussed on protocols with continuous variables, which are well-suited to implementations. The interest of continuous variables for quantum cryptography was recently recognized by being awarded a 10 M€ funding from the Quantum Flagship and SECRET will contribute to this project by studying the security of new key distribution protocols [88].

Recent results:

- Security proof for two-way continuous-variable quantum key distribution [22]: while many quantum key distribution protocols are one-way in the sense that quantum information is sent from one party to the other, it can be beneficial in terms of performance to consider two-way protocols where the quantum states perform a round-trip between the two parties. In this paper (to appear in *Physical Review A*), we show how to exploit the symmetries of the protocols in phase-space to establish their security against the most general attacks allowed by quantum theory.
- Investigating the optimality of ancilla-assisted linear optical Bell measurements [24]: Due to its experimental and theoretical simplicity, linear quantum optics has proved to be a promising route for the early implementation of important quantum communication protocols. A. Olivo and F. Grosshans study the efficiency of non ambiguous Bell measurements in this model and show both theoretical and numerical bounds depending on the number of ancilla qubits. This is important for understanding what resources are needed for building quantum repeaters, the last missing building block for secure long distance quantum key distribution networks.

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 worked on this topic for several years and Andrea Olivo was recruited as a PhD student to continue working on both theoretical and practical aspects of relativistic cryptography.

Recent results:

- Relativistic commitment and zero-knowledge proofs [30]: A. Chailloux and A. Leverrier constructed a relativistic zero-knowledge protocol for any NP-complete problem. The main technical tool is the analysis of quantum consecutive measurements, which allows us to prove security against quantum adversaries. R. Bricout and A. Chailloux also studied relativistic multi-round bit commitment schemes. They showed optimal classical cheating strategies for the canonical F_Q commitment scheme.

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 seemed for a long time Grover's algorithm, which allows for an exhaustive key search in the square root of the normal complexity. For this reason, it was usually believed that doubling key lengths suffices to maintain an equivalent security in the post-quantum world. However, a lot of work is certainly required in the field of symmetric cryptography in order to "quantize" the classical families of attacks in an optimized way, as well as to find new dedicated quantum attacks. M. Naya Plasencia has recently been awarded an ERC Starting grant for her project named QUASYModo on this topic.

Recent results:

- Hidden-shift quantum cryptanalysis [43]: X. Bonnetain and M. Naya-Plasencia have obtained new results that consider the tweak proposed at Eurocrypt 2017 of using modular additions to counter Simon's attacks. They have developed new algorithms that improve and generalize Kuperberg's algorithm for the hidden shift problem. Thanks to their improved algorithm, they have been able to build a quantum attack in the superposition model on Poly1305, proposed at FSE 2005, largely used and claimed to be quantumly secure. They also analyzed the security of some classical symmetric constructions with concrete parameters, to evaluate the impact and practicality of the proposed tweak, concluding that it does not seem to be efficient
- Quantum algorithm for the k-XOR problem [49]: The k-XOR (or generalized birthday) problem aims at finding k elements of n-bits, drawn at random, such that the XOR of all of them is 0. The algorithms proposed by Wagner more than 15 years ago remain the best known classical algorithms for solving it, when disregarding logarithmic factors. M. Naya-Plasencia and A. Schrottenloher, together with L. Grassi, studied this problem in the quantum setting and provided algorithms with the best known quantum time-complexities. In particular, they were able to considerably improve the 3-XOR algorithm.
- Quantum cryptanalysis of CSIDH and Ordinary Isogeny-based Schemes [68]: CSIDH is a recent proposal by Castryck et al. for post-quantum non-interactive key-exchange. It is similar in design to a scheme by Couveignes, Rostovtsev and Stolbunov, but it replaces ordinary elliptic curves by supersingular elliptic curves. Although CSIDH uses supersingular curves, it can attacked by a quantum subexponential hidden shift algorithm due to Childs et al. While the designers of CSIDH claimed that the parameters they suggested ensures security against this algorithm, X. Bonnetain and A. Schrottenloher showed that these security parameters were too optimistic: they improved the hidden shift algorithm and gave a precise complexity analysis in this context, which greatly reduced the complexity. For example, they showed that only 2^{35} quantum equivalents of a key-exchange are sufficient to break the 128-bit classical, 64-bit quantum security parameters proposed, instead of 2^{62}. They also extended their analysis to ordinary isogeny computations, and showed that an instance proposed by De Feo, Kieffer and Smith and expected to offer 56 bits of quantum security can be broken in 2^{38} quantum evaluations of a key exchange.

CAGE Project-Team

6. New Results

6.1. Geometry of vision and sub-Riemannian geometry: new results

Let us list here our new results in the geometry of vision axis and, more generally, on hypoelliptic diffusion and sub-Riemannian geometry.

- In [7] we present a new image inpainting algorithm, the Averaging and Hypoelliptic Evolution (AHE) algorithm, inspired by the one presented in [86] and based upon a (semi-discrete) variation of the Citti–Petitot–Sarti model of the primary visual cortex V1. In particular, we focus on reconstructing highly corrupted images (i.e. where more than the 80% of the image is missing).
- In [6] we deal with a severe ill posed problem, namely the reconstruction process of an image during tomography acquisition with (very) few views. We present different methods that we have been investigated during the past decade. They are based on variational analysis.
- [13] is the first paper of a series in which we plan to study spectral asymptotics for sub-Riemannian Laplacians and to extend results that are classical in the Riemannian case concerning Weyl measures, quantum limits, quantum ergodicity, quasi-modes, trace formulae. Even if hypoelliptic operators have been well studied from the point of view of PDEs, global geometrical and dynamical aspects have not been the subject of much attention. As we will see, already in the simplest case, the statements of the results in the sub-Riemannian setting are quite different from those in the Riemannian one. Let us consider a sub-Riemannian (sR) metric on a closed three-dimensional manifold with an oriented contact distribution. There exists a privileged choice of the contact form, with an associated Reeb vector field and a canonical volume form that coincides with the Popp measure. We establish a Quantum Ergodicity (QE) theorem for the eigenfunctions of any associated sR Laplacian under the assumption that the Reeb flow is ergodic. The limit measure is given by the normalized Popp measure.This is the first time that such a result is established for a hypoelliptic operator, whereas the usual Shnirelman theorem yields QE for the Laplace-Beltrami operator on a closed Riemannian manifold with ergodic geodesic flow. To prove our theorem, we first establish a microlocal Weyl law, which allows us to identify the limit measure and to prove the microlocal concentration of the eigenfunctions on the characteristic manifold of the sR Laplacian. Then, we derive a Birkhoff normal form along this characteristic manifold, thus showing that, in some sense, all 3D contact structures are microlocally equivalent. The quantum version of this normal form provides a useful microlocal factorization of the sR Laplacian. Using the normal form, the factorization and the ergodicity assumption, we finally establish a variance estimate, from which QE follows. We also obtain a second result, which is valid without any ergodicity assumption: every Quantum Limit (QL) can be decomposed in a sum of two mutually singular measures: the first measure is supported on the unit cotangent bundle and is invariant under the sR geodesic flow, and the second measure is supported on the characteristic manifold of the sR Laplacian and is invariant under the lift of the Reeb flow. Moreover, we prove that the first measure is zero for most QLs.
- In [22] we study the validity of the Whitney C^1 extension property for horizontal curves in sub-Riemannian manifolds endowed with 1-jets that satisfy a first-order Taylor expansion compatibility condition. We first consider the equiregular case, where we show that the extension property holds true whenever a suitable non-singularity property holds for the input-output maps on the Carnot groups obtained by nilpotent approximation. We then discuss the case of sub-Riemannian manifolds with singular points and we show that all step-2 manifolds satisfy the C^1 extension property. We conclude by showing that the C^1 extension property implies a Lusin-like approximation theorem for horizontal curves on sub-Riemannian manifolds.

- In [34] we prove the C^1 regularity for a class of abnormal length-minimizers in rank 2 sub-Riemannian structures. As a consequence of our result, all length-minimizers for rank 2 sub-Riemannian structures of step up to 4 are of class C^1.
- In [45] we address the double bubble problem for the anisotropic Grushin perimeter P_α, $\alpha \geq 0$, and the Lebesgue measure in $\mathbb{R}^2$, in the case of two equal volumes. We assume that the contact interface between the bubbles lays on either the vertical or the horizontal axis. Since no regularity theory is available in this setting, in both cases we first prove existence of minimizers via the direct method by symmetrization arguments and then characterize them in terms of the given area by first variation techniques. Angles at which minimal boundaries intersect satisfy the standard 120-degree rule up to a suitable change of coordinates. While for $\alpha = 0$ the Grushin perimeter reduces to the Euclidean one and both minimizers coincide with the symmetric double bubble found in [104], for $\alpha = 1$ vertical interface minimizers have Grushin perimeter strictly greater than horizontal interface minimizers. As the latter ones are obtained by translating and dilating the Grushin isoperimetric set found in [131], we conjecture that they solve the double bubble problem with no assumptions on the contact interface.
- In [51] we study the notion of geodesic curvature of smooth horizontal curves parametrized by arc-length in the Heisenberg group, that is the simplest sub-Riemannian structure. Our goal is to give a metric interpretation of this notion of geodesic curvature as the first corrective term in the Taylor expansion of the distance between two close points of the curve.

We would also like to mention the defense of the PhD thesis of Ludovic Saccchelli [3] on the subject.

6.2. Quantum control: new results

Let us list here our new results in quantum control theory.

- In [5] we consider a quantum particle in a potential $V(x)$ ($x \in \mathbb{R}^N$) in a time-dependent electric field $E(t)$ (the control). Boscain, Caponigro, Chambrion and Sigalotti proved in [83] that, under generic assumptions on V, this system is approximately controllable on the $L^2(\mathbb{R}^N, \mathbb{C})$-sphere, in sufficiently large time T. In the present article we show that approximate controllability does not hold in arbitrarily small time, no matter what the initial state is. This generalizes our previous result for Gaussian initial conditions. Moreover, we prove that the minimal time can in fact be arbitrarily large.
- In [11] we consider the bilinear Schrödinger equation with discrete-spectrum drift. We show, for $n \in \mathbb{N}$ arbitrary, exact controllability in projections on the first n given eigenstates. The controllability result relies on a generic controllability hypothesis on some associated finite-dimensional approximations. The method is based on Lie-algebraic control techniques applied to the finite-dimensional approximations coupled with classical topological arguments issuing from degree theory.
- In [14] we consider the one dimensional Schrödinger equation with a bilinear control and prove the rapid stabilization of the linearized equation around the ground state. The feedback law ensuring the rapid stabilization is obtained using a transformation mapping the solution to the linearized equation on the solution to an exponentially stable target linear equation. A suitable condition is imposed on the transformation in order to cancel the non-local terms arising in the kernel system. This conditions also insures the uniqueness of the transformation. The continuity and invertibility of the transformation follows from exact controllability of the linearized system.
- In [33] we discuss how to control a parameter-dependent family of quantum systems. Our technique is based on adiabatic approximation theory and on the presence of curves of conical eigenvalue intersections of the controlled Hamiltonian. As particular cases, we recover chirped pulses for two-level quantum systems and counter-intuitive solutions for three-level stimulated Raman adiabatic passage (STIRAP). The proposed technique works for systems evolving both in finite-dimensional and infinite-dimensional Hilbert spaces. We show that the assumptions guaranteeing ensemble controllability are structurally stable with respect to perturbations of the parametrized family of systems.

6.3. Stability and uncertain dynamics: new results

Let us list here our new results about stability and stabilization of control systems, on the properties of systems with uncertain dynamics.

- In [8] we consider a one-dimensional controlled reaction-diffusion equation, where the control acts on the boundary and is subject to a constant delay. Such a model is a paradigm for more general parabolic systems coupled with a transport equation. We prove that it is possible to stabilize (in H^1 norm) this process by means of an explicit predictor-based feedback control that is designed from a finite-dimensional subsystem. The implementation is very simple and efficient and is based on standard tools of pole-shifting. Our feedback acts on the system as a finite-dimensional predictor. We compare our approach with the backstepping method.
- In [14] we consider the one dimensional Schrödinger equation with a bilinear control and prove the rapid stabilization of the linearized equation around the ground state. The feedback law ensuring the rapid stabilization is obtained using a transformation mapping the solution of the linearized equation to the solution of an exponentially stable target linear equation. A suitable condition is imposed on the transformation in order to cancel the non-local terms arising in the kernel system. This conditions also insures the uniqueness of the transformation. The continuity and invertibility of the transformation follows from exact controllability of the linearized system.
- Based on the notion of generalized homogeneity, we develop in [17] a new algorithm of feedback control design for a plant modeled by a linear evolution equation in a Hilbert space with a possibly unbounded operator. The designed control law steers any solution of the closed-loop system to zero in a finite time. Method of homogeneous extension is presented in order to make the developed control design principles to be applicable for evolution systems with non-homogeneous operators. The design scheme is demonstrated for heat equation with the control input distributed on the segment $[0, 1]$.
- In [19] we analyse the asymptotic behaviour of integro-differential equations modeling N populations in interaction, all structured by different traits. Interactions are modeled by non-local terms involving linear combinations of the total number of individuals in each population. These models have already been shown to be suitable for the modeling of drug resistance in cancer, and they generalise the usual Lotka–Volterra ordinary differential equations. Our aim is to give conditions under which there is persistence of all species. Through the analysis of a Lyapunov function, our first main result gives a simple and general condition on the matrix of interactions, together with a convergence rate. The second main result establishes another type of condition in the specific case of mutualistic interactions. When either of these conditions is met, we describe which traits are asymptotically selected.
- The goal of [20] is to compute a boundary control of reaction-diffusion partial differential equation. The boundary control is subject to a constant delay, whereas the equation may be unstable without any control. For this system equivalent to a parabolic equation coupled with a transport equation, a prediction-based control is explicitly computed. To do that we decompose the infinite-dimensional system into two parts: one finite-dimensional unstable part, and one stable infinite-dimensional part. A finite-dimensional delay controller is computed for the unstable part, and it is shown that this controller succeeds in stabilizing the whole partial differential equation. The proof is based on a an explicit form of the classical Artstein transformation, and an appropriate Lyapunov function. A numerical simulation illustrate the constructive design method.
- [27] focuses on the (local) small-time stabilization of a Korteweg-de Vries equation on bounded interval, thanks to a time-varying Dirichlet feedback law on the left boundary. Recently, backstepping approach has been successfully used to prove the null controllability of the corresponding linearized system, instead of Carleman inequalities. We use the “adding an integrator" technique to gain regularity on boundary control term which clears the difficulty from getting stabilization in small-time.
- Motivated by improved ways to disrupt brain oscillations linked to Parkinson’s disease, we propose

in [29] an adaptive output feedback strategy for the stabilization of nonlinear time-delay systems evolving on a bounded set. To that aim, using the formalism of input-to-output stability (IOS), we first show that, for such systems, internal stability guarantees robustness to exogenous disturbances. We then use this feature to establish a general result on scalar adaptive output feedback of time-delay systems inspired by the "σ-modification" strategy. We finally apply this result to a delayed neuronal population model and assess numerically the performance of the adaptive stimulation.

- In [35] we consider open channels represented by Saint-Venant equations that are monitored and controlled at the downstream boundary and subject to unmeasured flow disturbances at the upstream boundary. We address the issue of feedback stabilization and disturbance rejection under Proportional-Integral (PI) boundary control. For channels with uniform steady states, the analysis has been carried out previously in the literature with spectral methods as well as with Lyapunov functions in Riemann coordinates. In [35], our main contribution is to show how the analysis can be extended to channels with non-uniform steady states with a Lyapunov function in physical coordinates.
- In [37], we study the exponential stabilization of a shock steady state for the inviscid Burgers equation on a bounded interval. Our analysis relies on the construction of an explicit strict control Lyapunov function. We prove that by appropriately choosing the feedback boundary conditions, we can stabilize the state as well as the shock location to the desired steady state in H^2-norm, with an arbitrary decay rate.
- Given a discrete-time linear switched system $\Sigma(A)$ associated with a finite set A of matrices, we consider in [40] the measures of its asymptotic behavior given by, on the one hand, its deterministic joint spectral radius $\rho_d(A)$ and, on the other hand, its probabilistic joint spectral radii $\rho_p(v, P, A)$ for Markov random switching signals with transition matrix P and a corresponding invariant probability v. Note that $\rho_d(A)$ is larger than or equal to $\rho_p(v, P, A)$ for every pair (v, P). In this paper, we investigate the cases of equality of $\rho_d(A)$ with either a single $\rho_p(v, P, A)$ or with the supremum of $\rho_p(v, P, A)$ over (v, P) and we aim at characterizing the sets A for which such equalities may occur.
- In [41], we introduce a method to get necessary and sufficient stability conditions for systems governed by 1-D nonlinear hyperbolic partial-differential equations with closed-loop integral controllers, when the linear frequency analysis cannot be used anymore. We study the stability of a general nonlinear transport equation where the control input and the measured output are both located on the boundaries. The principle of the method is to extract the limiting part of the stability from the solution using a projector on a finite-dimensional space and then use a Lyapunov approach. We improve a result of Trinh, Andrieu and Xu, and give an optimal condition for the design of the controller. The results are illustrated with numerical simulations where the predicted stable and unstable regions can be clearly identified.
- In [44] we construct explicit time-varying feedback laws leading to the global (null) stabilization in small time of the viscous Burgers equation with three scalar controls. Our feedback laws use first the quadratic transport term to achieve the small-time global approximate stabilization and then the linear viscous term to get the small-time local stabilization.
- In [46] we address the question of the exponential stability for the C^1 norm of general 1-D quasilinear systems with source terms under boundary conditions. To reach this aim, we introduce the notion of basic C^1 Lyapunov functions, a generic kind of exponentially decreasing function whose existence ensures the exponential stability of the system for the C^1 norm. We show that the existence of a basic C^1 Lyapunov function is subject to two conditions: an interior condition, intrinsic to the system, and a condition on the boundary controls. We give explicit sufficient interior and boundary conditions such that the system is exponentially stable for the C^1 norm and we show that the interior condition is also necessary to the existence of a basic C^1 Lyapunov function. Finally, we show that the results conducted in this article are also true under the same conditions for the exponential stability in the C^p norm, for any $p \geq 1$.

- In [47] we study the exponential stability for the C^1 norm of general 2×2 1-D quasilinear hyperbolic systems with source terms and boundary controls. When the eigenvalues of the system have the same sign, any nonuniform steady-state can be stabilized using boundary feedbacks that only depend on measurements at the boundaries and we give explicit conditions on the gain of the feedback. In other cases, we exhibit a simple numerical criterion for the existence of basic C^1 Lyapunov function, a natural candidate for a Lyapunov function to ensure exponential stability for the C^1 norm. We show that, under a simple condition on the source term, the existence of a basic C^1 (or C^p , for any $p \geq 1$) Lyapunov function is equivalent to the existence of a basic H^2 (or H^q , for any $q \geq 2$) Lyapunov function, its analogue for the H^2 norm. Finally, we apply these results to the nonlinear Saint-Venant equations. We show in particular that in the subcritical regime, when the slope is larger than the friction, the system can always be stabilized in the C^1 norm using static boundary feedbacks depending only on measurements of at the boundaries, which has a large practical interest in hydraulic and engineering applications.
- In [48] we study the exponential stability in the H^2 norm of the nonlinear Saint-Venant (or shallow water) equations with arbitrary friction and slope using a single Proportional-Integral (PI) control at one end of the channel. Using a local dissipative entropy we find a simple and explicit condition on the gain the PI control to ensure the exponential stability of any steady-states. This condition is independent of the slope, the friction, the length of the river, the inflow disturbance and, more surprisingly, the steady-state considered. When the inflow disturbance is time-dependent and no steady-state exist, we still have the Input-to-State stability of the system, and we show that changing slightly the PI control enables to recover the exponential stability of slowly varying trajectories.
- The exponential stability problem of the nonlinear Saint-Venant equations is addressed in [49]. We consider the general case where an arbitrary friction and space-varying slope are both included in the system, which lead to non-uniform steady-states. An explicit quadratic Lyapunov function as a weighted function of a small perturbation of the steady-states is constructed. Then we show that by a suitable choice of boundary feedback controls, that we give explicitly, the local exponential stability of the nonlinear Saint-Venant equations for the H^2-norm is guaranteed.
- [53] elaborates control strategies to prevent clustering effects in opinion formation models. This is the exact opposite of numerous situations encountered in the literature where, on the contrary, one seeks controls promoting consensus. In order to promote declustering, instead of using the classical variance that does not capture well the phenomenon of dispersion, we introduce an entropy-type functional that is adapted to measuring pairwise distances between agents. We then focus on a Hegselmann-Krause-type system and design declustering sparse controls both in finite-dimensional and kinetic models. We provide general conditions characterizing whether clustering can be avoided as function of the initial data. Such results include the description of black holes (where complete collapse to consensus is not avoidable), safety zones (where the control can keep the system far from clustering), basins of attraction (attractive zones around the clustering set) and collapse prevention (when convergence to the clustering set can be avoided).
- In [54] we consider the problem of controlling parabolic semilinear equations arising in population dynamics, either in finite time or infinite time. These are the monostable and bistable equations on $(0, L)$ for a density of individuals $0 \leq y(t, x) \leq 1$, with Dirichlet controls taking their values in $[0, 1]$. We prove that the system can never be steered to extinction (steady state 0) or invasion (steady state 1) in finite time, but is asymptotically controllable to 1 independently of the size L, and to 0 if the length L of the interval domain is less than some threshold value L^*, which can be computed from transcendental integrals. In the bistable case, controlling to the other homogeneous steady state $0 < \theta < 1$ is much more intricate. We rely on a staircase control strategy to prove that θ can be reached in finite time if and only if $L < L^?$. The phase plane analysis of those equations is instrumental in the whole process. It allows us to read obstacles to controllability, compute the threshold value for domain size as well as design the path of steady states for the control strategy.
- Given a linear control system in a Hilbert space with a bounded control operator, we establish in [56] a characterization of exponential stabilizability in terms of an observability inequality.

Such dual characterizations are well known for exact (null) controllability. Our approach exploits classical Fenchel duality arguments and, in turn, leads to characterizations in terms of observability inequalities of approximately null controllability and of α-null controllability. We comment on the relationships between those various concepts, at the light of the observability inequalities that characterize them.

- In [58] we use the backstepping method to study the stabilization of a 1-D linear transport equation on the interval $(0, L)$, by controlling the scalar amplitude of a piecewise regular function of the space variable in the source term. We prove that if the system is controllable in a periodic Sobolev space of order greater than 1, then the system can be stabilized exponentially in that space and, for any given decay rate, we give an explicit feedback law that achieves that decay rate.

Let us also mention the lecture notes [31] on stabilization of semilinear PDE's, which have been published this year.

6.4. Optimal control: new results

Let us list here our new results in optimal control theory beyond the sub-Riemannian framework.

- In [4] we focus on regional deterministic optimal control problems, i.e., problems where the dynamics and the cost functional may be different in several regions of the state space and present discontinuities at their interface. Under the assumption that optimal trajectories have a locally finite number of switchings (no Zeno phenomenon), we use the duplication technique to show that the value function of the regional optimal control problem is the minimum over all possible structures of trajectories of value functions associated with classical optimal control problems settled over fixed structures, each of them being the restriction to some submanifold of the value function of a classical optimal control problem in higher dimension. The lifting duplication technique is thus seen as a kind of desingularization of the value function of the regional optimal control problem. In turn, we extend to regional optimal control problems the classical sensitivity relations and we prove that the regularity of this value function is the same (i.e., is not more degenerate) than the one of the higher-dimensional classical optimal control problem that lifts the problem.
- The goal of [9] is to show how non-parametric statistics can be used to solve some chance constrained optimization and optimal control problems. We use the Kernel Density Estimation method to approximate the probability density function of a random variable with unknown distribution, from a relatively small sample. We then show how this technique can be applied and implemented for a class of problems including the Goddard problem and the trajectory optimization of an Ariane 5-like launcher.
- In control theory the term chattering is used to refer to fast oscillations of controls, such as an infinite number of switchings over a finite time interval. In [10] we focus on three typical instances of chattering: the Fuller phenomenon, referring to situations where an optimal control features an accumulation of switchings in finite time; the Robbins phenomenon, concerning optimal control problems with state constraints, where the optimal trajectory touches the boundary of the constraint set an infinite number of times over a finite time interval; and the Zeno phenomenon, for hybrid systems, referring to a trajectory that depicts an infinite number of location switchings in finite time. From the practical point of view, when trying to compute an optimal trajectory, for instance, by means of a shooting method, chattering may be a serious obstacle to convergence. In [10] we propose a general regularization procedure, by adding an appropriate penalization of the total variation. This produces a family of quasi-optimal controls whose associated cost converge to the optimal cost of the initial problem as the penalization tends to zero. Under additional assumptions, we also quantify quasi-optimality by determining a speed of convergence of the costs.
- In [12], a new robust and fast method is developed to perform transfers that minimize fuel consumption between two invariant manifolds of periodic orbits in the circular restricted three-body problem. The method starts with an impulse transfer between two invariant manifolds to build an optimal

control problem. This allows to choose an adequate fixed transfer time. Using the Pontryagin maximum principle, the resolution of the problem is formulated as that of finding the zero of a shooting function (indirect method). The algorithm couples different kinds of continuations (on cost, final state, and thrust) to improve robustness and to initialize the solver. The efficiency of the method is illustrated with numerical examples. Finally, the influence of the transfer time is studied numerically thanks to a continuation on this parameter, and it checks that, when transfer duration goes to zero, the control converges to the impulse transfer that it started with. It shows the robustness of the method and establishes a mathematical link between the two problems.

- In [15] we consider the controllability problem for finite-dimensional linear autonomous control systems, under state constraints but without imposing any control constraint. It is well known that, under the classical Kalman condition, in the absence of constraints on the state and the control, one can drive the system from any initial state to any final one in an arbitrarily small time. Furthermore, it is also well known that there is a positive minimal time in the presence of compact control constraints. We prove that, surprisingly, a positive minimal time may be required as well under state constraints, even if one does not impose any restriction on the control. This may even occur when the state constraints are unilateral, like the nonnegativity of some components of the state, for instance. Using the Brunovsky normal forms of controllable systems, we analyze this phenomenon in detail, that we illustrate by several examples. We discuss some extensions to nonlinear control systems and formulate some challenging open problems.
- In [18] we consider a system of two coupled integro-differential equations modeling populations of healthy and cancer cells under therapy. Both populations are structured by a phenotypic variable, representing their level of resistance to the treatment. We analyse the asymptotic behaviour of the model under constant infusion of drugs. By designing an appropriate Lyapunov function, we prove that both densities converge to Dirac masses. We then define an optimal control problem, by considering all possible infusion protocols and minimising the number of cancer cells over a prescribed time frame. We provide a quasi-optimal strategy and prove that it solves this problem for large final times. For this modeling framework, we illustrate our results with numerical simulations, and compare our optimal strategy with periodic treatment schedules.
- In [21] we use conductance based neuron models and the mathematical modeling of Optogenetics to define controlled neuron models and we address the minimal time control of these affine systems for the first spike from equilibrium. We apply tools of geometric optimal control theory to study singular extremals and we implement a direct method to compute optimal controls. When the system is too large to theoretically investigate the existence of singular optimal controls, we observe numerically the optimal bang-bang controls.
- In [23] we first derive a general integral-turnpike property around a set for infinite-dimensional non-autonomous optimal control problems with any possible terminal state constraints, under some appropriate assumptions. Roughly speaking, the integral-turnpike property means that the time average of the distance from any optimal trajectory to the turnpike set converges to zero, as the time horizon tends to infinity. Then, we establish the measure-turnpike property for strictly dissipative optimal control systems, with state and control constraints. The measure-turnpike property, which is slightly stronger than the integral-turnpike property, means that any optimal (state and control) solution remains essentially, along the time frame, close to an optimal solution of an associated static optimal control problem, except along a subset of times that is of small relative Lebesgue measure as the time horizon is large. Next, we prove that strict strong duality, which is a classical notion in optimization, implies strict dissipativity, and measure-turnpike. Finally, we conclude the paper with several comments and open problems.
- In [24], we investigate the asymptotic behavior of optimal designs for the shape optimization of 2D heat equations in long time horizons. The control is the shape of the domain on which heat diffuses. The class of 2D admissible shapes is the one introduced by Sverák, of all open subsets of a given bounded open set, whose complementary sets have a uniformly bounded number of connected components. Using a Γ-convergence approach, we establish that the parabolic optimal

designs converge as the length of the time horizon tends to infinity, in the complementary Hausdorff topology, to an optimal design for the corresponding stationary elliptic equation.

- In [25], we study the steady-state (or periodic) exponential turnpike property of optimal control problems in Hilbert spaces. The turnpike property, which is essentially due to the hyperbolic feature of the Hamiltonian system resulting from the Pontryagin maximum principle, reflects the fact that, in large time, the optimal state, control and adjoint vector remain most of the time close to an optimal steady-state. A similar statement holds true as well when replacing an optimal steady-state by an optimal periodic trajectory. To establish the result, we design an appropriate dichotomy transformation, based on solutions of the algebraic Riccati and Lyapunov equations. We illustrate our results with examples including linear heat and wave equations with periodic tracking terms.
- The Allee threshold of an ecological system distinguishes the sign of population growth either towards extinction or to carrying capacity. In practice human interventions can tune the Allee threshold for instance thanks to the sterile male technique and the mating disruption. In [26] we address various control objectives for a system described by a diffusion-reaction equation regulating the Allee threshold, viewed as a real parameter determining the unstable equilibrium of the bistable nonlinear reaction term. We prove that this system is the mean field limit of an interacting system of particles in which individual behaviours are driven by stochastic laws. Numerical simulations of the stochastic process show that population propagations are governed by wave-like solutions corresponding to traveling solutions of the macroscopic reaction-diffusion system. An optimal control problem for the macroscopic model is then introduced with the objective of steering the system to a target traveling wave. The relevance of this problem is motivated by the fact that traveling wave solutions model the fact that bounded space domains reach asymptotically an equilibrium configuration. Using well known analytical results and stability properties of traveling waves, we show that well-chosen piecewise constant controls allow to reach the target approximately in sufficiently long time. We then develop a direct computational method and show its efficiency for computing such controls in various numerical simulations. Finally we show the efficiency of the obtained macroscopic optimal controls in the microscopic system of interacting particles and we discuss their advantage when addressing situations that are out of reach for the analytical methods. We conclude the article with some open problems and directions for future research.
- Consider a general nonlinear optimal control problem in finite dimension, with constant state and/or control delays. By the Pontryagin Maximum Principle, any optimal trajectory is the projection of a Pontryagin extremal. In [39] we establish that, under appropriate assumptions, Pontryagin extremals depend continuously on the parameter delays, for adequate topologies. The proof of the continuity of the trajectory and of the control is quite easy, however, for the adjoint vector, the proof requires a much finer analysis. The continuity property of the adjoint with respect to the parameter delay opens a new perspective for the numerical implementation of indirect methods, such as the shooting method. We also discuss the sharpness of our assumptions.
- In [43] we are concerned about the controllability of a general linear hyperbolic system of the form $\partial_t w(t,x) = \Sigma(x)\partial_x w(t,x) + \gamma C(x) w(t,x)$ ($\gamma \in \mathbb{R}$) in one space dimension using boundary controls on one side. More precisely, we establish the optimal time for the null and exact controllability of the hyperbolic system for generic γ. We also present examples which yield that the generic requirement is necessary. In the case of constant Σ and of two positive directions, we prove that the null-controllability is attained for any time greater than the optimal time for all $\gamma \in \mathbb{R}$ and for all C which is analytic if the slowest negative direction can be alerted by both positive directions. We also show that the null-controllability is attained at the optimal time by a feedback law when $C \equiv 0$. Our approach is based on the backstepping method paying a special attention on the construction of the kernel and the selection of controls.
- In [52] we consider a state-constrained optimal control problem of a system of two non-local partial-differential equations, which is an extension of the one introduced in a previous work in mathematical oncology. The aim is to minimize the tumor size through chemotherapy while avoiding the emergence of resistance to the drugs. The numerical approach to solve the problem was the

combination of direct methods and continuation on discretization parameters, which happen to be insufficient for the more complicated model, where diffusion is added to account for mutations. In [52], we propose an approach relying on changing the problem so that it can theoretically be solved thanks to a Pontryagin Maximum Principle in infinite dimension. This provides an excellent starting point for a much more reliable and efficient algorithm combining direct methods and continuations. The global idea is new and can be thought of as an alternative to other numerical optimal control techniques.

We would also like to mention the defense of the PhD theses of Riccardo Bonalli [1] and Antoine Olivier [2] on the subject.

MATHERIALS Project-Team

6. New Results

6.1. Electronic structure calculations

Participants: Robert Benda, Éric Cancès, Virginie Ehrlacher, Antoine Levitt, Sami Siraj-Dine, 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 by an appropriate mathematical analysis, and improving the numerical approaches by a dedicated numerical analysis.

6.1.1. Mathematical analysis

The members of the team have continued their systematic study of the properties of materials in the reduced Hartree-Fock approximation, a model striking a good balance between mathematical tractability and the ability to reproduce qualitatively complex effects.

E. Cancès and G. Stoltz have studied with L. Cao models for certain extended defects in materials [37]. These extended defects typically correspond to taking out a slab of finite width in the three-dimensional homogeneous electron gas. The work is performed in the framework of the reduced Hartree-Fock model with either Yukawa or Coulomb interactions, using techniques previously developed to study local perturbations of the free-electron gas. It is shown that the model admits minimizers, and that Yukawa ground state energies and density matrices converge to ground state Coulomb energies and density matrices as the Yukawa parameter tends to zero. These minimizers are unique for Yukawa interactions, and are characterized by a self-consistent equation. Numerical simulations show evidence of Friedel oscillations in the total electronic density.

A. Levitt has examined the phenomenon of screening in materials. In [54] he has studied the effect of adding a small charge to a periodic system modeled by the reduced Hartree-Fock at finite temperature. He has showed that the reaction potential created by the rearrangement of the electrons counteracts exactly the free charge, so that the effective interaction in such systems is short-range. The proof proceeds by studying the properties of the linear response operator, which also sheds some light on the charge-sloshing instability seen in numerical methods to solve the self-consistent equations.

6.1.2. Numerical analysis

E. Cancès has pursued his long-term collaboration with Y. Maday (Sorbonne Université) on the numerical analysis of linear and nonlinear eigenvalue problems. Together with G. Dusson (Warwick, United Kingdom), B. Stamm (Aachen, Germany), and M. Vohralík (Inria SERENA), they have designed *a posteriori* error estimates for conforming numerical approximations of the Laplace eigenvalue problem with homogeneous Dirichlet boundary conditions. In [38], they prove *a priori* error estimates for the perturbation-based post-processing of the plane-wave approximation of Schrödinger equations introduced and tested numerically in previous works. They consider a Schrödinger operator $H = -\frac{1}{2}\Delta + V$ on $L^2(\Omega)$, where Ω is a cubic box with periodic boundary conditions. The quantities of interest are, on the one hand, the ground-state energy defined as the sum of the lowest N eigenvalues of H , and, on the other hand, the ground-state density matrix, that is the spectral projector on the vector space spanned by the associated eigenvectors. Such a problem is central in first-principle molecular simulation, since it corresponds to the so-called linear subproblem in Kohn-Sham density functional theory (DFT). Interpreting the exact eigenpairs of H as perturbations of the numerical eigenpairs obtained by a variational approximation in a plane-wave (i.e. Fourier) basis, they compute first-order corrections for the eigenfunctions, which are turned into corrections on the ground-state density matrix. This allows them to increase the accuracy of both the ground-state energy and the ground-state density matrix at a low computational extra-cost. Indeed, the computation of the corrections only requires the computation of the residual of the solution in a larger plane-wave basis and $2N$ Fast Fourier Transforms.

Implicit solvation models aim at computing the properties of a molecule in solution (most chemical reactions take place in the liquid phase) by replacing all the solvent molecules but the few ones strongly interacting with the solute, by an effective continuous medium accounting for long-range electrostatics. E. Cancès, Y. Maday (Sorbonne Université), and B. Stamm (Aachen, Germany) have introduced a few years ago 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 and B. Mennucci (Chemistry, Pisa, Italy) and J.-P. Piquemal (Sorbonne Université), they have implemented this algorithm in widely used computational software products (Gaussian and Tinker). Together with L. Lagardère (Sorbonne Université) and G. Scalmani (Gaussian Inc., USA), they illustrate in [29] the domain decomposition COSMO (ddCOSMO) implementation and how to couple it with an existing classical or quantum mechanical (QM) codes. They review in detail what input needs to be provided to ddCOSMO and how to assemble it, describe how the ddCOSMO equations are solved and how to process the results in order to assemble the required quantities, such as Fock matrix contributions for the QM case, or forces for the classical one. Throughout the paper, they make explicit references to the ddCOSMO module, which is an open source, Fortran 90 implementation of ddCOSMO that can be downloaded and distributed under the LGPL license.

E. Cancès, V. Ehrlacher and A. Levitt, together with D. Gontier (Dauphine) and D. Lombardi (Inria REO), have studied the convergence of properties of periodic systems as the size of the computing domain is increased. This convergence is known to be difficult in the case of metals. They have characterized in [39] the speed of convergence for a number of schemes in the metallic case, and have studied the properties of a widely used numerical method that adds an artificial electronic temperature.

A. Levitt has continued his study of Wannier functions in periodic systems. With A. Damle (Cornell, USA) and L. Lin (Berkeley, USA), they have proposed an efficient numerical method for the computation of maximally-localized Wannier functions in metals, and have showed on the example of the free electron gas that they are not in general exponentially localized [42]. With D. Gontier (Dauphine) and S. Siraj-Dine, they proposed a new method for the computation of Wannier functions which applies to any insulator, and in particular to the difficult case of topological insulators [45].

6.2. Computational Statistical Physics

Participants: Grégoire Ferré, Florent Hédin, Frédéric Legoll, Tony Lelièvre, Mouad Ramil, Julien Roussel, Laura Silva Lopes, Gabriel Stoltz, Pierre Terrier.

The objective of computational statistical physics is to compute macroscopic properties of materials starting from a microscopic description, using concepts of statistical physics (thermodynamic ensembles and molecular dynamics). The contributions of the team can be divided into four main topics: (i) the development of methods for sampling the configuration space; (ii) the numerical analysis of such methods; (iii) the efficient computation of dynamical properties which requires to sample metastable trajectories; (iv) coarse-graining techniques to reduce the computational cost of molecular dynamic simulations and gain some insights on the models.

6.2.1. Sampling of the configuration space: new algorithms and applications

New numerical methods in order to sample probability measures on the configuration space have been developed: either measures supported on submanifolds, or stationary states of stochastic dynamics. First, in [51], T. Lelièvre and G. Stoltz, together with M. Rousset (Inria Rennes, France) have studied how to sample probability measures supported on submanifolds, by adding an extra momentum variable to the state of the system, and discretizing the associated Hamiltonian dynamics with some stochastic perturbation in the extra variable. In order to avoid biases in the invariant probability measures sampled by discretizations of these stochastically perturbed Hamiltonian dynamics, a Metropolis rejection procedure can be considered. The so-obtained scheme belongs to the class of generalized Hybrid Monte Carlo (GHMC) algorithms. However, the usual method has to be generalized using a procedure suggested by Goodman, Holmes-Cerfon and Zappa for Metropolis random walks on submanifolds, where a reverse projection check is performed to enforce the reversibility of the algorithm for large timesteps and hence avoid biases in the invariant measure. A full mathematical analysis of such procedures is provided, as well as numerical experiments demonstrating the

importance of the reverse projection check on simple toy examples. Second, the work [55] by J. Roussel and G. Stoltz focuses on the use of control variates for non-equilibrium systems. Whereas most variance reduction methods rely on the knowledge of the invariant probability measure, this latter is not explicit out of equilibrium. Control variates offer an attractive alternative in this framework. J. Roussel and G. Stoltz have proposed a general strategy for constructing an efficient control variate, relying on physical simplifications of the dynamics. The authors provide an asymptotic analysis of the variance reduction in a perturbative framework, along with extensive numerical tests on three different systems.

In terms of applications of such sampling techniques, members of the project-team have been working on two different subjects: random matrices models and adaptive techniques to compute large deviation rate functionals. The paper [16] was written by G. Ferré and D. Chafaï (Université Paris Dauphine, France), following the simple idea: the eigenvalues of random matrices are distributed according to Boltzmann–Gibbs measures, but researchers in this field do not use techniques from statistical physics for numerical investigations. The authors therefore used a Hamiltonian Monte Carlo algorithm to investigate numerically conjectures about random matrices and related Coulomb gases. The next step is to add constraints to these systems to understand better the behavior of random matrices with constraints and the large size limit of their spectra (the algorithm mentioned above to sample probability measures supported on submanifolds may be useful in this context). The work [19] focuses on computing free energies and entropy functions, as they arise in large deviations theory, through adaptive techniques. It is actually in the spirit of techniques used in mathematical finance, adapted to the statistical mechanics context, and enriched with new estimators based on variational representations of entropy functions. These tools have been pioneered by H. Touchette (Stellenbosch University, South Africa), with whom the paper was written by G. Ferré.

6.2.2. Sampling of the configuration space: numerical analysis

Concerning the numerical analysis of sampling techniques of probability measures on the configuration space, let us mention three works.

First, in [44], G. Ferré and G. Stoltz study the numerical errors that arise when a stochastic differential equation (SDE) is discretized in order to compute scaled cumulant functions (or free energy) and ergodic properties of Feynman–Kac semigroups. These quantities naturally arise in large deviations theory, for estimating probabilities of rare events. This analysis is made difficult by the nonlinear (mean field) feature of the dynamics at hand. The obtained estimates generalize previous results on the numerical analysis of ergodic properties of discretized SDEs. As a theoretical extension of the previous work, the purpose of the work [43] by G. Ferré and G. Stoltz, in collaboration with M. Rousset (Inria Rennes, France), is to provide further theoretical investigations on the long time behavior of Feynman–Kac semigroups. More precisely, it aims at giving practical criteria for these nonlinear semigroups to have a limit, and makes precise in which sense this limit is to be understood. This was an open problem so far for systems evolving in unbounded configuration spaces, which was addressed through Lyapunov function techniques. Although theoretical, these results are of practical importance since, if these dynamics do not have a well-defined long time behavior, it is hopeless to try to compute rare events.

Finally, together with C. Andrieu (Univ. Bristol, United-Kingdom), A. Durmus (ENS Saclay, France) and N. Nüsken (Univ. Potsdam, Germany), J. Roussel derived in [32] spectral gap estimates for several Piecewise Deterministic Markov Processes (PDMPs), namely the Randomized Hamiltonian Monte Carlo, the Zig-Zag process and the Bouncy Particle Sampler. The hypocoercivity technique provides estimates with explicit dependence on the parameters of the dynamics. Moreover the general framework considered allows to compare quantitatively the bounds found for the different methods. Such PDMDs are currently more and more used as efficient sampling tools, but their theoretical properties are still not yet well understood.

6.2.3. Sampling of dynamical properties and rare events

The sampling of dynamical properties along molecular dynamics trajectories is crucial to get access to important quantities such as transition rates or reactive paths. This is difficult numerically because of the metastability of trajectories. Members of the project-team are following two numerical approaches to sample

metastable trajectories: the accelerated dynamics *à la* A.F. Voter and the adaptive multilevel splitting (AMS) technique to sample reactive paths between metastable states.

Concerning the mathematical analysis of the accelerated dynamics, in [50], T. Lelièvre reviews the recent mathematical approaches to justify these numerical methods, using the notion of quasi-stationary distribution. Moreover, in [49], T. Lelièvre together with D. Le Peutrec (Université de Paris Saclay, France) and G. Di Gesu and B. Nectoux (TU Wien, Austria) give an overview of the results obtained during the PhD of B. Nectoux. Using the quasi-stationary distribution approach and tools from semi-classical analysis, one can justify the use of kinetic Monte Carlo models parametrized by the Eyring-Kramers formulas to describe exit events from metastable states, for the overdamped Langevin dynamics. Concerning the implementation, in [22], F. Hédin and T. Lelièvre test the Generalized Parallel Replica algorithm to biological systems, and obtain strong linear scalability, providing up to 70% of the maximum possible speedup on several hundreds of CPUs. The "Parallel Replica" (ParRep) dynamics is known for allowing to simulate very long trajectories of metastable Langevin dynamics in the materials science community, but it relies on assumptions that can hardly be transposed to the world of biochemical simulations. The later developed "Generalized ParRep" variant solves those issues, and it had not been applied to significant systems of interest so far. Finally, let us mention the work [27] where T. Lelièvre together with J. Reygner (Ecole des Ponts, France) and L. Pillaud-Vivien (Inria Paris, France) analyze mathematically the Fleming-Viot particle process in the simple case of a finite state space. This Fleming-Viot particle process is a key ingredient of the Generalized ParRep algorithm mentioned above, in order to both approximate the convergence time to the quasi-stationary distribution, and to efficiently sample it.

Concerning the AMS technique, in [36], T. Lelièvre and C.-E. Bréhier (ENS Lyon, France) test new importance functions to compute rare events associated with the law of the solution to a stochastic differential equation at a given fixed time. This can be used for example to estimate the rate functional for large deviation principle applied to time averages.

6.2.4. Coarse-graining

In two related works, members of the project-team have studied the quality of the effective dynamics derived from a high dimension stochastic differential equation on a few degrees of freedom, using a projection approach *à la Mori-Zwanzig*. More precisely, in [48], F. Legoll, T. Lelièvre and U. Sharma obtain precise error bounds in the case of non reversible dynamics. This analysis also aims at discussing what is a good notion of mean force for non reversible systems. In [53], T. Lelièvre together with W. Zhang (ZIB, Germany) extend previous results on pathwise error estimates for such effective dynamics to the case of nonlinear vectorial reaction coordinates.

Once a good coarse-grained model has been obtained, one can try to use it in order to get a better integrator of the original dynamic in the spirit of a predictor-corrector method. In [52], T. Lelièvre together with G. Samaey and P. Zielinski (KU Leuven, Belgium) analyze such a micro-macro acceleration method for the Monte Carlo simulation of stochastic differential equations with time-scale separation between the (fast) evolution of individual trajectories and the (slow) evolution of the macroscopic function of interest.

6.3. Homogenization

Participants: Virginie Ehrlacher, Marc Josien, Claude Le Bris, Frédéric Legoll, Adrien Lesage, Pierre-Loïk Rothé.

6.3.1. Deterministic non-periodic systems

In homogenization theory, members of the project-team have pursued their ongoing systematic study of perturbations of periodic problems (by local and nonlocal defects). This has been done in two different directions.

For linear elliptic equations, C. Le Bris has written, in collaboration with X. Blanc (Paris Diderot, France) and P.-L. Lions (Collège de France, France), two manuscripts that present a more versatile proof of the existence of a corrector function for periodic problems with local defects, and also extend the results: the first manuscript [34] addresses the case of an equation (or a system) in divergence form, while the second manuscript [12] extends the analysis to advection-diffusion equations.

Second, they have also provided more details on the quality of approximation achieved by their theory. The fact that a corrector exists with suitable properties allows one to quantify the rate of convergence of the two-scale expansion using that corrector to the actual exact solution, as the small homogenization parameter ε vanishes. These works by C. Le Bris, in collaboration with X. Blanc and M. Josien (and in the context of the PhD thesis of the latter), will be presented in a series of manuscripts in preparation. The precise results have been announced in [11] and proven in [33]. A related study [47] has been performed by M. Josien and addresses issues regarding periodic Green functions.

Also in the context of homogenization theory, C. Le Bris and F. Legoll have initiated a collaboration with R. Cottereau (Ecole Centrale and now CNRS Marseille, France). The topic is in some sense a follow-up on both an earlier work of R. Cottereau and the series of works completed by C. Le Bris and F. Legoll in collaboration with K. Li and next S. Lemaire over the years. Schematically, the purpose of the work is to determine the homogenized coefficient for a medium without explicitly performing a homogenization approach nor using a MsFEM type approach. In earlier works, an approximation approach, somewhat engineering-style, was designed. The purpose now is to examine the performance of this approach in the context of the so-called Arlequin method, a very popular method in the mechanical engineering community. One couples a sub-region of the medium where a homogeneous model is employed, along with a complementary sub-region where the original multiscale model is solved explicitly. The coupling is performed using the Arlequin method. Then, one optimizes a suitable criterion so that optimization leads to an homogeneous sub-region indeed described by the homogenized coefficient seeked for. Some numerical analysis questions, together with practical perspectives for computational enhancements of the approach, are currently examined.

Finally, C. Le Bris has informally participated into the supervision of the master thesis of S. Wolf (Ecole Normale Supérieure, Paris, France), and in this context performed some works in interaction with the student and X. Blanc. The purpose is to investigate perturbations of periodic homogenization problems when the perturbation is geometric in nature. The test case considered is that of a domain perforated by holes the locations of which are not necessarily periodic, but only periodic up to a local perturbation. The results proven, on the prototypical Poisson equation, are natural extensions of the celebrated results by J.-L. Lions published in the late 1960s for the periodic case. This provides a proof of concept, showing that perturbations of a periodic geometry are also possible, a fact that will be more thoroughly investigated in the near future within the above mentioned collaboration.

6.3.2. Stochastic homogenization

The project-team has pursued its efforts in the field of stochastic homogenization of elliptic equations, aiming at designing numerical approaches that are practically relevant and keep the computational workload limited.

Using standard homogenization theory, one knows that the homogenized tensor, which is a deterministic matrix, depends on the solution of a stochastic equation, the so-called corrector problem, which is posed on the whole space $\mathbb{R}^d$. This equation is therefore delicate and expensive to solve. A standard approach consists in truncating the space $\mathbb{R}^d$ to some bounded domain, on which the corrector problem is numerically solved.

In collaboration with B. Stamm (Aachen University, Germany) and S. Xiang (now also at Aachen University, Germany), E. Cancès, V. Ehrlacher and F. Legoll have studied, both from a theoretical and a numerical standpoints, new alternatives for the approximation of the homogenized matrix. They all rely on the use of an embedded corrector problem, previously introduced by the authors, where a finite-size domain made of the highly oscillatory material is embedded in a homogeneous infinite medium whose diffusion coefficients have to be appropriately determined. In [40], they have shown that the different approximations introduced all converge to the homogenized matrix of the medium when the size of the embedded domain goes to infinity. In [41], they present an efficient algorithm for the resolution of such problems for particular heterogeneous materials, based on the reformulation of the embedded corrector problem as an integral equation, which is discretized using spherical harmonics and solved using the fast multipole method.

Besides the averaged behavior of the oscillatory solution u_ε on large space scales (which is given by its homogenized limit), a question of interest is to describe how u_ε fluctuates. This question is investigated in the PhD thesis of P.-L. Rothé, both from a theoretical and a numerical viewpoints. First, theoretical results

have been obtained for a weakly stochastic setting (where the coefficient is the sum of a periodic coefficient and a small random perturbation). It has been shown that, at the first order and when ε is small, the localized fluctuations (characterized by a test function g) of u_ε are Gaussian. The corresponding variance depends on the localization function g, on the right-hand side f of the problem satisfied by u_ε, and on a fourth order tensor Q which is defined in terms of the corrector. Since the corrector function is challenging to compute, so is Q. A numerical approach has hence been designed to approximate Q and its convergence has been proven. Second, numerical experiments in more general settings (i.e. full stochastic case) following the same approach have been performed. The results are promising, and consistent with the theoretical results obtained in the weakly stochastic setting. These results are collected in a manuscript in preparation.

In collaboration with T. Hudson (University of Warwick, United Kingdom), F. Legoll and T. Lelièvre have considered in [46] a scalar viscoelastic model in which the constitutive law is random and varies on a lengthscale which is small relative to the overall size of the solid. Using stochastic two-scale convergence, they have obtained the homogenized limit of the evolution, and have demonstrated that, under certain hypotheses, the homogenized model exhibits hysteretic behaviour which persists under asymptotically slow loading. This work is motivated by rate-independent stress-strain hysteresis observed in filled rubber.

6.3.3. Multiscale Finite Element approaches

From a numerical perspective, the Multiscale Finite Element Method (MsFEM) is a classical strategy to address the situation when the homogenized problem is not known (e.g. in difficult nonlinear cases), or when the scale of the heterogeneities, although small, is not considered to be zero (and hence the homogenized problem cannot be considered as a sufficiently accurate approximation).

During the year, several research tracks have been pursued in this general direction.

The final writing of the various works performed in the context of the PhD thesis of F. Madiot is still ongoing. The issues examined there are on the one hand the application (and adequate adjustment) of MsFEM approaches to the case of an advection-diffusion equation with a dominating convection term posed in a perforated domain, and on the other hand some more general study of a numerical approach based, again in the case of convection-dominated flows, on the introduction of the invariant measure associated to the problem. The final version of the two manuscripts describing the efforts in each of these directions should be completed in a near future.

The MsFEM approach uses a Galerkin approximation of the problem on a pre-computed basis, obtained by solving local problems mimicking the problem at hand at the scale of mesh elements, with carefully chosen right-hand sides and boundary conditions. The initially proposed version of MsFEM uses as basis functions the solutions to these local problems, posed on each mesh element, with null right-hand sides and with the coarse P1 elements as Dirichlet boundary conditions. Various improvements have next been proposed, such as the *oversampling* variant, which solves local problems on larger domains and restricts their solutions to the considered element. In collaboration with U. Hetmaniuk (University of Washington in Seattle, USA), C. Le Bris, F. Legoll and P.-L. Rothé have introduced and studied a MsFEM method improved differently. They have considered a variant of the classical MsFEM approach with enrichments based on Legendre polynomials, both in the bulk of the mesh elements and on their interfaces. A convergence analysis of this new variant has been performed. Promising numerical results have been obtained. These results are currently being collected in a manuscript in preparation.

One of the perspectives of the team, through the PhD thesis of A. Lesage, is the development of Multiscale Finite Element Methods for thin heterogeneous plates. The fact that one of the dimension of the domain of interest scales as the typical size of the heterogeneities within the material induces theoretical and practical difficulties that have to be carefully taken into account. The first steps of the work of V. Ehrlacher, F. Legoll and A. Lesage, in collaboration with A. Lebée (École des Ponts) have consisted in studying the homogenized limit (and the two-scale expansion) of problems posed on thin heterogeneous plates. The case of a diffusion equation has been first dealt with, while the more challenging case of elasticity is currently under study.

6.4. Complex fluids

Participants: Sébastien Boyaval, Dena Kazerani.

The aim of the research performed in the project-team about complex fluids is

- to guide the mathematical modeling with PDEs of real materials flows, multi-phase fluids such as suspensions of particles or stratified air-water flows in particular, and
- to propose efficient algorithms for the computation of flow solutions, mainly for the many applications in the hydraulic engineering context.

Concerning the first point, new results have been obtained in collaboration with A. Caboussat (HEG, Switzerland) and M. Picasso (EPFL, Switzerland), in the framework of the SEDIFLO project (funded by ANR) and of Arwa Mrad PhD thesis at EPFL. In [13], they have shown numerically inability of some classical incompressible density-dependent Navier-Stokes equations to take into account some multiphase concentration effects in a prototypical set-up of fluvial erosion (in comparison with physical experiments). Hence the need for *new* models, that better describe complex flows associated with heterogeneities in the fluid microstructure. Concerning the second point, new results have been obtained in collaboration with M. Grepl and K. Veroy (Aachen, Germany) regarding the numerical reduction of transport models for data assimilation [25], in the framework of M. Kaercher PhD thesis at Aachen.

MATHRISK Project-Team

7. New Results

7.1. Risk management in finance and insurance

7.1.1. Control of systemic risk in a dynamic framework

Interconnected systems are subject to contagion in time of distress. Recent effort has been dedicated to understanding the relation between network topology and the scope of distress propagation. It is critical to recognize that connectivity is a result of an optimization problem of agents, who derive benefits from connections and view the associated contagion risk as a cost. In our previous works on the control of contagion in financial systems (see e.g. [80], [41], [5]), a central party, for example a regulator or government, seeks to minimize contagion. In [54], in contrast, the financial institutions themselves are the decision makers, and their decision is made before the shock, with a rational expectation on the way the cascade will evolve following the shock. We are extending these studies in a *dynamic* framework by allowing a recovery feature in the financial system during the cascade process, captured by introducing certain extent of growth of the banks' assets between each round of contagion.

7.1.2. Option pricing in financial markets with imperfections and default

A. Sulem, M.C. Quenez and R. Dumitrescu have studied robust pricing in an imperfect financial market with default. In this setting, the pricing system is expressed as a nonlinear g-expectation $\mathcal{E}^g$ induced by a nonlinear BSDE with nonlinear driver g and default jump (see [24]). The case of American options in this market model is treated in [19]. The incomplete market case is under study.

7.1.3. American options

With Giulia Terenzi, D. Lamberton has been been working on American options in Heston's model. They have some 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). Their paper "Variational formulation of American option prices in the Heston model" [32] is now in minor revision for *SIAM Journal on Financial Mathematics*.

They also have some results on monotonicity and regularity properties of the price function.

D. Lamberton has also a paper on the binomial approximation of the American put, in which a new bound for the rate of convergence of the binomial approximation of the Black-Scholes American put price is derived [32].

Optimal stopping problems involving the maximum of a diffusion is currently under investigation. Partial results obtained by D. Lamberton and M. Zervos) enable them to treat reward functions with little regularity.

7.1.4. Monte-Carlo methods for the computation of the Solvency Capital Requirement (SCR) in Insurance

A. Alfonsi has obtained a grant from AXA Foundation on a Joint Research Initiative with a team of AXA France working on the strategic asset allocation. This team has to make recommendations on the investment over some assets classes as, for example, equity, real estate or bonds. In order to do that, each side of the balance sheet (assets and liabilities) is modeled in order to take into account their own dynamics but also their interactions. Given that the insurance products are long time contracts, the projections of the company's margins have to be done considering long maturities. When doing simulations to assess investment policies, it is necessary to take into account the SCR which is the amount of cash that has to be settled to manage the portfolio. Typically, the computation of the future values of the SCR involve expectations under conditional laws, which is greedy in computation time. The goal of this project is to develop efficient Monte-Carlo methods to compute the SCR for long investment strategies. A. Cherchali has started his PhD thesis in September 2017 on this topic.

A. Alfonsi and A. Cherchali are developing a model of the ALM management of insurance companies that takes into account the regulatory constraints on life-insurance. We are testing this model. The purpose is then to use this model to develop Monte-Carlo methods to approximate the SCR (Solvency Capital Requirement).

7.2. Optimal transport and applications

7.2.1. Martingale Optimal Transport.

B. Jourdain and W. Margheriti exhibit a new family of martingale couplings between two one-dimensional probability measures μ and ν in the convex order. This family is parametrised by two dimensional probability measures on the unit square with respective marginal densities proportional to the positive and negative parts of the difference between the quantile functions of μ and ν. It contains the inverse transform martingale coupling which is explicit in terms of the cumulative distribution functions of these marginal densities. The integral of $|x-y|$ with respect to each of these couplings is smaller than twice the W^1 distance between μ and ν. When the comonotoneous coupling between μ and ν is given by a map T, the elements of the family minimize $\int_{\mathbf{R}} |y-T(x)|M(dx,dy)$ among all martingale couplings M between μ and ν. When μ and ν are in the decreasing (resp. increasing) convex order, the construction can be generalized to exhibit super (resp. sub) martingale couplings.

A. Alfonsi and B. Jourdain show that any optimal coupling for the quadratic Wasserstein distance $W_2^2(\mu,\nu)$ between two probability measures μ and ν with finite second order moments on $\mathbf{R}^d$ is the composition of a martingale coupling with an optimal transport map $\mathcal{T}$. They check the existence of optimal couplings in which this map gives the unique optimal coupling between μ and $\mathcal{T}\#\mu$. Next, they prove that $\sigma \mapsto W_2^2(\sigma,\nu)$ is differentiable at μ in both Lions and the geometric senses iff there is a unique optimal coupling between μ and ν and this coupling is given by a map.

7.2.2. Numerical methods for optimal transport.

Optimal transport problems have got a recent attention in many different fields including physics, quantum chemistry and finance, where Martingale Optimal Transport problems allow to quantify the model risk. In practice, few numerical methods exist to approximate the optimal coupling measure and/or the optimal transport. In particular, to deal with large dimensions or with the optimal transport problems with many marginal laws, a natural direction is to develop Monte-Carlo methods.

A. Alfonsi, V. Ehrlacher (CERMICS, Inria Project-team MATHERIALS), D. Lombardi (Inria Project-team Reo) and R. Coyaud (PhD student of A. Alfonsi) are working on numerical approximations of the optimal transport between two (or more) probability measures.

7.3. Optimal Control of Mean field (S)PDEs

With Rui Chen and R. Dumitrescu, A. Sulem has studied mean-field Backward SDEs driven by a Brownian motion and an independant Poisson random measure and its interpretation in terms of global risk measures. Dual representation has been provided in the convex case. Optimal stopping for these BSDEs and links with reflected mean-field BSDEs has also been investigated.

A. Sulem, R. Dumitrescu and B. Øksendal have studied optimal control for mean-field stochastic **partial** differential equations (stochastic evolution equations) driven by a Brownian motion and an independent Poisson random measure, in the case of *partial information* control [20]. One important novelty is the introduction of *general mean-field* operators, acting on both the controlled state process and the control process. A sufficient and a necessary maximum principle for this type of control is formulated. Existence and uniqueness of the solution of such general forward and backward mean-field stochastic partial differential equations are proved. These results have been applied to find the explicit optimal control for an optimal harvesting problem.

7.4. Analysis of probabilistic numerical methods

7.4.1. Particles approximation of mean-field SDEs

O. Bencheikh and Benjamin Jourdain have proved that the weak error between a stochastic differential equation with nonlinearity in the sense of McKean given by moments and its approximation by the Euler discretization with time-step h of a system of N interacting particles is $\mathcal{O}(N^{-1}+h)$. Numerical experiments confirm this behaviour and show that it extends to more general mean-field interaction.

7.4.2. Approximation of Markov processes

V. Bally worked on general approximation schemes in total variation distance for diffusion processes in collaboration with his former Phd student Clément Rey [52] This work includes high order schemes as Victoir-Ninomya for example. Further development in this direction is under study in collaboration with A. Alfonsi. Moreover, in collaboration with his former Phd student V. Rabiet and with D. Goreac (University Paris Est Marne la Vallée), V. Bally is studying approximations schemes for Piecewise Deterministic Markov Processes (see [17], [51]). In this framework the goal is to replace small jumps by a Brownien component - such a procedure is popular for "usual" jump equations, but the estimate of the error in the case of PDMP's is much more delicate. A significant example is the Bolzmann equation [28].

7.4.3. High order approximation for diffusion processes

A. Alfonsi and V. Bally are working on a generic method to achieve any weak order of convergence for approximating SDEs.

7.4.4. Adaptive MCMC methods

The Self-Healing Umbrella Sampling (SHUS) algorithm is an adaptive biasing algorithm which has been proposed in order to efficiently sample a multimodal probability measure.

In [21], G. Fort, B. Jourdain, T. Lelièvre and G. Stoltz extend previous works [68], [66], [67] and study a larger class of algorithms where the target distribution is biased using only a fraction of the free energy and which includes a discrete version of well-tempered metadynamics.

MOKAPLAN Project-Team

6. New Results

6.1. Rank optimality for the Burer-Monteiro factorization

I. Waldspurger, A. Watersw

In [39], Numerically solving a large scale semidefinite program, in full generality, is a challenge: The complexity of generic algorithms blows up quickly with the size of the unknown matrix. Fortunately, in many situations, the solution of the program has low rank, and this can be exploited to achieve algorithmic speedups. The most classical way to do this is the Burer-Monteiro factorization, introduced in [77]. It consists in writing the unknown matrix as the product of low-rank factors, and optimizing the factors instead of the matrix itself. The first theoretical guarantees for this method appeared in [69], where it was shown that this strategy almost always succeeds when the size of the factors is of the order of the square root of the full matrix. In our article, we show that, up to a marginal improvement, this result is optimal: Contrarily to what numerical experiments might suggest, there exist situations where the method fails if the size of the factors is chosen smaller.

6.2. Representer theorems in variational problems

C. Boyer, A. Chambolle, Y. De Castro, V. Duval, F. De Gournay, P. Weiss

In [29], we have established a general principle which states that regularizing an inverse problem with a convex function yields solutions which are convex combinations of a small number of *atoms*. These atoms are identified with the extreme points and elements of the extreme rays of the regularizer level sets. An extension to a broader class of quasi-convex regularizers is also discussed. As a side result, we characterize the minimizers of the total gradient variation, describing the solutions of total variation problem as a superposition of indicator functions of simply connected sets. That result provides an explanation of the so-called *staircasing* phenomenon.

6.3. The Sliding Frank-Wolfe algorithm for Super-resolution Microscopy Imaging

Q. Denoyelle, V. Duval, G. Peyré, E. Soubies

In [32], we have studied the theoretical and numerical performance of the Sliding Frank-Wolfe, a novel optimization algorithm to solve the BLASSO sparse spikes super-resolution problem. The BLASSO is a continuous (*i.e.* off-the-grid or grid-less) counterpart to the well-known ℓ^1 sparse regularisation method (also known as LASSO or Basis Pursuit). Our algorithm is a variation on the classical Frank-Wolfe (also known as conditional gradient) which follows a recent trend of interleaving convex optimization updates (corresponding to adding new spikes) with non-convex optimization steps (corresponding to moving the spikes). Our main theoretical result is that this algorithm terminates in a finite number of steps under a mild non-degeneracy hypothesis. We then target applications of this method to several instances of single molecule fluorescence imaging modalities, among which certain approaches rely heavily on the inversion of a Laplace transform. Our second theoretical contribution is the proof of the exact support recovery property of the BLASSO to invert the 1-D Laplace transform in the case of positive spikes. On the numerical side, we conclude this paper with an extensive study of the practical performance of the Sliding Frank-Wolfe on different instantiations of single molecule fluorescence imaging, including convolutive and non-convolutive (Laplace-like) operators. This shows the versatility and superiority of this method with respect to alternative sparse recovery techniques.

6.4. Approximation of variational problems with a convexity constraint by PDEs of Abreu type

G. Carlier, T. Radice

In [31], motivated by some variational problems subject to a convexity constraint, we consider an approximation using the logarithm of the Hessian determinant as a barrier for the constraint. We show that the minimizer of this penalization can be approached by solving a second boundary value problem for Abreu's equation which is a well-posed nonlinear fourth-order elliptic problem. More interestingly, a similar approximation result holds for the initial constrained variational problem.

6.5. Variational methods for tomographic reconstruction with few views

M. Bergounioux, I. Abraham, R. Abraham, G. Carlier, E. Le Pennec, E. Trélat

In [16], we deal with a severe ill posed problem, namely the reconstruction process of an image during tomography acquisition with (very) few views. We present different methods that we investigated during the past decade. They are based on variational analysis. This is a survey paper and we refer to the quoted papers for more details.

6.6. A differential approach to the multi-marginal Schrödinger system

G. Carlier, M. Laborde

In [30], we develop an elementary and self-contained differential approach, in an L^∞ setting, for well-posedness (existence, uniqueness and smooth dependence with respect to the data) for the multi-marginal Schrödinger system which arises in the entropic regularization of optimal transport problems.

6.7. Minimal convex extensions and finite difference discretization of the quadratic Monge-Kantorovich problem

J-D. Benamou, V. Duval

In [15] we present an adaptation of the MA-LBR scheme to the Monge-Ampr`e equation with second boundary value condition, provided the target is a convex set. This yields a fast adaptive method to numerically solve the Optimal Transport problem between two absolutely continuous measures, the second of which has convex support. The proposed numerical method actually captures a specific Brenier solution which is minimal in some sense. We prove the convergence of the method as the grid stepsize vanishes and we show with numerical experiments that it is able to reproduce subtle properties of the Optimal Transport problem.

6.8. Second order models for optimal transport and cubic splines on the Wasserstein space

J-D. Benamou, T. O. Gallouët, F-X. Vialard

On the space of probability densities, we extend in [28] the Wasserstein geodesics to the case of higher-order interpolation such as cubic spline interpolation. After presenting the natural extension of cubic splines to the Wasserstein space, we propose a simpler approach based on the relaxation of the variational problem on the path space. We explore two different numerical approaches, one based on multi-marginal optimal transport and entropic regularization and the other based on semi-discrete optimal transport.

6.9. An entropy minimization approach to second-order variational mean-field games

J-D. Benamou, G. Carlier, S. Di Marino, L. Nenna

In [26] we propose a new viewpoint on variational mean-field games with diffusion and quadratic Hamiltonian. We show the equivalence of such mean-field games with a relative entropy minimization at the level of probabilities on curves. We also address the time-discretization of such problems, establish Gamma-Convergence results as the time step vanishes and propose an efficient algorithm relying on this entropic interpretation as well as on the Sinkhorn scaling algorithm.

6.10. Generalized incompressible flows, multi-marginal transport and Sinkhorn algorithm

J-D. Benamou, G. Carlier, L. Nenna

Starting from Brenier's relaxed formulation of the incompress-ible Euler equation in terms of geodesics in the group of measure-preserving diffeomorphisms, we propose in [27] a numerical method based on Sinkhorn's algorithm for the entropic regularization of optimal transport. We also make a detailed comparison of this entropic regular-ization with the so-called Bredinger entropic interpolation problem (see [1]). Numerical results in dimension one and two illustrate the feasibility of the method.

6.11. Testing Gaussian Process with Applications to Super-Resolution

J.-M. Azaïs, Y. De Castro, S. Mourareau

In [13], we introduce exact testing procedures on the mean of a Gaussian process X derived from the outcomes of ℓ_1-minimization over the space of complex valued measures. The process X can be thought as the sum of two terms: first, the convolution between some kernel and a target atomic measure (mean of the process); second, a random perturbation by an additive centered Gaussian process. The first testing procedure considered is based on a dense sequence of grids on the index set of X and we establish that it converges (as the grid step tends to zero) to a randomized testing procedure: the decision of the test depends on the observation X and also on an independent random variable. The second testing procedure is based on the maxima and the Hessian of X in a grid-less manner. We show that both testing procedures can be performed when the variance is unknown (and the correlation function of X is known). These testing procedures can be used for the problem of deconvolution over the space of complex valued measures, and applications in frame of the Super-Resolution theory are presented. As a byproduct, numerical investigations may demonstrate that our grid-less method is more powerful (it detects sparse alternatives) than tests based on very thin grids.

6.12. Approximate Optimal Designs for Multivariate Polynomial Regression

Y. De Castro, F. Gamboa, D. Henrion, R. Hess, J.-B Lasserre

In [19], we introduce a new approach aiming at computing approximate optimal designs for multivariate polynomial regressions on compact (semi-algebraic) design spaces. We use the moment-sum-of-squares hierarchy of semidefinite programming problems to solve numerically the approximate optimal design problem. The geometry of the design is recovered via semidefinite programming duality theory. This article shows that the hierarchy converges to the approximate optimal design as the order of the hierarchy increases. Furthermore, we provide a dual certificate ensuring finite convergence of the hierarchy and showing that the approximate optimal design can be computed numerically with our method. As a byproduct, we revisit the equivalence theorem of the experimental design theory: it is linked to the Christoffel polynomial and it characterizes finite convergence of the moment-sum-of-square hierarchies.

6.13. Simulation of multiphase porous media flows with minimizing movement and finite volume schemes

C .Cancès, T. O. Gallouët, M. Laborde, L. Monsaingeon

In [17]: the Wasserstein gradient flow structure of the PDE system governing multiphase flows in porous media was recently highlighted in [85]. The model can thus be approximated by means of the minimizing movement (or JKO) scheme. We solve the JKO scheme using the ALG2-JKO scheme proposed in [55]. The numerical results are compared to a classical upstream mobility Finite Volume scheme, for which strong stability properties can be established.

6.14. An unbalanced optimal transport splitting scheme for general advection-reaction-diffusion problems

T. O. Gallouët, M. Laborde, L. Monsaingeon

In [21] the authors show that unbalanced optimal transport provides a convenient framework to handle reaction and diffusion processes in a unified metric framework. We use a constructive method, alternating minimizing movements for the Wasserstein distance and for the Fisher-Rao distance, and prove existence of weak solutions for general scalar reaction-diffusion-advection equations. We extend the approach to systems of multiple interacting species, and also consider an application to a very degenerate diffusion problem involving a Gamma-limit. Moreover, some numerical simulations are included.

6.15. Generalized compressible fluid flows and solutions of the Camassa-Holm variational model

T. O. Gallouët, A. Natale, F-X. Vialard

In [35] : The Camassa-Holm equation on a domain M $\in \mathbb{R}^d$, in one of its possible multi-dimensional generalizations, describes geodesics on the group of diffeomorphisms with respect to the H(div) metric. It has been recently reformulated as a geodesic equation for the L2 metric on a subgroup of the diffeomorphism group of the cone over M. We use such an interpretation to construct an analogue of Brenier's generalized incompressible Euler flows for the Camassa-Holm equation. This involves describing the fluid motion using probability measures on the space of paths on the cone, so that particles are allowed to split and cross. Differently from Brenier's model, however, we are also able to account for compressibility by employing an explicit probabilistic representation of the Jacobian of the flow map. We formulate the boundary value problem associated to the Camassa-Holm equation using such generalized flows. We prove existence of solutions and that, for short times, smooth solutions of the Camassa-Holm equations are the unique solutions of our model. We propose a numerical scheme to construct generalized solutions on the cone and present some numerical results illustrating the relation between the generalized Camassa-Holm and incompressible Euler solutions.

6.16. The Camassa-Holm equation as an incompressible Euler equation: a geometric point of view

T. O. Gallouët, F-X. Vialard

In [23]: The group of diffeomorphisms of a compact manifold endowed with the L^2 metric acting on the space of probability densities gives a unifying framework for the incompressible Euler equation and the theory of optimal mass transport. Recently, several authors have extended optimal transport to the space of positive Radon measures where the Wasserstein-Fisher-Rao distance is a natural extension of the classical L^2-Wasserstein distance. In this paper, we show a similar relation between this unbalanced optimal transport problem and the Hdiv right-invariant metric on the group of diffeomorphisms, which corresponds to the Camassa-Holm (CH) equation in one dimension. On the optimal transport side, we prove a polar factorization theorem on the automorphism group of half-densities. Geometrically, our point of view provides an isometric embedding of the group of diffeomorphisms endowed with this right-invariant metric in the automorphisms group of the fiber bundle of half densities endowed with an L^2 type of cone metric. This leads to a new formulation of the (generalized) CH equation as a geodesic equation on an isotropy subgroup of this automorphisms group; On S_1, solutions to the standard CH thus give particular solutions of the incompressible Euler equation on a group of homeomorphisms of $\mathbb{R}^2$ which preserve a radial density that has a singularity at 0. An other application consists in proving that smooth solutions of the Euler-Arnold equation for the Hdiv right-invariant metric are length minimizing geodesics for sufficiently short times.

6.17. Variational Second-Order Interpolation on the Group of Diffeomorphisms with a Right-Invariant Metric

F-X. Vialard

In [38] we propose a variational framework in which the minimization of the acceleration on the group of diffeomorphisms endowed with a right-invariant metric is well-posed. It relies on constraining the acceleration to belong to a Sobolev space of higher-order than the order of the metric in order to gain compactness. It provides the theoretical guarantee of existence of minimizers which is compulsory for numerical simulations.

6.18. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences

J. Feydy, T. Séjourné, F.X. Vialard, S-I. Amari, A. Trouvé, G. Peyré

In [33]: Comparing probability distributions is a fundamental problem in data sciences. Simple norms and divergences such as the total variation and the relative entropy only compare densities in a point-wise manner and fail to capture the geometric nature of the problem. In sharp contrast, Maximum Mean Discrepancies (MMD) and Optimal Transport distances (OT) are two classes of distances between measures that take into account the geometry of the underlying space and metrize the convergence in law. This paper studies the Sinkhorn divergences, a family of geometric divergences that interpolates between MMD and OT. Relying on a new notion of geometric entropy, we provide theoretical guarantees for these divergences: positivity, convexity and metrization of the convergence in law. On the practical side, we detail a numerical scheme that enables the large scale application of these divergences for machine learning: on the GPU, gradients of the Sinkhorn loss can be computed for batches of a million samples.

QUANTIC Project-Team

6. New Results

6.1. Simulation of quantum walks and fast mixing with classical processes

Participants: A. Sarlette

This is the final result of a line of work where we show that the mixing behavior of quantum walks on graphs can always be simulated by a classical "lifted Markov chain". This implies that quantum walks must satisfy a conductance bound on mixing speed, like classical Markov chains. Also current efficient quantum walk constructions are linked to classical processes that provide the same convergence speed. This excludes a simple characterization of quantum walk advantages in terms of bare mixing speed, as has been done by some previous authors comparing just to simple Markov chains. The question of efficient design of walks on graphs, on the basis of local graph queries and for specific applications, is thus brought back to the center of the focus for quantum walks. This collaborative work with F. Ticozzi (U. of Padova) has been published in [11].

As a follow-up on this work, we have developed algorithms in the latter sense: quantum walks on the basis of local design and which do speed up some applications. These last results have been presented as posters at conferences and will hopefully be part of next year's publications.

6.2. Adiabatic elimination for multi-partite open quantum systems with non-trivial zero-order dynamics

Participants: Paolo Forni, Alain Sarlette, Pierre Rouchon

We pursue the work initiated in our group during the thesis of Rémi Azouit, where we apply center manifold theory in order to reduce the model of a quantum system to its slowly contracting dynamics. Such model reduction is ubiquitous in models of coupled quantum systems where part of the system relaxes quickly towards an equilibrium situation, and acts as an environment for a system of interest. The extension presented in this work is the answer to a question by experimental physicists at Laboratoire Kastler Brossel (LKB), where they apply a strong drive which, in an 'intuitive model', would saturate so-called two-level-system impurities and thereby imply a particular behavior of frequency shift and dissipation on the target system (slow dynamics) as a function of drive characteristics. A good model for this situation involves, beyond a strongly dissipative environment, also a fast non-dissipative dynamics on the slowly contracting subsystem. Adding the latter into the model reduction was the purpose of this result. We analyze the experimental results and show that the model reduction allows us to explain the observed trends. This result led to a publication in collaboration with physicists Thibault Capelle, Emmanuel Flurin and Samule Deleglise from LKB [20].

Further extensions of adiabatic elimination formulas have been worked out during this year and will hopefully be part of next year's publications.

6.3. Exponential stochastic stabilization of a two-level quantum system via strict Lyapunov control

Participants: Gerardo Cardona, Alain Sarlette, Pierre Rouchon

In this result, we address the fundamental task of stabilizing the state of a quantum system towards a target eigenstate of a continuous-time quantum nondemolition measurement. The starting point is that a static output feedback does not allow us to stabilize this system, while more complicated procedures were not able to provide a convergence rate. Our main idea is to introduce a dynamic feedback controller of moderate complexity, where (i) feedback gains depend on estimated state and progressively go to zero as one approaches the target; and (ii) the feedback involves noise (in this paper from the measurement back-action but in further extensions possibly just independent noise). With this controller we show, providing a Lyapunov function close to the Bures distance measure, that the system converges exponentially towards the target eigenstate. This result, restricted to a proof-of-principle on the qubit, was published in [19].

This has laid the basis for further work, presented on posters and to be published next year, where we have shown that:

- the optimal convergence rate, equal to information gain, can be achieved with this feedback;

- the procedure extends to N-level systems, with noise just independent instead of coming from the measurement backaction;

- the procedure can be exploited towards continuous-time measurement-based quantum error correction

6.4. Structural instability of driven Josephson circuits prevented by an inductive shunt

Participants: Lucas Verney, Raphaël Lescanne, Zaki Leghtas, Mazyar Mirrahimi.

Superconducting circuits are a versatile platform to implement a multitude of Hamiltonians which perform quantum computation, simulation and sensing tasks. A key ingredient for realizing a desired Hamiltonian is the irradiation of the circuit by a strong drive. These strong drives provide an insitu control of couplings, which cannot be obtained by near-equilibrium Hamiltonians. However, as shown in our result, out-of-equilibrium systems are easily plagued by complex dynamics leading to instabilities. Predicting and preventing these instabilities is crucial, both from a fundamental and application perspective. We propose an inductively shunted transmon as the elementary circuit optimized for strong parametric drives. Developing a novel numerical approach that avoids the built-in limitations of perturbative analysis, we demonstrate that adding the inductive shunt significantly extends the range of pump powers over which the circuit behaves in a stable manner. This collaborative work between the Quantic team and Michel Devoret at Yale has been recently submitted for publication [25].

6.5. Observing the escape of a driven quantum Josephson circuit into unconfined states

Participants: Raphaël Lescanne, Lucas Verney, Mazyar Mirrahimi, Zaki Leghtas.

Josephson circuits have been ideal systems to study complex non-linear dynamics which can lead to chaotic behavior and instabilities. More recently, Josephson circuits in the quantum regime, particularly in the presence of microwave drives, have demonstrated their ability to emulate a variety of Hamiltonians that are useful for the processing of quantum information. In this experimental work, we show that these drives lead to an instability which results in the escape of the circuit mode into states that are not confined by the Josephson cosine potential. We observe this escape in a ubiquitous circuit: a transmon embedded in a 3D cavity. When the transmon occupies these free-particle-like states, the circuit behaves as though the junction had been removed, and all non-linearities are lost. This work deepens our understanding of strongly driven Josephson circuits, which is important for fundamental and application perspectives, such as the engineering of Hamiltonians by parametric pumping. This collaborative work between Quantic team, Benjamin Huard's team at ENS Lyon and Michel Devoret at Yale, has been recently submitted for publication [21].

6.6. Dynamics of a qubit while simultaneously monitoring its relaxation and dephasing

Participants: Zaki Leghtas.

Decoherence originates from the leakage of quantum information into external degrees of freedom. For a qubit, the two main decoherence channels are relaxation and dephasing. Here, we report an experiment on a superconducting qubit where we retrieve part of the lost information in both of these channels. We demonstrate that raw averaging of the corresponding measurement records provides a full quantum tomography of the qubit state where all three components of the effective spin-1/2 are simultaneously measured. From single realizations of the experiment, it is possible to infer the quantum trajectories followed by the qubit state conditioned on relaxation and/or dephasing channels. The incompatibility between these quantum measurements of the qubit leads to observable consequences in the statistics of quantum states. The high level of controllability of superconducting circuits enables us to explore many regimes from the Zeno effect to underdamped Rabi oscillations depending on the relative strengths of driving, dephasing, and relaxation. This work is a collaboration between the Quantic team and the group of Benjamin Huard at ENS Lyon and was published in [13].

6.7. Demonstration of an effective ultrastrong coupling between two oscillators

Participants: Zaki Leghtas

When the coupling rate between two quantum systems becomes as large as their characteristic frequencies, it induces dramatic effects on their dynamics and even on the nature of their ground state. The case of a qubit coupled to a harmonic oscillator in this ultrastrong coupling regime has been investigated theoretically and experimentally. Here, we explore the case of two harmonic oscillators in the ultrastrong coupling regime. Probing the properties of their ground state remains out of reach in natural implementations. Therefore, we have realized an analog quantum simulation of this coupled system by dual frequency pumping a nonlinear superconducting circuit. The pump amplitudes directly tune the effective coupling rate. We observe spectroscopic signature of a mode hybridization that is characteristic of the ultrastrong coupling. We experimentally demonstrate a key property of the ground state of this simulated ultrastrong coupling between modes by observing simultaneous single- and two-mode squeezing of the radiated field below vacuum fluctuations. This work is a collaboration between the Quantic team and the group of Benjamin Huard at ENS Lyon and was published in [14].

6.8. Fault-tolerant detection of a quantum error

Participants: Mazyar Mirrahimi

A critical component of any quantum error–correcting scheme is detection of errors by using an ancilla system. However, errors occurring in the ancilla can propagate onto the logical qubit, irreversibly corrupting the encoded information. We experimentally demonstrate a fault-tolerant error-detection scheme that suppresses spreading of ancilla errors by a factor of 5, while maintaining the assignment fidelity. The same method is used to prevent propagation of ancilla excitations, increasing the logical qubit dephasing time by an order of magnitude. Our approach is hardware-efficient, as it uses a single multilevel transmon ancilla and a cavity-encoded logical qubit, whose interaction is engineered in situ by using an off-resonant sideband drive. The results demonstrate that hardware-efficient approaches that exploit system-specific error models can yield advances toward fault-tolerant quantum computation. This work is a collaboration between the Quantic team and the group of Robert Schoelkopf at Yale university and was published in [17].

6.9. Coherent oscillations inside a quantum manifold stabilized by dissipation

Participants: Zaki Leghtas, Mazyar Mirrahimi

Manipulating the state of a logical quantum bit usually comes at the expense of exposing it to decoherence. Fault-tolerant quantum computing tackles this problem by manipulating quantum information within a stable manifold of a larger Hilbert space, whose symmetries restrict the number of independent errors. The remaining errors do not affect the quantum computation and are correctable after the fact. Here we implement the autonomous stabilization of an encoding manifold spanned by Schrödinger cat states in a superconducting cavity. We show Zeno-driven coherent oscillations between these states analogous to the Rabi rotation of a qubit protected against phase flips. Such gates are compatible with quantum error correction and hence are crucial for fault-tolerant logical qubits. This experimental work follows our previous theoretical proposal [70]. It is a collaboration between the Quantic team and the group of Michel Devoret at Yale university and was published in [18].

6.10. To catch and reverse a quantum jump mid-flight

Participants: Mazyar Mirrahimi

A quantum system driven by a weak deterministic force while under strong continuous energy measurement exhibits quantum jumps between its energy levels. This celebrated phenomenon is emblematic of the special nature of randomness in quantum physics. The times at which the jumps occur are reputed to be fundamentally unpredictable. However, certain classical phenomena, like tsunamis, while unpredictable in the long term, may possess a degree of predictability in the short term, and in some cases it may be possible to prevent a disaster by detecting an advance warning signal. Can there be, despite the indeterminism of quantum physics, a possibility to know if a quantum jump is about to occur or not? We answer this question affirmatively by experimentally demonstrating that the completed jump from the ground to an excited state of a superconducting artificial atom can be tracked, as it follows its predictable "flight," by monitoring the population of an auxiliary level coupled to the ground state. Furthermore, we show that the completed jump is continuous, deterministic, and coherent. Exploiting this coherence, we catch and reverse a quantum jump mid-flight, thus preventing its completion. This real-time intervention is based on a particular lull period in the population of the auxiliary level, which serves as our advance warning signal. Our experimental results, which agree with theoretical predictions essentially without adjustable parameters, support the modern quantum trajectory theory and provide new ground for the exploration of real-time intervention techniques in the control of quantum systems, such as early detection of error syndromes. This work is a collaboration between the Quantic team and the group of Michel Devoret at Yale university and is recently submitted for publication [22].

6.11. Remote entanglement stabilization and concentration by quantum reservoir engineering

Participants: Nicolas Didier, Jérémie Guillaud, Mazyar Mirrahimi

Quantum information processing in a modular architecture requires the distribution, stabilization, and distillation of entanglement in a qubit network. We present autonomous entanglement stabilization protocols between two superconducting qubits that are coupled to distant cavities. The coupling between cavities is mediated and controlled via a three-wave mixing device that generates either a two-mode squeezed state or a delocalized mode between the remote cavities depending on the pump applied to the mixer. Local drives on the qubits and the cavities steer and maintain the system to the desired qubit Bell state. Most spectacularly, even a weakly squeezed state can stabilize a maximally entangled Bell state of two distant qubits through an autonomous entanglement concentration process. Moreover, we show that such reservoir-engineering-based protocols can stabilize entanglement in the presence of qubit-cavity asymmetries and losses. This work was published in [12].

SIERRA Project-Team

7. New Results

7.1. On the Global Convergence of Gradient Descent for Over-parameterized Models using Optimal Transport

Many tasks in machine learning and signal processing can be solved by minimizing a convex function of a measure. This includes sparse spikes deconvolution or training a neural network with a single hidden layer. For these problems, in [25] we study a simple minimization method: the unknown measure is discretized into a mixture of particles and a continuous-time gradient descent is performed on their weights and positions. This is an idealization of the usual way to train neural networks with a large hidden layer. We show that, when initialized correctly and in the many-particle limit, this gradient flow, although non-convex, converges to global minimizers. The proof involves Wasserstein gradient flows, a by-product of optimal transport theory. Numerical experiments show that this asymptotic behavior is already at play for a reasonable number of particles, even in high dimension.

7.2. Sharp Analysis of Learning with Discrete Losses

In [49], we study a least-squares framework to systematically design learning algorithms for discrete losses, with quantitative characterizations in terms of statistical and computational complexity. In particular we improve existing results by providing explicit dependence on the number of labels for a wide class of losses and faster learning rates in conditions of low-noise. Theoretical results are complemented with experiments on real datasets, showing the effectiveness of the proposed general approach.

7.3. Gossip of Statistical Observations using Orthogonal Polynomials

Consider a network of agents connected by communication links, where each agent holds a real value. The gossip problem consists in estimating the average of the values diffused in the network in a distributed manner. Current techniques for gossiping are designed to deal with worst-case scenarios, which is irrelevant in applications to distributed statistical learning and denoising in sensor networks. In [39], we design second-order gossip methods tailor-made for the case where the real values are i.i.d. samples from the same distribution. In some regular network structures, we are able to prove optimality of our methods, and simulations suggest that they are efficient in a wide range of random networks. Our approach of gossip stems from a new acceleration framework using the family of orthogonal polynomials with respect to the spectral measure of the network graph.

7.4. Marginal Weighted Maximum Log-likelihood for Efficient Learning of Perturb-and-Map models

In [20], We consider the structured-output prediction problem through probabilistic approaches and generalize the "perturb-and-MAP" framework to more challenging weighted Hamming losses, which are crucial in applications. While in principle our approach is a straightforward marginalization, it requires solving many related MAP inference problems. We show that for log-supermodular pairwise models these operations can be performed efficiently using the machinery of dynamic graph cuts. We also propose to use double stochastic gradient descent, both on the data and on the perturbations, for efficient learning. Our framework can naturally take weak supervision (e.g., partial labels) into account. We conduct a set of experiments on medium-scale character recognition and image segmentation, showing the benefits of our algorithms.

7.5. Slice inverse regression with score functions

In [6], we consider non-linear regression problems where we assume that the response depends non-linearly on a linear projection of the covariates. We propose score function extensions to sliced inverse regression problems, both for the first- order and second-order score functions. We show that they provably improve estimation in the population case over the non-sliced versions and we study finite sample estimators and their consistency given the exact score functions. We also propose to learn the score function as well, 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. We illustrate our results on a series of experiments.

7.6. Constant Step Size Stochastic Gradient Descent for Probabilistic Modeling

Stochastic gradient methods enable learning probabilistic models from large amounts of data. While large step-sizes (learning rates) have shown to be best for least-squares (e.g., Gaussian noise) once combined with parameter averaging, these are not leading to convergent algorithms in general. In this paper, we consider generalized linear models, that is, conditional models based on exponential families. In [14], we propose averaging moment parameters instead of natural parameters for constant-step-size stochastic gradient descent. For finite-dimensional models, we show that this can sometimes (and surprisingly) lead to better predictions than the best linear model. For infinite-dimensional models, we show that it always converges to optimal predictions, while averaging natural parameters never does. We illustrate our findings with simulations on synthetic data and classical benchmarks with many observations.

7.7. Nonlinear Acceleration of Momentum and Primal-Dual Algorithms

In [40], We describe a convergence acceleration scheme for multistep optimization algorithms. The extrapolated solution is written as a nonlinear average of the iterates produced by the original optimization algorithm. Our scheme does not need the underlying fixed-point operator to be symmetric, hence handles e.g. algorithms with momentum terms such as Nesterov's accelerated method, or primal-dual methods. The weights are computed via a simple linear system and we analyze performance in both online and offline modes. We use Crouzeix's conjecture to show that acceleration performance is controlled by the solution of a Chebyshev problem on the numerical range of a non-symmetric operator modelling the behavior of iterates near the optimum. Numerical experiments are detailed on image processing problems, logistic regression and neural network training for CIFAR10 and ImageNet.

7.8. Nonlinear Acceleration of Deep Neural Networks

Regularized nonlinear acceleration (RNA) is a generic extrapolation scheme for optimization methods, with marginal computational overhead. It aims to improve convergence using only the iterates of simple iterative algorithms. However, so far its application to optimization was theoretically limited to gradient descent and other single-step algorithms. Here, we adapt RNA to a much broader setting including stochastic gradient with momentum and Nesterov's fast gradient. In [54], we use it to train deep neural networks, and empirically observe that extrapolated networks are more accurate, especially in the early iterations. A straightforward application of our algorithm when training ResNet-152 on ImageNet produces a top-1 test error of 20.88, improving by 0.8 the reference classification pipeline. Furthermore, the code runs offline in this case, so it never negatively affects performance.

7.9. Nonlinear Acceleration of CNNs

The Regularized Nonlinear Acceleration (RNA) algorithm is an acceleration method capable of improving the rate of convergence of many optimization schemes such as gradient descend, SAGA or SVRG. Until now, its analysis is limited to convex problems, but empirical observations shows that RNA may be extended to wider settings. In [36], we investigate further the benefits of RNA when applied to neural networks, in particular for the task of image recognition on CIFAR10 and ImageNet. With very few modifications of exiting frameworks, RNA improves slightly the optimization process of CNNs, after training.

7.10. Robust Seriation and Applications To Cancer Genomics

The seriation problem seeks to reorder a set of elements given pairwise similarity information, so that elements with higher similarity are closer in the resulting sequence. When a global ordering consistent with the similarity information exists, an exact spectral solution recovers it in the noiseless case and seriation is equivalent to the combinatorial 2-SUM problem over permutations, for which several relaxations have been derived. However, in applications such as DNA assembly, similarity values are often heavily corrupted, and the solution of 2-SUM may no longer yield an approximate serial structure on the elements. In [52], we introduce the robust seriation problem and show that it is equivalent to a modified 2-SUM problem for a class of similarity matrices modeling those observed in DNA assembly. We explore several relaxations of this modified 2-SUM problem and compare them empirically on both synthetic matrices and real DNA data. We then introduce the problem of seriation with duplications, which is a generalization of Seriation motivated by applications to cancer genome reconstruction. We propose an algorithm involving robust seriation to solve it, and present preliminary results on synthetic data sets.

7.11. Reconstructing Latent Orderings by Spectral Clustering

Spectral clustering uses a graph Laplacian spectral embedding to enhance the cluster structure of some data sets. When the embedding is one dimensional, it can be used to sort the items (spectral ordering). A number of empirical results also suggests that a multidimensional Laplacian embedding enhances the latent ordering of the data, if any. This also extends to circular orderings, a case where unidimensional embeddings fail. In [51], we tackle the task of retrieving linear and circular orderings in a unifying framework, and show how a latent ordering on the data translates into a filamentary structure on the Laplacian embedding. We propose a method to recover it, illustrated with numerical experiments on synthetic data and real DNA sequencing data.

7.12. Lyapunov Functions for First-Order Methods: Tight Automated Convergence Guarantees

In [21], we present a novel way of generating Lyapunov functions for proving linear convergence rates of first-order optimization methods. Our approach provably obtains the fastest linear convergence rate that can be verified by a quadratic Lyapunov function (with given states), and only relies on solving a small-sized semidefinite program. Our approach combines the advantages of performance estimation problems and integral quadratic constraints, and relies on convex interpolation.

7.13. Efficient First-order Methods for Convex Minimization: a Constructive Approach

In [44], we describe a novel constructive technique for devising efficient first-order methods for a wide range of large-scale convex minimization settings, including smooth, non-smooth, and strongly convex minimization. The design technique takes a method performing a series of subspace-searches and constructs a family of methods that share the same worst-case guarantees as the original method, and includes a fixed-step first-order method. We show that this technique yields optimal methods in the smooth and non-smooth cases and derive new methods for these cases, including methods that forego knowledge of the problem parameters, at the cost of a one-dimensional line search per iteration. In the strongly convex case, we show how numerical tools can be used to perform the construction, and show that resulting method offers an improved convergence rate compared to Nesterov's celebrated fast gradient method.

7.14. Operator Splitting Performance Estimation: Tight contraction factors and optimal parameter selection

In [53], we propose a methodology for studying the performance of common splitting methods through semidefinite programming. We prove tightness of the methodology and demonstrate its value by presenting

two applications of it. First, we use the methodology as a tool for computer-assisted proofs to prove tight analytical contraction factors for Douglas–Rachford splitting that are likely too complicated for a human to find bare-handed. Second, we use the methodology as an algorithmic tool to computationally select the optimal splitting method parameters by solving a series of semidefinite programs.

7.15. Finite-sample Analysis of M-estimators using Self-concordance

In [50], we demonstrate how *self-concordance* of the loss allows to obtain asymptotically optimal rates for M-estimators in finite-sample regimes. We consider two classes of losses: (i) self-concordant losses, i.e., whose third derivative is uniformly bounded with the $3/2$ power of the second; (ii) *pseudo* self-concordant losses, for which the power is removed. These classes contain some losses arising in generalized linear models, including the logistic loss; in addition, the second class includes some common pseudo-Huber losses. Our results consist in establishing the *critical sample size* sufficient to reach the asymptotically optimal excess risk in both cases. Denoting d the parameter dimension, and d_e the effective dimension taking into account possible model misspecification, we find the critical sample size to be $O(d_e \cdot d)$ for the first class of losses, and $O(\rho \cdot d_e \cdot d)$ for the second class, where ρ is the problem-dependent parameter that characterizes the risk curvature at the best predictor θ_*. In contrast to the existing results, we only impose *local* assumptions on the data distribution, assuming that the calibrated design, i.e., the design scaled with the square root of the second derivative of the loss, is subgaussian at the best predictor. Moreover, we obtain the improved bounds on the critical sample size, scaling *near-linearly* in $\max(d_e, d)$, under the extra assumption that the calibrated design is subgaussian in the Dikin ellipsoid of θ_*. Motivated by these findings, we construct canonically self-concordant analogues of the Huber and logistic losses with improved statistical properties. Finally, we extend some of the above results to ℓ_1-penalized M-estimators in high-dimensional setups.

7.16. Uniform regret bounds over R^d for the sequential linear regression problem with the square loss

In [45] we consider the setting of online linear regression for arbitrary deterministic sequences, with the square loss. We are interested in obtaining regret bounds that hold uniformly over all vectors R^d. When the feature sequence is known at the beginning of the game, they provided closed-form regret bounds of $2dB^2 \ln T + O(1)$, where T is the number of rounds and B is a bound on the observations. Instead, we derive bounds with an optimal constant of 1 in front of the $dB^2 \ln T$ term. In the case of sequentially revealed features, we also derive an asymptotic regret bound of $dB^2 \ln T$ for any individual sequence of features and bounded observations. All our algorithms are variants of the online nonlinear ridge regression forecaster, either with a data-dependent regularization or with almost no regularization.

7.17. Efficient online algorithms for fast-rate regret bounds under sparsity.

In [46] we consider the problem of online convex optimization in two different settings: arbitrary and i.i.d. sequence of convex loss functions. In both settings, we provide efficient algorithms whose cumulative excess risks are controlled with fast-rate sparse bounds. First, the excess risks bounds depend on the sparsity of the objective rather than on the dimension of the parameters space. Second, their rates are faster than the slow-rate $1/\sqrt{T}$

7.18. Exponential convergence of testing error for stochastic gradient methods

In [32], we consider binary classification problems with positive definite kernels and square loss, and study the convergence rates of stochastic gradient methods. We show that while the excess testing loss (squared loss) converges slowly to zero as the number of observations (and thus iterations) goes to infinity, the testing error (classification error) converges exponentially fast if low-noise conditions are assumed.

7.19. Statistical Optimality of Stochastic Gradient Descent on Hard Learning Problems through Multiple Passes

In [33], we consider stochastic gradient descent (SGD) for least-squares regression with potentially several passes over the data. While several passes have been widely reported to perform practically better in terms of predictive performance on unseen data, the existing theoretical analysis of SGD suggests that a single pass is statistically optimal. While this is true for low-dimensional easy problems, we show that for hard problems, multiple passes lead to statistically optimal predictions while single pass does not; we also show that in these hard models, the optimal number of passes over the data increases with sample size. In order to define the notion of hardness and show that our predictive performances are optimal, we consider potentially infinite-dimensional models and notions typically associated to kernel methods, namely, the decay of eigenvalues of the covariance matrix of the features and the complexity of the optimal predictor as measured through the covariance matrix. We illustrate our results on synthetic experiments with non-linear kernel methods and on a classical benchmark with a linear model.

7.20. Central Limit Theorem for stationary Fleming–Viot particle systems in finite spaces

In [11], we consider the Fleming–Viot particle system associated with a continuous-time Markov chain in a finite space. Assuming irreducibility, it is known that the particle system possesses a unique stationary distribution, under which its empirical measure converges to the quasistationary distribution of the Markov chain. We complement this Law of Large Numbers with a Central Limit Theorem. Our proof essentially relies on elementary computations on the infinitesimal generator of the Fleming–Viot particle system, and involves the so-called π-return process in the expression of the asymptotic variance. Our work can be seen as an infinite-time version, in the setting of finite space Markov chains, of results by Del Moral and Miclo [ESAIM: Probab. Statist., 2003] and Cérou, Delyon, Guyader and Rousset [arXiv:1611.00515, arXiv:1709.06771].

7.21. SeaRNN: Improved RNN training through Global-Local Losses

In [16], we propose SEARNN, a novel training algorithm for recurrent neural networks (RNNs) inspired by the "learning to search" (L2S) approach to structured prediction. RNNs have been widely successful in structured prediction applications such as machine translation or parsing, and are commonly trained using maximum likelihood estimation (MLE). Unfortunately, this training loss is not always an appropriate surrogate for the test error: by only maximizing the ground truth probability, it fails to exploit the wealth of information offered by structured losses. Further, it introduces discrepancies between training and predicting (such as exposure bias) that may hurt test performance. Instead, SEARNN leverages test-alike search space exploration to introduce global-local losses that are closer to the test error. We first demonstrate improved performance over MLE on two different tasks: OCR and spelling correction. Then, we propose a subsampling strategy to enable SEARNN to scale to large vocabulary sizes. This allows us to validate the benefits of our approach on a machine translation task.

7.22. Improved asynchronous parallel optimization analysis for stochastic incremental methods

As datasets continue to increase in size and multi-core computer architectures are developed, asynchronous parallel optimization algorithms become more and more essential to the field of Machine Learning. Unfortunately, conducting the theoretical analysis of asynchronous methods is difficult, notably due to the introduction of delay and inconsistency in inherently sequential algorithms. Handling these issues often requires resorting to simplifying but unrealistic assumptions. Through a novel perspective, in [10] we revisit and clarify 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 introduced "perturbed iterate" framework that resolves it. We demonstrate the usefulness of our new framework by analyzing three distinct

asynchronous parallel incremental optimization algorithms: Hogwild (asynchronous SGD), KROMAGNON (asynchronous SVRG) and ASAGA, a novel asynchronous parallel version of the incremental gradient algorithm SAGA that enjoys fast linear convergence rates. We are able to both remove problematic assumptions and obtain better theoretical results. Notably, we prove that ASAGA and KROMAGNON 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 speedups as well as the hardware overhead. Finally, we investigate the overlap constant, an ill-understood but central quantity for the theoretical analysis of asynchronous parallel algorithms. We find that it encompasses much more complexity than suggested in previous work, and often is order-of-magnitude bigger than traditionally thought.

7.23. Asynchronous optimisation for Machine Learning

The impressive breakthroughs of the last two decades in the field of machine learning can be in large part attributed to the explosion of computing power and available data. These two limiting factors have been replaced by a new bottleneck: algorithms. The focus of this thesis [3] is thus on introducing novel methods that can take advantage of high data quantity and computing power. We present two independent contributions.

First, we develop and analyze novel fast optimization algorithms which take advantage of the advances in parallel computing architecture and can handle vast amounts of data. We introduce a new framework of analysis for asynchronous parallel incremental algorithms, which enable correct and simple proofs. We then demonstrate its usefulness by performing the convergence analysis for several methods, including two novel algorithms.

Asaga is a sparse asynchronous parallel variant of the variance-reduced algorithm Saga which enjoys fast linear convergence rates on smooth and strongly convex objectives. We prove that it can be linearly faster than its sequential counterpart, even without sparsity assump- tions.

ProxAsaga is an extension of Asaga to the more general setting where the regularizer can be non-smooth. We prove that it can also achieve a linear speedup. We provide extensive experiments comparing our new algorithms to the current state-of-art.

Second, we introduce new methods for complex struc- tured prediction tasks. We focus on recurrent neural net- works (RNNs), whose traditional training algorithm for RNNs – based on maximum likelihood estimation (MLE) – suffers from several issues. The associated surrogate training loss notably ignores the information contained in structured losses and introduces discrepancies between train and test times that may hurt performance.

To alleviate these problems, we propose SeaRNN, a novel training algorithm for RNNs inspired by the "learning to search" approach to structured prediction.SeaRNN leverages test-alike search space exploration to introduce global-local losses that are closer to the test error than the MLE objective.

We demonstrate improved performance over MLE on three challenging tasks, and provide several subsampling strategies to enable SeaRNN to scale to large-scale tasks, such as machine translation. Finally, after contrasting the behavior of SeaRNN models to MLE models, we conduct an in-depth comparison of our new approach to the related work.

7.24. M^*-Regularized Dictionary Learning

In [38], we derive a performance measure for dictionaries in compressed sensing, based on the M^* of the corresponding norm. We use this measure to regularize dictionary learning algorithms and study the performance of our methods on both compression and inpainting experiments.

7.25. Optimal Algorithms for Non-Smooth Distributed Optimization in Networks

In [35], we consider the distributed optimization of non-smooth convex functions using a network of computing units. We investigate this problem under two regularity assumptions: (1) the Lipschitz continuity

of the global objective function, and (2) the Lipschitz continuity of local individual functions. Under the local regularity assumption, we provide the first optimal first-order decentralized algorithm called multi-step primal-dual (MSPD) and its corresponding optimal convergence rate. A notable aspect of this result is that, for non-smooth functions, while the dominant term of the error is in $O(1/\sqrt{t})$, the structure of the communication network only impacts a second-order term in $O(1/t)$, where t is time. In other words, the error due to limits in communication resources decreases at a fast rate even in the case of non-strongly-convex objective functions. Under the global regularity assumption, we provide a simple yet efficient algorithm called distributed randomized smoothing (DRS) based on a local smoothing of the objective function, and show that DRS is within a $d^{1/4}$ multiplicative factor of the optimal convergence rate, where d is the underlying dimension.

7.26. Relating Leverage Scores and Density using Regularized Christoffel Functions

Statistical leverage scores emerged as a fundamental tool for matrix sketching and column sampling with applications to low rank approximation, regression, random feature learning and quadrature. Yet, the very nature of this quantity is barely understood. Borrowing ideas from the orthogonal polynomial literature, we introduce in [31] the regularized Christoffel function associated to a positive definite kernel. This uncovers a variational formulation for leverage scores for kernel methods and allows to elucidate their relationships with the chosen kernel as well as population density. Our main result quantitatively describes a decreasing relation between leverage score and population density for a broad class of kernels on Euclidean spaces. Numerical simulations support our findings.

7.27. Averaging Stochastic Gradient Descent on Riemannian Manifolds

In [37] we consider the minimization of a function defined on a Riemannian manifold M accessible only through unbiased estimates of its gradients. We develop a geometric framework to transform a sequence of slowly converging iterates generated from stochastic gradient descent (SGD) on M to an averaged iterate sequence with a robust and fast $O(1/n)$ convergence rate. We then present an application of our framework to geodesically-strongly-convex (and possibly Euclidean non-convex) problems. Finally, we demonstrate how these ideas apply to the case of streaming k-PCA, where we show how to accelerate the slow rate of the randomized power method (without requiring knowledge of the eigengap) into a robust algorithm achieving the optimal rate of convergence.

7.28. Localized Structured Prediction

Key to structured prediction is exploiting the problem structure to simplify the learning process. A major challenge arises when data exhibit a local structure (e.g., are made by "parts") that can be leveraged to better approximate the relation between (parts of) the input and (parts of) the output. Recent literature on signal processing, and in particular computer vision, has shown that capturing these aspects is indeed essential to achieve state-of-the-art performance. While such algorithms are typically derived on a case-by-case basis, in [42] we propose the first theoretical framework to deal with part-based data from a general perspective. We derive a novel approach to deal with these problems and study its generalization properties within the setting of statistical learning theory. Our analysis is novel in that it explicitly quantifies the benefits of leveraging the part-based structure of the problem with respect to the learning rates of the proposed estimator.

7.29. Optimal rates for spectral algorithms with least-squares regression over Hilbert spaces

In [12], we study regression problems over a separable Hilbert space with the square loss, covering non-parametric regression over a reproducing kernel Hilbert space. We investigate a class of spectral-regularized algorithms, including ridge regression, principal component analysis, and gradient methods. We prove optimal,

high-probability convergence results in terms of variants of norms for the studied algorithms, considering a capacity assumption on the hypothesis space and a general source condition on the target function. Consequently, we obtain almost sure convergence results with optimal rates. Our results improve and generalize previous results, filling a theoretical gap for the non-attainable cases.

7.30. Differential Properties of Sinkhorn Approximation for Learning with Wasserstein Distance

Applications of optimal transport have recently gained remarkable attention thanks to the computational advantages of entropic regularization. However, in most situations the Sinkhorn approximation of the Wasserstein distance is replaced by a regularized version that is less accurate but easy to differentiate. In [17] we characterize the differential properties of the original Sinkhorn distance, proving that it enjoys the same smoothness as its regularized version and we explicitly provide an efficient algorithm to compute its gradient. We show that this result benefits both theory and applications: on one hand, high order smoothness confers statistical guarantees to learning with Wasserstein approximations. On the other hand, the gradient formula allows us to efficiently solve learning and optimization problems in practice. Promising preliminary experiments complement our analysis.

7.31. Learning with SGD and Random Features

Sketching and stochastic gradient methods are arguably the most common techniques to derive efficient large scale learning algorithms. In [15], we investigate their application in the context of nonparametric statistical learning. More precisely, we study the estimator defined by stochastic gradient with mini batches and random features. The latter can be seen as form of nonlinear sketching and used to define approximate kernel methods. The considered estimator is not explicitly penalized/constrained and regularization is implicit. Indeed, our study highlights how different parameters, such as number of features, iterations, step-size and mini-batch size control the learning properties of the solutions. We do this by deriving optimal finite sample bounds, under standard assumptions. The obtained results are corroborated and illustrated by numerical experiments.

7.32. Manifold Structured Prediction

Structured prediction provides a general framework to deal with supervised problems where the outputs have semantically rich structure. While classical approaches consider finite, albeit potentially huge, output spaces, in [19] we discuss how structured prediction can be extended to a continuous scenario. Specifically, we study a structured prediction approach to manifold valued regression. We characterize a class of problems for which the considered approach is statistically consistent and study how geometric optimization can be used to compute the corresponding estimator. Promising experimental results on both simulated and real data complete our study.

7.33. On Fast Leverage Score Sampling and Optimal Learning

Leverage score sampling provides an appealing way to perform approximate computations for large matrices. Indeed, it allows to derive faithful approximations with a complexity adapted to the problem at hand. Yet, performing leverage scores sampling is a challenge in its own right requiring further approximations. In [18], we study the problem of leverage score sampling for positive definite matrices defined by a kernel. Our contribution is twofold. First we provide a novel algorithm for leverage score sampling and second, we exploit the proposed method in statistical learning by deriving a novel solver for kernel ridge regression. Our main technical contribution is showing that the proposed algorithms are currently the most efficient and accurate for these problems.

7.34. Accelerated Decentralized Optimization with Local Updates for Smooth and Strongly Convex Objectives

In [47], we study the problem of minimizing a sum of smooth and strongly convex functions split over the nodes of a network in a decentralized fashion. We propose a decentralized accelerated algorithm that only requires local synchrony. Its rate depends on the condition number κ of the local functions as well as the network topology and delays. Under mild assumptions on the topology of the graph, our algorithm takes a time $O((\tau_{\max} + \Delta_{\max})\sqrt{\kappa/\gamma}\ln(\epsilon^{-1}))$ to reach a precision ϵ where γ is the spectral gap of the graph, $\tau_{\max}$ the maximum communication delay and $\Delta_{\max}$ the maximum computation time. Therefore, it matches the rate of SSDA, which is optimal when $\tau_{\max} = \Omega(\Delta_{\max})$. Applying our algorithm to quadratic local functions leads to an accelerated randomized gossip algorithm of rate $O(\sqrt{\theta_{\text{gossip}}/n})$ where θ_{gossip} is the rate of the standard randomized gossip. To the best of our knowledge, it is the first asynchronous gossip algorithm with a provably improved rate of convergence of the second moment of the error. We illustrate these results with experiments in idealized settings.

ANGE Project-Team

7. New Results

7.1. Numerical methods for fluid flows

Participants: Jacques Sainte-Marie, Virgile Dubos, Cindy Guichard, Martin Parisot, Marie-Odile Bristeau, Fabien Souillé, Edwige Godlewski, Yohan Penel.

7.1.1. Advancing dynamical cores of oceanic models across all scales

Oceanic numerical models are used to understand and predict a wide range of processes from global paleoclimate scales to short-term prediction in estuaries and shallow coastal areas. One of the overarching challenges, and the main topic of the COMMODORE workshop, is the appropriate design of the dynamical cores given the wide variety of scales of interest and their interactions with atmosphere, sea-ice, biogeochemistry, and even societal processes. The construction of a dynamical core is a very long effort which takes years and decades of research and development and which requires a collaborative mixture of scientific disciplines. In [14], we present a significant number of fundamental choices, such as which equations to solve, which horizontal and vertical grid arrangement is adequate, which discrete algorithms allows jointly computational efficiency and sufficient accuracy, etc.

7.1.2. A Well-balanced Finite Volume Scheme for Shallow Water Equations with Porosity

Our work [20] aims to study the ability of a single porosity-based shallow water model to modelize the impact of vegetation in open-channel flows. More attention on flux and source terms discretizations are required in order to archive the well-balancing and shock capturing. We present a new Godunov-type finite volume scheme based on a simple-wave approximation and compare it with some other methods in the literature. A first application with experimental data was performed.

7.1.3. The gradient discretisation method

The monograph [21] is dedicated to the presentation of the gradient discretisation method (GDM) and to some of its applications. It is intended for masters students, researchers and experts in the field of the numerical analysis of partial differential equations.The GDM is a framework which contains classical and recent discretisation schemes for diffusion problems of different kinds: linear or non-linear, steady-state or time-dependent.

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

This work [29] is devoted to the numerical resolution in multidimensional framework of a hierarchy of reduced models of the free surface Euler equations. In a first part, entropy-satisfying scheme is proposed fo the monolayer dispersive model [Green, Naghdi '76] and [Bristeau, Mangeney, Sainte-Marie, Seguin '15]. In a second part, the strategy is extended to the layerwise models proposed in [Fernandez-Nieto, Parisot, Penel, Sainte-Marie]. To illustrate the accuracy and the robustness of the strategy, several numerical experiments are performed. In particular, the strategy is able to deal with dry areas without particular treatment.

7.1.5. Numerical approximation of the 3d hydrostatic Navier-Stokes system with free surface

In this work [23], we propose a stable and robust strategy to approximate the 3d incompressible hydrostatic Euler and Navier-Stokes systems with free surface. Compared to shallow water approximation of the Navier-Stokes system, the idea is to use a Galerkin type approximation of the velocity field with piecewise constant basis functions in order to obtain an accurate description of the vertical profile of the horizontal velocity.

7.1.6. Congested shallow water model: floating object

In [27], we are interested in the floating body problem on a large space scale. We focus on objects floating freely in the water such as icebergs or wave energy converters. The formulation of the fluid-solid interaction using the congested shallow water model for the fluid and Newton's second law of motion for the solid is given and a strong coupling between the two systems is explained. The energy transfer between the solid and the water is focused on since it is of major interest for energy production. A numerical resolution based on the coupling of a finite volume scheme for the fluid and a Newmark scheme for the solid is presented. An entropy correction based on an adapted choice of discretization for the coupling terms is made in order to ensure a dissipation law at the discrete level. Simulations are presented to validate the method and to show the feasibility of more complex cases.

7.1.7. Numerical strategies for a dispersive layer-averaged model

A hierarchy of models has been derived in [12] to approximate the Euler equations by means of a layer-averaging procedure. This results in several dispersive models with one velocity field per layer. The structure of the equations induces issues of efficiency. The standard splitting between hydrostatic and non-hydrostatic components leads to a prohibitive computational costs. In a work in progress, we are investigating a new strategy to solve the projection step in a cheaper way. This is assessed by means of steady nontrivial solutions of the dispersive equations.

7.1.8. Methods of Reflections

The basic idea of the method of reflections appeared almost two hundred years ago; it is a method of successive approximations for the interaction of particles within a fluid, and it seems intuitively related to the Schwarz domain decomposition methods, the subdomains being the complements of the particle domains. We show in [25] that indeed there is a direct correspondence between the methods of reflections and Schwarz methods in the two particle/subdomain case. This allows us to give a new convergence analysis based on maximum principle techniques with precise convergence estimates that one could not obtain otherwise. We then show however also that in the case of more than two particles/subdomains, the methods of reflections and the Schwarz methods are really different methods, with different convergence properties. We finally also introduce for the first time coarse corrections for the methods of reflections to make them scalable in the case when the number of particles becomes large.

7.2. Modelling

Participants: Marie-Odile Bristeau, Jacques Sainte-Marie, Fabien Souillé, Emmanuel Audusse, Léa Boitin, Martin Parisot, Di Martino Bernard, Anne Mangeney.

7.2.1. How do microalgae perceive light in a high-rate pond? Towards more realistic Lagrangian experiments

In [10], we present a multidisciplinary downscaling study, where we first reconstructed single cell trajectories in an open raceway using an original hydrodynamical model offering a powerful discretization of the Navier–Stokes equations tailored to systems with free surfaces. The trajectory of a particular cell was selected and the associated high-frequency light pattern was computed. This light pattern was then experimentally reproduced in an Arduino-driven computer controlled cultivation system with a low density Dunaliella salina culture. The effect on growth and pigment content was recorded for various frequencies of the light pattern, by setting different paddle wheel velocities.

7.2.2. Modeling and simulation of sediment transport

A previous derivation of the sediment layer model has then been extended. Depending on the scaling chosen for the physical parameters, different models are obtained. The model we are interested in is the non-local model (with a viscosity term). Several numerical schemes are implemented and studied to simulate this model. Only one of these schemes is satisfactory. Simulations of the coupled water-sediment systems are made. The influence of the viscosity is emphasized. Turning on the non-local term allows to simulate dune growth and propagation.

Following the previous work, a numerical scheme for the sediment layer is proposed. The numerical scheme is tested. The influence of the viscosity on the behaviour of the sediment layer is studied. A numerical strategy for the resolution of the coupled model (water layer and sediment layer) is implemented. The behaviour of the coupled system is numerically assessed. Academic test cases are performed.

7.2.3. The Navier-Stokes system with temperature and salinity for free-surface flows

We model free surface flows where density variations coming e.g. from temperature or salinity differences play a significant role. Starting from the compressible Navier-Stokes system, a model is derived by performing the incompressible limit (the dependence of the density on the pressure is removed). A layer-averaged formulation of the model is proposed. The layer-averaged model satisfies a dissipative energy balance. A numerical scheme is proposed. It verifies several stability properties (positivity, well-balancing, maximum principle on the density). Numerical simulations are performed. The differences with models relying on the classical Boussinesq approximation are shown.

7.2.4. Various analytical solutions for the incompressible Euler and Navier-Stokes systems with free surface

In this paper [24], we propose several time dependent analytical solutions for the incompressible Euler and Navier-Stokes systems with free surface. The given analytical solutions concerns the hydrostatic and nonhydrostatic Euler and Navier-Stokes systems.

7.3. Functional analysis of PDE models in Fluid Mechanics

Participants: Bilal Al Taki, Boris Haspot.

7.3.1. New functional inequality and its application

In [22], we prove by simple arguments a new kind of Logarithmic Sobolev inequalities generalizing two known inequalities founded in some papers related to fluid dynamics models. As a by product, we show how our inequality can help in obtaining some important a priori estimates for the solution of the Navier-Stokes-Korteweg system.

7.3.2. Vortex solutions for the compressible Navier-Stokes equations with general viscosity coefficients in 1D: regularizing effects or not on the density

We consider Navier-Stokes equations for compressible viscous fluids in the one-dimensional case with general viscosity coefficients. We prove the existence of global weak solution when the initial momentum $\rho_0 u_0$ belongs to the set of the finite measure $\mathcal{M}(\mathbf{R})$ and when the initial density ρ_0 is in the set of bounded variation functions $BV(\mathbf{R})$. In particular it allows to deal with initial momentum which are Dirac masses and initial density which admit shocks. We can observe in particular that this type of initial data have infinite energy. Furthermore we show that if the coupling between the density and the velocity is sufficiently strong then the initial density which admits initially shocks is instantaneously regularized and becomes continuous. This coupling is expressed via the regularity of the so called effective velocity $v = u + \frac{\mu(\rho)}{\rho^2}\psi_x\rho$ with $\mu(\rho)$ the viscosity coefficient. Inversely if the coupling between the initial density and the initial velocity is too weak (typically $\rho_0 v_0 \in \mathcal{M}(\mathbf{R})$) then we prove the existence of weak energy in finite time but the density remains a priori discontinuous on the time interval of existence.

7.3.3. Strong solution for Korteweg system

In this paper we investigate the question of the local existence of strong solution for the Korteweg system in critical spaces when $N \geq 1$ provided that the initial data are small. More precisely the initial momentum $\rho_0 u_0$ belongs to $\mathrm{bmo}_T^{-1}(\mathbf{R}^N)$ for $T > 0$ and the initial density ρ_0 is in $L^\infty(\mathbf{R}^N)$ and far away from the vacuum. This result extends the so called Koch-Tataru theorem for the incompressible Navier-Stokes equations to the case of the Korteweg system. It is also interesting to observe that any initial shock on the density is instantaneously regularized inasmuch as the density becomes Lipschitz for any $\rho(t,\cdot)$ with $t > 0$. We also prove the existence of global strong solution for small initial data $(\rho_0 - 1, \rho_0 u_0)$ in the homogeneous Besov

spaces $(\dot{B}^{\mathbf{N}-1}_{2,\infty}(\mathbf{R}^{\mathbf{N}} \cap \dot{B}^{\mathbf{N}}_{2,\infty}(\mathbf{R}^{\mathbf{N}} \cap L^{\infty}(\mathbf{R}^{\mathbf{N}})) \times (\dot{B}^{\mathbf{N}-1}_{2,\infty}(\mathbf{R}^{\mathbf{N}}))^{\mathbf{N}}$. This result allows in particular to extend in dimension $N = 2$ the notion of Oseen solutions defined for incompressible Navier-Stokes equations to the case of the Korteweg system when the vorticity of the momentum $\rho_0 u_0$ is a Dirac mass $\alpha\delta_0$ with α sufficiently small. However unlike the Navier Stokes equations

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

Participants: Vivien Mallet, Ngoc Bao Tran Le, Antoine Lesieur, Frédéric Allaire, Hammond Janelle.

7.4.1. Uncertainty quantification of on-road traffic emissions

Road traffic emissions of air pollutants depend on both traffic flow and vehicle emission factors. At metropolitan scale, traffic flow can be obtained by traffic assignment models, and emission factors can be computed from the traffic flow using COPERT IV formulas. Global sensitivity analyses, especially the computation of Sobol' indices, was carried out for the traffic model and the air pollutant emissions. In the process, the traffic model was replaced by a metamodel, or surrogate model, in order to reduce the high computational burden. The results identified the most important input parameters, e.g., the demand associated with small travel distances (for the traffic flow) or the gasoline car share (for the air pollutant emissions). Furthermore, the uncertainties in traffic flow and pollutant emissions was quantified by propagating into the model the uncertainties in the input parameters. Large ensembles of traffic flows were generated and evaluated with traffic flow measurements.

7.4.2. Uncertainty quantification in atmospheric dispersion of radionuclides

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

7.4.3. Metamodeling corrected by observational data

An air quality model at urban scale computes the air pollutant concentrations at street resolution based on various emissions, meteorology, imported pollution and city geometry. Because of the computational cost of such model, we previously designed a metamodel using dimension reduction and statistical emulation. Novel work was dedicated to the correction of this metamodel using observational data. The proposed approach builds a corrected metamodel that is still much faster than the original model, but also performs better when compared to new observations.

7.4.4. Sensitivity analysis and metamodeling of an urban noise model

Urban noise mapping models simulate the propagation of noise, originating from emission sources (e.g., road traffic), in all street of a city, based on its geometry. They are subject to uncertainties due to incomplete and erroneous data. We carried out screening studies in order to evaluate the sensitivity of the computed noise to the uncertain data. Further work dealt with the development of a metamodel, which will open the way to uncertainty quantification. The work was carried out with the model NoiseModelling and applied to the noise mapping of Lorient (France).

7.4.5. Monte Carlo simulation and ensemble evaluation for wildland fire propagation

We worked on Monte Carlo simulations of wildland fires. The objective was to evaluate how the uncertainties lying in all the inputs of a fire propagation model can be propagated through the model. A careful review of the literature allowed us to define varying intervals for all the uncertain inputs. The Monte Carlo simulations were then evaluated with ensemble scores, using the observations of the final contours for a number of real cases. The ensemble scores were inspired by classical scores used in meteorology, but were adapted to the nature of the fire observations.

7.4.6. Metamodeling of a complete air quality simulation chain

With the objective of uncertainty quantification, we worked in [15] on the generation of a metamodel for the simulation of urban air quality, using a complete simulation chain including dynamic traffic assignment, the computation of air pollutant emissions and the dispersion of the pollutant in a city. The traffic model and the dispersion model are computationally costly and operate in high dimension. We employed dimension reduction, and coupled it with Kriging in order to build a metamodel for the complete simulation chain.

7.5. Software Devlopments

Participant: Cédric Doucet.

Improvements in the FRESHKISS3D code Several improvements have been achieved in FreshKiss3D :

1. installation step is simpler due to the usage of a YAML file listing third-party libraries;
2. Mac OS is now supported;
3. continuous integration is performed on Ubuntu 16 and OSX with different compilers (GCC, Clang) and different builds (debug, release);
4. a major bug in the computation of fluxes has been fixed;
5. the number of third-party libraries has been minimized (geomalgo, metis4py);
6. build automation is now based on CMake (instead of Waf);
7. documentation has been updated and it is now published during the continuous integration process by means of Gitlab pages;
8. continuous integration has been optimized (better slaves, parallelization)

ARAMIS Project-Team

7. New Results

7.1. Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data

Participants: Jorge Samper-González, Ninon Burgos, Simona Bottani, Sabrina Fontanella, Pascal Lu, Arnaud Marcoux, Alexandre Routier, Jérémy Guillon, Michael Bacci, Junhao Wen, Anne Bertrand, Hugo Bertin, Marie-Odile Habert, Stanley Durrleman, Theodoros Evgeniou, Olivier Colliot [Correspondant].

A large number of papers have introduced novel machine learning and feature extraction methods for automatic classification of Alzheimer's disease (AD). However, while the vast majority of these works use the public dataset ADNI for evaluation, they are difficult to reproduce because different key components of the validation are often not readily available. These components include selected participants and input data, image preprocessing and cross-validation procedures. The performance of the different approaches is also difficult to compare objectively. In particular, it is often difficult to assess which part of the method (e.g. preprocessing, feature extraction or classification algorithms) provides a real improvement, if any. We proposed a framework for reproducible and objective classification experiments in AD using three publicly available datasets (ADNI, AIBL and OASIS). The framework comprises: i) automatic conversion of the three datasets into a standard format (BIDS); ii) a modular set of preprocessing pipelines, feature extraction and classification methods, together with an evaluation framework, that provide a baseline for benchmarking the different components. We demonstrated the use of the framework for a large-scale evaluation on 1960 participants using T1 MRI and FDG PET data. In this evaluation, we assessed the influence of different modalities, preprocessing, feature types (regional or voxel-based features), classifiers, training set sizes and datasets. Performances were in line with the state-of-the-art. FDG PET outperformed T1 MRI for all classification tasks. No difference in performance was found for the use of different atlases, image smoothing, partial volume correction of FDG PET images, or feature type. Linear SVM and L2-logistic regression resulted in similar performance and both outperformed random forests. The classification performance increased along with the number of subjects used for training. Classifiers trained on ADNI generalized well to AIBL and OASIS. All the code of the framework and the experiments is publicly available: general-purpose tools have been integrated into the Clinica software (http://www.clinica.run/) and the paper-specific code is available at: https://gitlab.icm-institute.org/aramislab/AD-ML.

More details in [30].

7.2. An automated pipeline for the analysis of PET data on the cortical surface

Participants: Arnaud Marcoux, Ninon Burgos, Anne Bertrand, Marc Teichmann, Alexandre Routier, Junhao Wen, Jorge Samper-González, Simona Bottani, Stanley Durrleman, Marie-Odile Habert, Olivier Colliot [Correspondant].

We developed a fully automatic pipeline for the analysis of PET data on the cortical surface. Our pipeline combines tools from FreeSurfer and PETPVC, and consists of i) co-registration of PET and T1-w MRI (T1) images, ii) intensity normalization, iii) partial volume correction, iv) robust projection of the PET signal onto the subject's cortical surface, v) spatial normalization to a template, and vi) atlas statistics. We evaluated the performance of the proposed workflow by performing group comparisons and showed that the approach was able to identify the areas of hypometabolism characteristic of different dementia syndromes: Alzheimer's disease (AD) and both the semantic and logopenic variants of primary progressive aphasia. We also showed that these results were comparable to those obtained with a standard volume-based approach. We then performed individual classifications and showed that vertices can be used as features to differentiate cognitively normal and AD subjects. This pipeline is integrated into Clinica, an open-source software platform for neuroscience studies available at http://www.clinica.run/.

More details in [24].

7.3. Comparative study of algorithms for synthetic CT generation from MRI: Consequences for MRI-guided radiation planning in the pelvic region

Participants: Hossein Arabi, Jason A. Dowling, Ninon Burgos [Correspondant], Xiao Han, Peter B. Greer, Nikolaos Koutsouvelis, Habib Zaidi.

Magnetic resonance imaging (MRI)-guided radiation therapy (RT) treatment planning is limited by the fact that the electron density distribution required for dose calculation is not readily provided by MR imaging. We compare a selection of novel synthetic CT generation algorithms recently reported in the literature, including segmentation-based, atlas-based and machine learning techniques, using the same cohort of patients and quantitative evaluation metrics. Six MRI-guided synthetic CT generation algorithms were evaluated: one segmentation technique into a single tissue class (water-only), four atlas-based techniques, namely, median value of atlas images (ALMedian), atlas-based local weighted voting (ALWV), bone enhanced atlas-based local weighted voting (ALWV-Bone), iterative atlas-based local weighted voting (ALWV-Iter), and a machine learning technique using deep convolution neural network (DCNN). Organ auto-contouring from MR images was evaluated for bladder, rectum, bones, and body boundary. Overall, DCNN exhibited higher segmentation accuracy resulting in Dice indices while ALMedian showed the lowest accuracy. DCNN reached the best performance in terms of accurate derivation of synthetic CT values within each organ, followed by the advanced atlas-based methods. ALMedian led to the highest error. Considering the dosimetric evaluation results, ALWV-Iter, ALWV, DCNN and ALWV-Bone led to similar mean dose estimation within each organ at risk and target volume with less than 1% dose discrepancy. However, the two-dimensional gamma analysis demonstrated higher pass rates for ALWV-Bone, DCNN, ALMedian and ALWV-Iter at 1%/1 mm criterion. Overall, machine learning and advanced atlas-based methods exhibited promising performance by achieving reliable organ segmentation and synthetic CT generation. DCNN appears to have slightly better performance by achieving accurate automated organ segmentation and relatively small dosimetric errors (followed closely by advanced atlas-based methods, which in some cases achieved similar performance). However, the DCNN approach showed higher vulnerability to anatomical variation, where a greater number of outliers was observed with this method. Considering the dosimetric results obtained from the evaluated methods, the challenge of electron density estimation from MR images can be resolved with a clinically tolerable error.

More details in [4].

7.4. Double diffeomorphism: combining morphometry and structural connectivity analysis

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

The brain is composed of several neural circuits which may be seen as anatomical complexes composed of grey matter structures interconnected by white matter tracts. Grey and white matter components may be modelled as 3D surfaces and curves respectively. Neurodevelopmental disorders involve morphological and organizational alterations which can not be jointly captured by usual shape analysis techniques based on single diffeomorphisms. We propose a new deformation scheme, called double diffeomorphism, which is a combination of two diffeomorphisms. The first one captures changes in structural connectivity, whereas the second one recovers the global morphological variations of both grey and white matter structures. This deformation model is integrated into a Bayesian framework for atlas construction. We evaluate it on a data-set of 3D structures representing the neural circuits of patients with Gilles de la Tourette syndrome (GTS). We show that this approach makes it possible to localise, quantify and easily visualise the pathological anomalies altering the morphology and organization of the neural circuits. Furthermore, results also indicate that the proposed deformation model better discriminates between controls and GTS patients than a single diffeomorphism.

More details in [15].

7.5. Learning distributions of shape trajectories from longitudinal datasets: a hierarchical model on a manifold of diffeomorphisms

Participants: Alexandre Bône, Olivier Colliot, Stanley Durrleman [Correspondant].

We propose a method to learn a distribution of shape tra-jectories from longitudinal data, i.e. the collection of individual objects repeatedly observed at multiple time-points. The method allows to compute an average spatiotemporal trajectory of shape changes at the group level, and the individual variations of this trajectory both in terms of geometry and time dynamics. First, we formulate a non-linear mixed-effects statistical model as the combination of a generic statistical model for manifold-valued longitudinal data, a deformation model defining shape trajectories via the action of a finite-dimensional set of diffeomorphisms with a man-ifold structure, and an efficient numerical scheme to compute parallel transport on this manifold. Second, we introduce a MCMC-SAEM algorithm with a specific approach to shape sampling, an adaptive scheme for proposal variances , and a log-likelihood tempering strategy to estimate our model. Third, we validate our algorithm on 2D simulated data, and then estimate a scenario of alteration of the shape of the hippocampus 3D brain structure during the course of Alzheimer's disease. The method shows for instance that hippocampal atrophy progresses more quickly in female subjects, and occurs earlier in APOE4 mutation carriers. We finally illustrate the potential of our method for classifying pathological trajectories versus normal ageing.

More details in [38].

7.6. Spatiotemporal Propagation of the Cortical Atrophy: Population and Individual Patterns

Participants: Igor Koval, Jean-Baptiste Schiratti, Alexandre Routier, Michael Bacci, Olivier Colliot, Stéphanie Allassonnière, Stanley Durrleman.

Repeated failures in clinical trials for Alzheimer's disease (AD) have raised a strong interest for the prodromal phase of the disease. A better understanding of the brain alterations during this early phase is crucial to diagnose patients sooner, to estimate an accurate disease stage, and to give a reliable prognosis. According to recent evidence, structural alterations in the brain are likely to be sensitive markers of the disease progression. Neuronal loss translates in specific spatiotemporal patterns of cortical atrophy, starting in the enthorinal cortex and spreading over other cortical regions according to specific propagation pathways. We developed a digital model of the cortical atrophy in the left hemisphere from prodromal to diseased phases, which is built on the temporal alignment and combination of several short-term observation data to reconstruct the long-term history of the disease. The model not only provides a description of the spatiotemporal patterns of cortical atrophy at the group level but also shows the variability of these patterns at the individual level in terms of difference in propagation pathways, speed of propagation, and age at propagation onset. Longitudinal MRI datasets of patients with mild cognitive impairments who converted to AD are used to reconstruct the cortical atrophy propagation across all disease stages. Each observation is considered as a signal spatially distributed on a network, such as the cortical mesh, each cortex location being associated to a node. We consider how the temporal profile of the signal varies across the network nodes. We introduce a statistical mixed-effect model to describe the evolution of the cortex alterations. To ensure a spatiotemporal smooth propagation of the alterations, we introduce a constrain on the propagation signal in the model such that neighboring nodes have similar profiles of the signal changes. Our generative model enables the reconstruction of personalized patterns of the neurodegenerative spread, providing a way to estimate disease progression stages and predict the age at which the disease will be diagnosed. The model shows that, for instance, APOE carriers have a significantly higher pace of cortical atrophy but not earlier atrophy onset.

More details in [19].

7.7. A Fanning Scheme for the Parallel Transport Along Geodesics on Riemannian Manifolds

Participants: Maxime Louis, Benjamin Charlier, Paul Jusselin, Susovan Pal, Stanley Durrleman.

Parallel transport on Riemannian manifolds allows one to connect tangent spaces at different points in an isometric way and is therefore of importance in many contexts, such as for statistics on manifolds. The existing methods to compute parallel transport require either the computation of Riemannian logarithms, such as the Schild's ladder, or the Christoffel symbols. The Logarithm is rarely given in closed form, and therefore costly to compute whereas the number of Christoffel symbols explodes with the dimension of the manifold, making both these methods intractable. From an identity between parallel transport and Jacobi fields, we propose a numerical scheme to approximate the parallel transport along a geodesic. We find and prove an optimal convergence rate for the scheme, which is equivalent to Schild's ladder's. We investigate potential variations of the scheme and give experimental results on the Euclidean two-sphere and on the manifold of symmetric positive-definite matrices.

More details in [23].

7.8. Reduction of recruitment costs in preclinical AD trials. Validation of automatic pre-screening algorithm for brain amyloidosis.

Participants: Manon Ansart, Stéphane Epelbaum, Geoffroy Gagliardi, Olivier Colliot, Didier Dormont, Bruno Dubois, Harald Hampel, Stanley Durrleman [Correspondant].

We propose a method for recruiting asymptomatic Amyloid positive individuals in clinical trials, using a two-step process. We first select during a pre-screening phase a subset of individuals which are more likely to be amyloid positive based on the automatic analysis of data acquired during routine clinical practice, before doing a confirmatory PET-scan to these selected individuals only. This method leads to an increased number of recruitments and to a reduced number of PET-scans, resulting in a decrease in overall recruitment costs. We validate our method on 3 different cohorts, and consider 5 different classification algorithms for the pre-screening phase. We show that the best results are obtained using solely cognitive, genetic and socio-demographic features, as the slight increased performance when using MRI or longitudinal data is balanced by the cost increase they induce. We show that the proposed method generalizes well when tested on an independent cohort, and that the characteristics of the selected set of individuals are identical to the characteristics of a population selected in a standard way. The proposed approach shows how Machine Learning can be used effectively in practice to optimize recruitment costs in clinical trials.

More details in [3].

7.9. Multiplex core-periphery organization of the human connectome

Participants: Federico Battiston, Jeremy Guillon, Mario Chavez, Vito Latora, Fabrizio de Vico Fallani [Correspondant].

What is the core of the human brain is a fundamental question that has been mainly addressed by studying the anatomical connections between differently specialized areas, thus neglecting the possible contributions from their functional interactions. While many methods are available to identify the core of a network when connections between nodes are all of the same type, a principled approach to define the core when multiple types of connectivity are allowed is still lacking. Here, we introduce a general framework to define and extract the core–periphery structure of multi-layer networks by explicitly taking into account the connectivity patterns at each layer. We first validate our algorithm on synthetic networks of different size and density, and with tunable overlap between the cores at different layers. We then use our method to merge information from structural and functional brain networks, obtaining in this way an integrated description of the core of the human connectome. Results confirm the role of the main known cortical and subcortical hubs, but also suggest the presence of new areas in the sensori-motor cortex that are crucial for intrinsic brain functioning. Taken together these findings provide fresh evidence on a fundamental question in modern neuroscience and offer new opportunities to explore the mesoscale properties of multimodal brain networks.

More details in [6].

7.10. Integrating EEG and MEG signals to improve motor imagery classification in brain-computer interfaces

Participants: Marie-Constance Corsi, Mario Chavez, Denis Schwartz, Laurent Hugueville, Ankit Khambhati, Danielle Bassett, Fabrizio de Vico Fallani [Correspondant].

We adopted a fusion approach that combines features from simultaneously recorded electroencephalogram (EEG) and magnetoencephalogram (MEG) signals to improve classification performances in motor imagery-based brain–computer interfaces (BCIs). We applied our approach to a group of 15 healthy subjects and found a significant classification performance enhancement as compared to standard single-modality approaches in the alpha and beta bands. Taken together, our findings demonstrate the advantage of considering multimodal approaches as complementary tools for improving the impact of noninvasive BCIs.

More details in [10].

7.11. Role of inter-hemispheric connections in functional brain networks

Participants: Johann Martinez [Correspondant], Javier Buldu, David Papo, Fabrizio de Vico Fallani, Mario Chavez.

Today the human brain can be modeled as a graph where nodes represent different regions and links stand for statistical interactions between their activities as recorded by different neuroimaging techniques. Empirical studies have lead to the hypothesis that brain functions rely on the coordination of a scattered mosaic of functionally specialized brain regions (modules or sub-networks), forming a web-like structure of coordinated assemblies (a network of networks). The study of brain dynamics would therefore benefit from an inspection of how functional sub-networks interact between them. In this paper, we model the brain as an interconnected system composed of two specific sub-networks, the left (L) and right (R) hemispheres, which compete with each other for centrality, a topological measure of importance in a networked system. Specifically, we considered functional brain networks derived from high-density electroencephalographic (EEG) recordings and investigated how node centrality is shaped by interhemispheric connections. Our results show that the distribution of centrality strongly depends on the number of functional connections between hemispheres and the way these connections are distributed. Additionally, we investigated the consequences of node failure on hemispherical centrality, and showed how the abundance of inter-hemispheric links favors the functional balance of centrality distribution between the hemispheres.

More details in [25].

7.12. Statistical shape analysis of large datasets based on diffeomorphic iterative centroids

Participants: Claire Cury, Joan Glaunès, Olivier Colliot.

We proposed an approach for template-based shape analysis of large datasets, using diffeomorphic centroids as atlas shapes. Diffeomorphic centroid methods fit in the Large Deformation Diffeomorphic Metric Mapping (LDDMM) framework and use kernel metrics on currents to quantify surface dissimilarities. The statistical analysis is based on a Kernel Principal Component Analysis (Kernel PCA) performed on the set of initial momentum vectors which parametrize the deformations. We tested the approach on different datasets of hippocampal shapes extracted from brain magnetic resonance imaging (MRI), compared three different centroid methods and a variational template estimation. The largest dataset is composed of 1,000 surfaces, and we are able to analyse this dataset in 26 h using a diffeomorphic centroid. Our experiments demonstrate that computing diffeomorphic centroids in place of standard variational templates leads to similar shape analysis results and saves around 70% of computation time. Furthermore, the approach is able to adequately capture the variability of hippocampal shapes with a reasonable number of dimensions, and to predict anatomical features of the hippocampus, only present in 17% of the population, in healthy subjects.

More details in [12].

7.13. Multi-modal brain fingerprinting: a manifold approximation based framework

Participants: Kuldeep Kumar, Olivier Colliot, Christian Desrosiers.

We proposed an efficient framework, based on manifold approximation, for generating brain fingerprints from multi-modal data. The proposed framework represents images as bags of local features, which are used to build a subject proximity graph. Compact fingerprints are obtained by projecting this graph in a low-dimensional manifold, using spectral embedding. Experiments using the T1/T2-weighted MRI, diffusion MRI, and resting state fMRI data of 945 Human Connectome Project subjects demonstrate the benefit of combining multiple modalities, with multi-modal fingerprints more discriminative than those generated from individual modalities. Results also highlight the link between fingerprint similarity and genetic proximity, monozygotic twins having more similar fingerprints than dizygotic or non-twin siblings. This link is also reflected in the differences of feature correspondences between twin/sibling pairs, occurring in major brain structures and across hemispheres. The robustness of the proposed framework to factors like image alignment and scan resolution, as well as the reproducibility of results on retest scans, suggest the potential of multi-modal brain fingerprinting for characterizing individuals in a large cohort analysis. In addition, taking inspiration from the computer vision community, the proposed rank retrieval evaluation based on the task of twin/sibling identification and using Mean Average Precision (MAP) can be used for a standardized comparison of future brain fingerprints.

More details in [20].

7.14. Structural, Microstructural, and Metabolic Alterations in Primary Progressive Aphasia Variants

Participants: Alexandre Routier [Correspondant], Marie-Odile Habert, Olivier Colliot, Marc Teichmann.

Neuroimaging studies have described the brain alterations in primary progressive aphasia (PPA) variants (semantic, logopenic, nonfluent/agrammatic). However, few studies combined T1, FDG-PET, and diffusion MRI techniques to study atrophy, hypometabolism, and tract alterations across the three PPA main variants. We therefore explored a large early-stage cohort of semantic, logopenic and nonfluent/agrammatic variants (N = 86) and of 23 matched healthy controls with anatomical MRI (cortical thickness), FDG PET (metabolism) and diffusion MRI (white matter tracts analyses), aiming at identifying cortical and sub-cortical brain alterations, and confronting these alterations across imaging modalities and aphasia variants. In the semantic variant, there was cortical thinning and hypometabolism in anterior temporal cortices, with left-hemisphere predominance, extending toward posterior temporal regions, and affecting tracts projecting to the anterior temporal lobes (inferior longitudinal fasciculus, uncinate fasciculus) and tracts projecting to or running nearby posterior temporal cortices: (superior longitudinal fasciculus, inferior frontal-occipital fasciculus). In the logopenic variant metabolic alterations were more extensive than atrophy affecting mainly the left temporal-parietal junction and extending toward more anterior temporal cortices. Metabolic and tract data were coherent given the alterations of the left superior and inferior longitudinal fasciculus and the left inferior frontal-occipital fasciculus. In the nonfluent/agrammatic variant cortical thinning and hypometabolism were located in the left frontal cortex but Broca's area was only affected on metabolic measures. Metabolic and tract alterations were coherent as reflected by damage to the left uncinate fasciculus connecting with Broca's area. Our findings provide a full-blown statistically robust picture of brain alterations in early-stage variants of primary progressive aphasia which has implications for diagnosis, classification and future therapeutic strategies. They demonstrate that in logopenic and semantic variants patterns of brain damage display a non-negligible overlap in temporal regions whereas they are substantially distinct in the nonfluent/agrammatic variant (frontal regions). These results also indicate that frontal networks (combinatorial syntax/phonology) and temporal networks (lexical/semantic representations) constitute distinct anatomo-functional entities with differential vulnerability to degenerative processes in aphasia variants. Finally, the identification of the specific damage

patterns could open an avenue for trans-cranial stimulation approaches by indicating the appropriate target-entry into the damaged language system.

More details in [29].

7.15. Neurite density is reduced in the presymptomatic phase of C9orf72 disease

Participants: Junhao Wen, Hui Zhang, Daniel Alexander, Stanley Durrleman, Olivier Colliot, Isabelle Le Ber, Anne Bertrand [Correspondant].

In this study, we aimed to assess the added value of neurite orientation dispersion and density imaging (NODDI) compared to conventional DTI and anatomical MRI to detect changes in presymptomatic carriers of chromosome 9 open reading frame 72 (C9orf72) mutation. The PREV-DEMALS study is a prospective, multicenter, observational study of first-degree relatives of individuals carrying the C9orf72 mutation. Sixty-seven participants (38 presymptomatic C9orf72 mutation carriers [C9+], 29 non carriers [C9-]) were included in the present cross-sectional study. Each participant underwent one single-shell, multi-shell diffusion MRI and 3DT1 MRI. Volumetric measures, DTI and NODDI metrics were calculated within regions of interest. Differences in white matter integrity, gray matter volume and free water fraction between C9+ and C9- individuals were assessed using linear mixed-effects models. Compared with C9-, C9+ demonstrated white matter abnormalities in 10 tracts with neurite density index, and only 5 tracts with DTI metrics. Effect size was significantly higher for the neurite density index than for DTI metrics in two tracts. No tract had a significantly higher effect size for DTI than for NODDI. For gray matter cortical analysis, free water fraction was increased in 13 regions in C9+, whereas 11 regions displayed volumetric atrophy. In conclusion, NODDI provides higher sensitivity and greater tissue-specificity compared to conventional DTI for identifying white matter abnormalities in the presymptomatic C9orf72 carriers. Our results encourage the use of neurite density as biomarker of the preclinical phase.

More details in [34].

7.16. Learning myelin content in multiple sclerosis from multimodal MRI through adversarial training

Participants: Wen Wei, Emilie Poirion, Benedetta Bodini, Stanley Durrleman, Nicholas Ayache, Bruno Stankoff, Olivier Colliot [Correspondant].

Multiple sclerosis (MS) is a demyelinating disease of the central nervous system (CNS). A reliable measure of the tissue myelin content is therefore essential to understand the physiopathology of MS, track progression and assess treatment efficacy. Positron emission tomography (PET) with [11 C]PIB has been proposed as a promising biomarker for measuring myelin content changes in-vivo in MS. However, PET imaging is expensive and invasive due to the injection of a radioactive tracer. On the contrary, magnetic resonance imaging (MRI) is a non-invasive, widely available technique, but existing MRI sequences do not provide, to date, a reliable, specific, or direct marker of either demyelination or remyelination. In this work, we therefore propose Sketcher-Refiner Generative Adversarial Networks (GANs) with specifically designed adversarial loss functions to predict the PET-derived myelin content map from a combination of MRI modalities. The prediction problem is solved by a sketch-refinement process in which the sketcher generates the preliminary anatomical and physiological information and the refiner refines and generates images reflecting the tissue myelin content in the human brain. We evaluated the ability of our method to predict myelin content at both global and voxel-wise levels. The evaluation results show that the demyelination in lesion regions and myelin content in normal-appearing white matter (NAWM) can be well predicted by our method. The method has the potential to become a useful tool for clinical management of patients with MS.

More details in [40].

7.17. COGEVIS: A New Scale to Evaluate Cognition in Patients with Visual Deficiency

Participants: Claire Meyniel, Dalila Samri, Farah Stefano, Joel Crevoisier, Florence Bonté, Raffaela Migliaccio, Laure Delaby, Anne Bertrand, Marie-Odile Habert, Bruno Dubois, Baram Bodaghi, Stéphane Epelbaum [Correspondant].

We evaluated the cognitive status of visually impaired patients referred to low vision rehabilitation (LVR) based on a standard cognitive battery and a new evaluation tool, named the COGEVIS, which can be used to assess patients with severe visual deficits. We studied patients aged 60 and above, referred to the LVR Hospital in Paris. Neurological and cognitive evaluations were performed in an expert memory center. Thirty-eight individuals, 17 women and 21 men with a mean age of 70.3(SD=1.3 years) and a mean visual acuity of 0.12(SD=0.02), were recruited over a one-year period. Sixty-three percent of participants had normal cognitive status. Cognitive impairment was diagnosed in 37.5% of participants. The COGEVIS score cutoff point to screen for cognitive impairment was 24 (maximum score of 30) with a sensitivity of 66.7% and a specificity of 95%. Evaluation following 4 months of visual rehabilitation showed an improvement of Instrumental Activities of Daily Living ($p = 0.004$), National Eye Institute Visual Functioning Questionnaire ($p = 0.035$), and Montgomery-Åsberg Depression Rating Scale ($p = 0.037$). This study introduces a new short test to screen for cognitive impairment in visually impaired patients.

More details in [27].

7.18. Neural correlates of episodic memory in the Memento cohort

Participants: Stéphane Epelbaum [Correspondant], Vincent Bouteloup, Jean François Mangin, Valentina La Corte, Raffaela Migliaccio, Hugo Bertin, Marie Odile Habert, Clara Fischer, Chabha Azouani, Ludovic Fillon, Marie Chupin, Bruno Vellas, Florence Pasquier, Frederic Blanc, Audrey Gabelle, Mathieu Ceccaldi, Pierre Krolak-Salmon, Jacques Hugon, Olivier Hanon, Olivier Rouaud, Renaud David, Genevieve Chene, Bruno Dubois, Carole Dufouil.

The free and cued selective reminding test is used to identify memory deficits in mild cognitive impairment and demented patients. It allows assessing three processes: encoding, storage, and recollection of verbal episodic memory. We investigated the neural correlates of these three memory processes in a large cohort study. The Memento cohort enrolled 2323 outpatients presenting either with subjective cognitive decline or mild cognitive impairment who underwent cognitive, structural MRI and, for a subset, fluorodeoxyglucose-positron emission tomography evaluations. Encoding was associated with a network including parietal and temporal cortices; storage was mainly associated with entorhinal and parahippocampal regions, bilaterally; retrieval was associated with a widespread network encompassing frontal regions. The neural correlates of episodic memory processes can be assessed in large and standardized cohorts of patients at risk for Alzheimer's disease. Their relation to pathophysiological markers of Alzheimer's disease remains to be studied.

7.19. Cognitive and neuroimaging features and brain amyloidosis in individuals at risk of Alzheimer's disease

Participants: Bruno Dubois [Correspondant], Stéphane Epelbaum, Francis Nyasse, Hovagim Bakardjian, Geoffroy Gagliardi, Olga Uspenskaya, Marion Houot, Simone Lista, Federica Cacciamani, Marie Claude Potier, Anne Bertrand, Foudil Lamari, Habib Benali, Jean François Mangin, Olivier Colliot, Remy Genthon, Marie-Odile Habert, Harald Hampel.

Improved understanding is needed of risk factors and markers of disease progression in preclinical Alzheimer's disease. We assessed associations between brain amyloidosis and various cognitive and neuroimaging parameters with progression of cognitive decline in individuals with preclinical Alzheimer's disease. The INSIGHT-preAD is an ongoing single-centre observational study at the Salpêtrière Hospital, Paris, France. Eligible participants were age 70-85 years with subjective memory complaints but unimpaired cognition and memory (Mini-Mental State Examination [MMSE] score $\geq$ 27, Clinical Dementia Rating score 0, and Free and

Cued Selective Reminding Test [FCSRT] total recall score ≥ 41). We stratified participants by brain amyloid deposition on 18F-florbetapir PET (positive or negative) at baseline. All patients underwent baseline assessments of demographic, cognitive, and psychobehavioural, characteristics, APOE $\epsilon 4$ allele carrier status, brain structure and function on MRI, brain glucose-metabolism on 18F-fluorodeoxyglucose (18F-FDG) PET, and event-related potentials on electroencephalograms (EEGs). Actigraphy and CSF investigations were optional. Participants were followed up with clinical, cognitive, and psychobehavioural assessments every 6 months, neuropsychological assessments, EEG, and actigraphy every 12 months, and MRI, and 18F-FDG and 18F-florbetapir PET every 24 months. We assessed associations of amyloid deposition status with test outcomes at baseline and 24 months, and with clinical status at 30 months. Progression to prodromal Alzheimer's disease was defined as an amnestic syndrome of the hippocampal type. From May 25, 2013, to Jan 20, 2015, we enrolled 318 participants with a mean age of 76·0 years (SD 3·5). The mean baseline MMSE score was 28·67 (SD 0·96), and the mean level of education was high (score >6 [SD 2] on a scale of 1-8, where 1=infant school and 8=higher education). 88 (28showed amyloid deposition and the remainder did not. The amyloid subgroups did not differ for any psychobehavioural, cognitive, actigraphy, and structural and functional neuroimaging results after adjustment for age, sex, and level of education More participants positive for amyloid deposition had the APOE $\epsilon 4$ allele (33 [38%] vs 29 [13%], p<0·0001). Amyloid concentration in CSF significantly correlated with mean 18F-florbetapir uptake at baseline (r=-0·62, p<0·0001) and the ratio of amyloid to amyloid (r=-0·61, p<0·0001), and identified amyloid deposition status with high accuracy (mean area under the curve values 0·89, 95% CI 0·80-0·98 and 0·84, 0·72-0·96, respectively). No difference was seen in MMSE (28·3 [SD 2·0] vs 28·9 [1·2], p=0·16) and Clinical Dementia Rating scores (0·06 [0·2] vs 0·05 [0·3]; p=0·79) at 30 months (n=274) between participants positive or negative for amyloid. Four participants (all positive for amyloid deposition at baseline) progressed to prodromal Alzheimer's disease. They were older than other participants positive for amyloid deposition at baseline (mean 80·2 years [SD 4·1] vs 76·8 years [SD 3·4]) and had greater 18F-florbetapir uptake at baseline (mean standard uptake value ratio 1·46 [SD 0·16] vs 1·02 [SD 0·20]), and more were carriers of the APOE $\epsilon 4$ allele (three [75%] of four vs 33 [39%] of 83). They also had mild executive dysfunction at baseline (mean FCSRT free recall score 21·25 [SD 2·75] vs 29·08 [5·44] and Frontal Assessment Battery total score 13·25 [1·50] vs 16·05 [1·68]). Brain amyloidosis alone did not predict progression to prodromal Alzheimer's disease within 30 months. Longer follow-up is needed to establish whether this finding remains consistent.

More details in [13].

MAMBA Project-Team

7. New Results

7.1. Modelling Polymerization Processes

Nucleation Phenomena.
A new stochastic model of polymerization including the nucleation has been analyzed in [4]. A Functional Central Limit Theorem for the Becker-Döring model in an infinite dimensional state space is established in [25].

An oscillatory model of polymerisation-depolymerisation.
In 2017, we evidenced the presence of several polymeric species by using data assimilation methods to fit experimental data from H. Rezaei's lab [64]. In collaboration with Klemens Fellner from the university of Graz, we now propose a new model, variant of the Becker-Döring system but containing two monomeric species, capable of displaying sustained though damped oscillations [39].

Time asymptotics for nucleation, growth and division equations.
We revisited the well-known Lifshitz-Slyozov model, which takes into account only polymerisation and depolymerisation, and progressively enriched the model. Taking into account depolymerisation and fragmentation reaction term may surprisingly stabilisde the system, since a steady size-distribution of polymers may then emerge, so that "Ostwald ripening" does not happen [8].

Cell population dynamics and its control
The PhD thesis work of Camille Pouchol (co-supervisors Jean Clairambault, Michèle Sabbah, INSERM, and Emmanuel Trélat, Inria CAGE and LJLL) has been continued, leading after his first article published in the J. Maths Pures Appl. [136], summarised in [31], to his PhD defence in June [1], and to a diversification of his research activities in various directions related to population dynamics and optimal control with Antoine Olivier, Emmanuel Trélat and Enrique Zuazua [51], [56] or to more general questions [55].

Measure solutions for the growth-fragmentation equation
As recalled in the section "Foundations", entropy methods for population dynamics have been successfully developed around B. Perthame and co-authors. We recently extend such methods to the growth-fragmentation equation, in collaboration with P. Gwiazda, E. Wiedemann and T. Debiec [40], using the framework of generalised Young measures.

7.2. Large Stochastic Networks

The equilibrium properties of allocation algorithms for networks with a large number of nodes with finite capacity are investigated in [46] and in [60].

7.3. Control Strategies for Sterile Insect Techniques

We proposed different models to serve as a basis for the design of control strategies relying on releases of sterile male mosquitoes (*Aedes spp*) and aiming at elimination of wild vector population. Different types of releases were considered (constant, periodic or impulsive) and sufficient conditions to reach elimination were provided in each case [57], [3], [35]. We also estimated sufficient and minimal treatment times. A feedback approach was introduced, in which the impulse amplitude is chosen as a function of the actual wild population [57], [3], [35].

7.4. Optimal replacement strategies, application to Wolbachia

We modelled and designed optimal release control strategy with the help of a least square problem. In a nutshell, one wants to minimize the number of uninfected mosquitoes at a given time horizon, under relevant biological constraints. We derived properties of optimal controls and studied a limit problem providing useful asymptotic properties of optimal controls [49], [3].

7.5. Oscillatory regimes in population models

Understanding mosquitoes life cycle is of great interest presently because of the increasing impact of vector borne diseases. Observations yields evidence of oscillations in these populations independent of seasonality, still unexplained. We proposed [58], [3] a simple mathematical model of egg hatching enhancement by larvae which produces such oscillations that conveys a possible explanation.

On the other hand, population oscillations may be induced by seasonal changes. We considered a biological population whose environment varies periodically in time, exhibiting two very different "seasons", favorable and unfavorable. We addressed the following question: the system's period being fixed, under what conditions does there exist a critical duration above which the population cannot sustain and extincts, and below which the system converges to a unique periodic and positive solution? We obtained [59], [3] sufficient conditions for such a property to occur for monotone differential models with concave nonlinearities, and applied the obtained criterion to a two-dimensional model featuring juvenile and adult insect populations.

7.6. Feedback control principles for population replacement by Wolbachia

The issue of effective scheduling of the releases of *Wolbachia*-infected mosquitoes is an interesting problem for Control theory. Having in mind the important uncertainties present in the dynamics of the two populations in interaction, we attempted to identify general ideas for building release strategies, which should apply to several models and situations [34]. These principles were exemplified by two interval observer-based feedback control laws whose stabilizing properties were demonstrated when applied to a model retrieved from [76].

7.7. Bacterial motion by run and tumble

Collective motion of chemotactic bacteria such as Escherichia coli relies, at the individual level, on a continuous reorientation by runs and tumbles. It has been established that the length of run is decided by a stiff response to a temporal sensing of chemical cues along the pathway. We describe in [21] a novel mechanism for pattern formation stemming from the stiffness of chemotactic response relying on a kinetic chemotaxis model which includes a recently discovered formalism for the bacterial chemotaxis. We prove instability both for a microscopic description in the space-velocity space and for the macroscopic equation, a flux-limited Keller-Segel equation, which has attracted much attention recently. A remarkable property is that the unstable frequencies remain bounded, as it is the case in Turing instability. Numerical illustrations based on a powerful Monte Carlo method show that the stationary homogeneous state of population density is destabilized and periodic patterns are generated in realistic ranges of parameters. These theoretical developments are in accordance with several biological observations.

This motivates also our study of traveling wave and aggregation in population dynamics of chemotactic cells based on the FLKS model with a population growth term [7]. Our study includes both numerical and theoretical contributions. In the numerical part, we uncover a variety of solution types in the one-dimensional FLKS model additionally to standard Fisher/KPP type traveling wave. The remarkable result is a counter-intuitive backward traveling wave, where the population density initially saturated in a stable state transits toward an un- stable state in the local population dynamics. Unexpectedly, we also find that the backward traveling wave solution transits to a localized spiky solution as increasing the stiffness of chemotactic response. In the theoretical part, we obtain a novel analytic formula for the minimum traveling speed which includes the counter-balancing effect of chemotactic drift vs. reproduction/diffusion in the propagating front. The front propagation speeds of numerical results only slightly deviate from the minimum traveling speeds, except for the localized spiky solutions, even for the backward traveling waves. We also discover an analytic solution of unimodal traveling wave in the large-stiffness limit, which is certainly unstable but exists in a certain range of parameters.

7.8. Numerical methods for cell aggregation by chemotaxis

Three-dimensional cultures of cells are gaining popularity as an in vitro improvement over 2D Petri dishes. In many such experiments, cells have been found to organize in aggregates. We present new results of three-dimensional in vitro cultures of breast cancer cells exhibiting patterns. Understanding their formation is of particular interest in the context of cancer since metastases have been shown to be created by cells moving in clusters. In the paper [37], we propose that the main mechanism which leads to the emergence of patterns is chemotaxis, i.e., oriented movement of cells towards high concentration zones of a signal emitted by the cells themselves. Studying a Keller-Segel PDE system to model chemotactical auto-organization of cells, we prove that it is subject to Turing instability if a time-dependent condition holds. This result is illustrated by two-dimensional simulations of the model showing spheroidal patterns. They are qualitatively compared to the biological results and their variability is discussed both theoretically and numerically.

This motivates to study parabolic-elliptic Keller-Segel equation with sensitivity saturation, because of its pattern formation ability, is a challenge for numerical simulations. We provide in [16] two finite-volume schemes that are shown to preserve, at the discrete level, the fundamental properties of the solutions, namely energy dissipation, steady states, positivity and conservation of total mass. These requirements happen to be critical when it comes to distinguishing between discrete steady states, Turing unstable transient states, numerical artifacts or approximate steady states as obtained by a simple upwind approach. These schemes are obtained either by following closely the gradient flow structure or by a proper exponential rewriting inspired by the Scharfetter-Gummel discretization. An interesting fact is that upwind is also necessary for all the expected properties to be preserved at the semi-discrete level. These schemes are extended to the fully discrete level and this leads us to tune precisely the terms according to explicit or implicit discretizations. Using some appropriate monotonicity properties (reminiscent of the maximum principle), we prove well-posedness for the scheme as well as all the other requirements. Numerical implementations and simulations illustrate the respective advantages of the three methods we compare.

7.9. Focus on cancer

Modelling Acute Myeloid Leukaemia (AML) and its control by anticancer drugs by PDEs and Delay Differential equations

This theme has continued to be developed in collaboration with Catherine Bonnet, Inria DISCO (Saclay) [12], [29]. Without control by drugs, but with representation of mutualistic interactions between tumor cells and their surrounding support stroll cells, it has also, in collaboration with Delphine Salort and Thierry Jaffredo (LCQB-IBPS) given rise to a recent work by Thanh Nam Nguyen, hired as HTE and ERC postdoctoral fellow at LCQB, submitted as full article [50].

Adaptive dynamics setting to model and circumvent evolution towards drug resistance in cancer by optimal control

The research topic “Evolution and cancer”, designed in the framework of adaptive dynamics to represent and overcome acquired drug resistance in cancer, initiated in [119], [118] and later continued in [90], [89], [117], has been recently summarised in [31] and has been the object of the PhD thesis work of Camille Pouchol, see above “Cell population dynamics and its control” . It is now oriented, thanks to work underway by Cécile Carrère, Jean Clairambault, Tommaso Lorenzi and Grégoire Nadin, in particular towards the mathematical representation of *bet hedging* in cancer, namely a supposed optimal strategy consisting for cancer cell populations under life-threatening cell stress in diversifying their phenotypes according to several resistance mechanisms, such as overexpression of ABC transporters (P-glycoprotein and many others), of DNA repair enzymes or of intracellular detoxication processes. According to different deadly insults the cancer cell population is exposed to, some phenotypes may be selected, any such successful subpopulation being able to store the cell population genome (or subclones of it if the cell population is already genetically heterogeneous) and make it amenable to survival and renewed replication.

Philosophy of cancer biology

This new research topic in Mamba, dedicated to explore possibly underinvestigated, from the mathematical modelling point of view, parts of the field of cancer growth, evolution and therapy, has been the object of a presentation by Jean Clairambault at the recent workshop "Philosophy of cancer biology' (https://www.philinbiomed.org/event/philosophy-of-cancer-biology-workshop/). This workshop gathered most members worldwide of this small, but very active in publishing, community of philosophers of science whose field of research is "philosophy of cancer", as they call it themselves. This topic offers a clear point of convergence between mathematics, biology and social and human sciences.

7.10. Deformable Cell Modeling: biomechanics and Liver regeneration

- Biomechanically mediated growth control of cancer cells The key intriguing novelty was that the same agent-based model after a single parameter has been calibrated with growth data for multicellular spheroids without application of external mechanical stress by adapting a single parameter, permitted to correctly predict the growth speed of multicellular spheroids of 5 different cell lines subject of external mechanical stress. Hereby the same mechanical growth control stress function was used without any modification [44]. The prediction turned out to be correct independent of the experimental method used to exert the stress, whereby once a mechanical capsule has been used, once dextran has been used in the experiments.
- Regeneration of liver with the Deformable Cell Model. The key novelty was the implementation of the model itself, but an interesting novel result is that the DCM permits closure of a pericentral liver lobule lesion generated by drug-induced damage with about 5 times smaller active migration force due to the ability of the cell to strongly deform and squeeze into narrow spaces between the capillaries. This finding stresses that a precise mechanical description is important in view of quantitatively correct modeling results [142]. The deformable cell model however could be used to calibrate the interaction forces of the computationally much cheaper center-based model to arrive at almost the same results.

REO Project-Team

7. New Results

7.1. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Muriel Boulakia, Ludovic Boilevin-Kayl, Chen-Yu Chiang, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Céline Grandmont, Damiano Lombardi, Marc Thiriet, Marina Vidrascu.

In [31], we consider a system modeling the interaction between a viscous incompressible fluid and an elastic structure. The fluid motion is represented by the classical Navier-Stokes equations while the elastic displacement is described by the linearized elasticity equation. The elastic structure is immersed in the fluid and the whole system is confined into a bounded domain of dimension 3. Our main result is the local in time existence and uniqueness of a strong solution of the corresponding system. This result holds without any restrictive assumptions on the domains geometry.

The numerical simulation of a thin-walled structure immersed in an incompressible fluid can be addressed by various methods. In [16], three of them are considered: the Arbitrary Lagrangian-Eulerian (ALE) method, the Fictitious Domain/Lagrange multipliers (FD) method and the Nitsche-XFEM method. Taking ALE as a reference, the advantages and limitations of FD and Nitsche-XFEM are carefully discussed on three benchmark test cases which have been chosen to be representative of typical difficulties encountered in valves or living cells simulations.

Fictitious domain approximations of fluid-structure interaction problems are generally discretized in time using strongly coupled schemes. This guarantees unconditional stability but at the price of solving a computationally demanding coupled system at each time-step. The design of loosely coupled schemes (i.e., methods that invoke the fluid and solid solvers only once per time-step) is of fundamental interest, especially for three-dimensional simulations, but the existing approaches are known to suffer from severe stability and/or time accuracy issues. In [28], we propose a new approach that overcomes these difficulties in the case of the coupling with immersed thin-walled structures.

In [27], we derive a Nitsche-based formulation for fluid-structure interaction (FSI) problems with contact. The approach is based on the work of Chouly and Hild [SIAM Journal on Numerical Analysis. 2013;51(2):1295-1307] for contact problems in solid mechanics. We present two numerical approaches, both of them formulating the FSI interface and the contact conditions simultaneously in equation form on a joint interface-contact surface. The first approach uses a relaxation of the contact conditions to allow for a small mesh-dependent gap between solid and wall. The second alternative introduces an artificial fluid below the contact surface. The resulting systems of equations can be included in a consistent fashion within a monolithic variational formulation, which prevents the so-called "chattering" phenomenon. To deal with the topology changes in the fluid domain at the time of impact, we use a fully Eulerian approach for the FSI problem. We compare the effect of slip and no-slip interface conditions and study the performance of the method by means of numerical examples.

7.2. Numerical methods for biological flows

Participants: Ludovic Boilevin-Kayl, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Florian Joly, Alexandre This, Marc Thiriet, Irene Vignon Clementel.

Cirrhosis is the common end-stage of chronic liver disease, with architectural distortion increasing the intrahepatic vascular resistance, leading to portal hypertension and systemic circulatory disorders. In [13] we investigate the impact of the changing vascular resistances on the hepatic and global circulation hemodynamics during cirrhogenesis. Morphological quantification of vascular trees from corosion casts of rats developing the disease provide the input for a lumped parameter model of the liver that was coupled to a model of the entire circulation of the rat. The simulations explain how vascular changes due to cirrhosis severely disrupt both hepatic and global hemodynamics.

Image-based models derived from CT angiography are being used clinically to simulate blood flow in the coronary arteries of individual patients to aid in the diagnosis of disease and planning treatments. However, image resolution limits vessel segmentation to larger epicardial arteries. In [20], we propose an algorithm for the generation of a patient-specific cardiac vascular network from epicardial vessels down to arterioles. We extend a tree generation method based on satisfaction of functional principles, to account for competing vascular trees, with flow-related and geometrical constraints adapting the simultaneous tree growths to patient priors.

Growth and remodeling of the embryo pharyngeal arch artery (PAA) network into the extracardiac great vessels is poorly understood but a major source of clinically serious malformations. In [21] we develop a methodological pipeline from high-resolution nano-computed tomography imaging and live-imaging flow measurements to multiscale pulsatile computational models. We identify local morphological variation along the PAAs and their association with specific hemodynamic changes in embryos of different stages, advancing our understanding of morphogenesis.

In [22] we evaluate atrioventricular valve regurgitation (AVVR) in babies born with an already very challenging heart condition, i.e., with single ventricle physiology. Although the second surgery that single ventricle patients undergo is thought to decrease AVVR, there is much controversy in the clinical literature about AVVR treatment. The effect of AVVR on Stage 1 haemodynamics and resulting acute changes from conversion to Stage 2 circulation in single ventricle patients are analyzed through lumped parameter models. Several degrees of AVVR severity are analyzed, for two types of valve regurgitation: incomplete leaflet closure and valve prolapse.

The medical imaging community is eager to define quantitative biophysical parameters. As part of a book addressing this question, in [26], we give a short overview of the mathematical modeling of blood flow at different resolutions, from the large vessel scale (three-dimensional, one-dimensional, and zero-dimensional modeling) to microcirculation and tissue perfusion.

In order to reduce the complexity of heart hemodynamics simulations, uncoupling approaches are often considered for the modeling of the immersed valves as an alternative to complex fluid-structure interaction (FSI) models. A possible shortcoming of these simplified approaches is the difficulty to correctly capture the pressure dynamics during the isovolumetric phases. In [35], we propose an enhanced resistive immersed surfaces (RIS) model of cardiac valves which overcomes this issue. The benefits of the model are investigated and tested in blood flow simulations of the left heart.

7.3. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Jean-Frédéric Gerbeau, Damiano Lombardi, Fabien Raphel.

In [19] a method to assess the variability of phenomena described by PDEs is proposed. In particular, the probability density distribution of the parameters of a model is estimated, in such a way that the statistics of the model output match the observed ones. The investigated approach is based on a differential entropy regularised moment matching.

In [25] we investigated how, by a semi-empirical design of composite biomarkers, the classification of the action of a drug on the electrical activity of a cell can be improved. The data used are measured with a Micro-Electrodes-Array.

In [33] a method is investigated, to design composite biomarkers by exploiting a database of in silico experiments. In particular, a dictionary approach is proposed. The composite biomarker is expressed as a linear combination of linear and non-linear forms applied to the observable. The coefficients of the combination are determined by solving a ℓ^1 regularised optimisation problem.

7.4. Lung and respiration modeling

Participants: Laurent Boudin, Céline Grandmont, Marina Vidrascu, Marc Thiriet, Irene Vignon Clementel.

In [34] we analyse multiscale models arising in the description of physiological flows such as blood flow in arteries or air flow in the bronchial tree. The fluid in the 3D part is described by the Stokes or the Navier-Stokes system which is coupled to 0D models or so-called Windkessel models. The resulting Navier-Stokes-Windkessel coupled system involves Neumann non-local boundary conditions that depends on the considered applications. We first show that the different types of Windkessel models share a similar formalism. Next we derive stability estimates for the continuous coupled Stokes-Windkessel or Navier-Stokes-Windkessel problem as well as stability estimates for the semi-discretized systems with either implicit or explicit coupling.We exhibit different kinds of behavior depending on the considered 0D model. Moreover even if no energy estimates can be derived in energy norms for the Navier-Stokes-Windkessel system, leading to possible numerical instabilities for large applied pressures, we show that stability estimates for both the continuous and semi-discrete problems, can be obtained in appropriate norms for small enough data by introducing a new well chosen Stokes-like operator. These sufficient stability conditions on the data may give a hint on the order of magnitude of the data enabling stable computations without stabilization method for the problem.

In [17], we consider a multi-species kinetic model which leads to the Maxwell-Stefan equations under a standard diffusive scaling (small Knudsen and Mach numbers). We propose a suitable numerical scheme which approximates both the solution of the kinetic model in rarefied regime and the one in the diffusion limit. We prove some a priori estimates (mass conservation and nonnegativity) and well-posedness of the discrete problem. We also present numerical examples where we observe the asymptotic-preserving behavior of the scheme.

In [30], we are interested in a system of fluid equations for mixtures with a stiff relaxation term of Maxwell-Stefan diffusion type. We use the formalism developed by Chen, Levermore, Liu to obtain a limit system of Fick type where the species velocities tend to align to a bulk velocity when the relaxation parameter remains small.

In [29], we consider the Boltzmann operator for mixtures with cutoff Maxwellian, hard potentials, or hard spheres collision kernels. In a perturbative regime around the global Maxwellian equilibrium, the linearized Boltzmann multi-species operator L is known to possess an explicit spectral gap, in the global equilibrium weighted L^2 space. We study a new operator L_ε obtained by linearizing the Boltzmann operator for mixtures around local Maxwellian distributions, where all the species evolve with different small macroscopic velocities of order $\varepsilon > 0$. This is a non-equilibrium state for the mixture. We establish a quasi-stability property for the Dirichlet form of L_ε in the global equilibrium weighted L^2 space. More precisely, we consider the explicit upper bound that has been proved for the entropy production functional associated to L and we show that the same estimate holds for the entropy production functional associated to L_ε, up to a correction of order ε.

7.5. Miscellaneous

Participants: Damiano Lombardi, Irene Vignon Clementel.

In [32] numerical quadrature schemes for the integration of observable quantities in the Brillouin zone for the periodic Schrödinger operator are investigated.

The indocyanine green (ICG) clearance, presented as plasma disappearance rate is, presently, a reliable method to estimate the hepatic function. However, this technique is not instantaneously available and thus cannot been used intra-operatively (during liver surgery). Near-infrared spectroscopy enables to assess hepatic ICG concentration over time in the liver tissue. In [14], we propose to extract more information from the liver intensity dynamics by interpreting it through a dedicated pharmacokinetics model. Parameters for different liver states are estimated from in-vivo measurements in rabbits (El-Desoky et al. 1999), and their link with liver function is investigated.

The hepatic hemodynamics is an essential parameter in surgical planning as well as in various disease processes. The transit time ultrasound (TTUS) perivascular flow probe technology is widely used in clinical practice to evaluate the hepatic inflow, yet invasive. The phase-contrast-MRI (PC-MRI) is not invasive and potentially applicable in assessing the hepatic blood flow. In [15], we compare the hepatic inflow rates using the PC-MRI and the TTUS probe, and evaluated their predictive value of post-hepatectomy adverse events in a porcine experimental model of partial hepatectomy.

SERENA Project-Team

7. New Results

7.1. Unfitted hybrid-high-order methods

Participants: Alexandre Ern, Guillaume Delay.

Our team contributes actively to the development of hybrid high-order (HHO) methods. Such methods support polyhedral meshes with hanging nodes, but one requirement is that the mesh cells have planar faces. This is difficult when it comes to solving with high accuracy a problem posed on a domain with curved boundaries or a problem involving a curved interface separating two materials with different properties. One key idea to treat these problems is to use an unfitted mesh, so that the curved boundary or the curved interface freely cuts through the mesh cells. This greatly simplifies the meshing process, but at the same time poses the question on how the HHO method can address the approximation of functions that are not smooth within some mesh cells. The major idea in our approach, which is inspired from similar approaches developed in the context of the more classical finite element method, is to double the discrete unknowns attached to the cut mesh faces and to introduce a consistent Nitsche-type formulation to enforce either the boundary condition or the jump conditions across the interface in a weak manner. In this context, we started a collaboration with Erik Burman (University College London) and we elaborated in [20] the numerical analysis of HHO methods in an unfitted context; further analysis for Stokes and Helmholtz equations has started recently within the postdoc of Guillaume Delay and a collaboration on the subject with CEA is on the way.

7.2. An exponential time stepping scheme for the simulation of diffusion processes

Participant: Géraldine Pichot.

We present in [56] a new Monte Carlo algorithm to simulate diffusion processes in presence of discontinuous convective and diffusive terms. The algorithm is based on the knowledge of close form analytic expressions of the resolvents of the diffusion processes which are usually easier to obtain than close form analytic expressions of the density. In the particular case of diffusion processes with piecewise constant coefficients, known as Skew Diffusions, such close form expressions for the resolvent are available. Then we apply our algorithm to this particular case and we show that the approximate densities of the particles given by the algorithm replicate well the particularities of the true densities (discontinuities, bimodality, ...) Besides, numerical experiments show a quick convergence.

7.3. Localization of dual and distance norms

Participants: Martin Vohralík, Patrick Ciarlet Jr., Jan Blechta, Josef Málek.

Dual norms like the dual norm of the residual and the distance norm to the Sobolev space H_0^1 seem to be fundamentally global over the entire computational domain. In [23], together with P. Ciarlet, we prove, in extension of some older results, that they are both equivalent to the Hilbertian sums of their localizations over patches of elements. Together with J. Blechta and J. Málek, we extend in [45] this result from the space H_0^1 with Hilbertian structure to the Sobolev space $W_0^{1,p}$, with the exponent p bigger than or equal to one, and to an arbitrary bounded linear functional on $W_0^{1,p}$.

7.4. Adaptivity with guaranteed error contraction

Participants: Martin Vohralík, Alexandre Ern, Patrik Daniel, Iain Smears.

In [26], we conceive novel adaptive refinement strategies which automatically decide between mesh refinement and polynomial degree increase. We numerically observe that the error decreases exponentially as a function of the number of degrees of freedom, for smooth as well as for singular numerical solutions. The salient feature of our approach is, however, that we ensure that the error on the next hp-refinement step will be reduced at least by a factor that is given. We then extend in [53] this result to the case where the underlying algebraic solver is inexact. To the best of our knowledge, these results, obtained in the framework of the Ph.D. thesis of Patrik Daniel, is the first ever where such an error contraction bound is computable and guaranteed. Numerically, its precision turns out to be very high (overestimation by a factor very close to the optimal value of one). It immediately implies convergence of the adaptive method, and we would like to use it in the near future for optimality proofs.

ALPINES Project-Team

7. New Results

7.1. First kind Galerkin boundary element method for the Hodge-Laplacian in three dimensions

Boundary value problems for the Euclidean Hodge-Laplacian in three dimension $-\Delta_{HL} = \mathbf{curlcurl} - \mathbf{grad}\mathrm{div}$ lead to variational formulations set in subspaces of $\mathbf{H}(\mathbf{curl},\Omega) \cap \mathbf{H}(\mathrm{div},\Omega)$, $\Omega \subset \mathbb{R}^3$ a bounded Lipschitz domain. Via a representation formula and Calderón identities we derive corresponding first-kind boundary integral equations set in trace spaces of $H^1(\Omega)$, $\mathbf{H}(\mathbf{curl},\Omega)$, and $\mathbf{H}(\mathrm{div},\Omega)$. They give rise to saddle-point variational formulations and feature kernels whose dimensions are linked to fundamental topological invariants of Ω.

Kernels of the same dimensions also arise for the linear systems generated by low-order conforming Galerkin boundary element (BE) discretization. On their complements, we can prove stability of the discretized problems, nevertheless. We prove that discretization does not affect the dimensions of the kernels and also illustrate this fact by numerical tests.

7.2. Boundary integral multi-trace formulations and Optimised Schwarz Methods

In the present contribution, we consider Helmholtz equation with material coefficients being constant in each subdomain of a geometric partition of the propagation medium (discarding the presence of junctions), and we are interested in the numerical solution of such a problem by means of local multi-trace boundary integral formulations (local-MTF). For a one dimensional problem and configurations with two subdomains, it has been recently established that applying a Jacobi iterative solver to local-MTF is exactly equivalent to an Optimised Schwarz Method (OSM) with a non-local impendance. In the present contribution, we show that this correspondance still holds in the case where the subdomain partition involves an arbitrary number of subdomains. From this, we deduce that the depth of the adjacency graph of the subdomain partition plays a critical role in the convergence of linear solvers applied to local-MTF: we prove it for the case of homogeneous propagation medium and show, through numerical evidences, that this conclusion still holds for heterogeneous media. Our study also shows that, considering variants of local-MTF involving a relaxation parameter, there is a fixed value of this relaxation parameter that systematically leads to optimal speed of convergence for linear solvers.

7.3. Poroelasticity

In [38], we design and study a fully coupled numerical scheme for the poroelasticity problem modelled through Biot's equations. The classical way to numerically solve this system is to use a finite element method for the mechanical equilibrium equation and a finite volume method for the fluid mass conservation equation. However, to capture specific properties of underground media such as heterogeneities, discontinuities and faults, meshing procedures commonly lead to badly shaped cells for finite element based modelling. Consequently, we investigate the use of the recent virtual element method which appears as a potential discretization method for the mechanical part and could therefore allow the use of a unique mesh for the both mechanical and fluid flow modelling. Starting from a first insight into virtual element method applied to the elastic problem in the context of geomechanical simulations, we apply in addition a finite volume method to take care of the fluid conservation equation. We focus on the first order virtual element method and the two point flux approximation for the finite volume part. A mathematical analysis of this original coupled scheme is provided, including existence and uniqueness results and a priori estimates. The method is then illustrated by some computations on two or three dimensional grids inspired by realistic application cases.

7.4. Hybrid discontinuous Galerkin discretisation and domain decomposition preconditioners for the Stokes problem

Solving the Stokes equation by an optimal domain decomposition method derived algebraically involves the use of nonstandard interface conditions whose discretisation is not trivial. For this reason the use of approximation methods such as hybrid discontinuous Galerkin appears as an appropriate strategy: on the one hand they provide the best compromise in terms of the number of degrees of freedom in between standard continuous and discontinuous Galerkin methods, and on the other hand the degrees of freedom used in the nonstandard interface conditions are naturally defined at the boundary between elements. In this paper, we introduce the coupling between a well chosen discretisation method (hybrid discontinuous Galerkin) and a novel and efficient domain decomposition method to solve the Stokes system. We present the detailed analysis of the hybrid discontinuous Galerkin method for the Stokes problem with non standard boundary conditions. This analysis is supported by numerical evidence. In addition, the advantage of the new preconditioners over more classical choices is also supported by numerical experiments. The full paper [18] is available at https://hal.archives-ouvertes.fr/hal-01967577

7.5. A class of efficient locally constructed preconditioners based on coarse spaces

In [14] we present a class of robust and fully algebraic two-level preconditioners for SPD matrices. We introduce the notion of algebraic local SPSD splitting of an SPD matrix and we give a characterization of this splitting. This splitting leads to construct algebraically and locally a class of efficient coarse spaces which bound the spectral condition number of the preconditioned matrix by a number defined a priori. We also introduce the notion of filtering subspace. This concept helps compare the dimension minimality of coarse spaces. Some PDEs-dependant preconditioners correspond to a special case. The examples of the algebraic coarse spaces in this paper are not practical due to expensive construction. We propose a heuristic approximation that is not costly. Numerical experiments illustrate the efficiency of the proposed method.

7.6. Enlarged Krylov methods for reducing communication

Krylov methods are widely used for solving large sparse linear systems of equations. On distributed architectures, their performance is limited by the communication needed at each iteration of the algorithm. In [34], we study the use of so-called enlarged Krylov subspaces for reducing the number of iterations, and therefore the overall communication, of Krylov methods. In particular, we consider a reformulation of the Conjugate Gradient method using these enlarged Krylov subspaces: the enlarged Conjugate Gradient method. We present the parallel design of two variants of the enlarged Conjugate Gradient method as well as their corresponding dynamic versions where the number of search directions is dynamically reduced during the iterations. For a linear elasticity problem with heterogeneous coefficients using a block Jacobi preconditioner, we show that this implementation scales up to 16,384 cores, and is up to 6,9 times faster than the PETSc implementation of PCG.

In [15] we propose a variant of the GMRES method for solving linear systems of equations with one or multiple right-hand sides. Our method is based on the idea of the enlarged Krylov subspace to reduce communication. It can be interpreted as a block GMRES method. Hence, we are interested in detecting inexact breakdowns. We introduce a strategy to perform the test of detection. Furthermore, we propose an eigenvalues deflation technique aiming to have two benefits. The first advantage is to avoid the plateau of convergence after the end of a cycle in the restarted version. The second is to have a very fast convergence when solving the same system with different right-hand sides, each given at a different time (useful in the context of CPR preconditioner). With the same memory cost, we obtain a saving of up to 50% in the number of iterations to reach convergence with respect to the original method.

7.7. Recycling Krylov subspaces and reducing deflation subspaces for solving a sequence of linear systems

In [32] we present deflation strategies related to recycling Krylov subspace methods for solving one or a sequence of linear systems of equations. Besides well-known strategies of deflation, Ritz and harmonic Ritz based deflation, we introduce an SVD-based deflation technique. We consider the recycling in two contexts, recycling the Krylov subspace between the cycles of restarts and recycling a deflation subspace when the matrix changes in a sequence of linear systems. Numerical experiments on real-life reservoir simulations demonstrate the impact of our proposed strategy.

7.8. Solving linear equations with messenger-field and conjugate gradient techniques: an application to CMB data analysis

In [26] we discuss linear system solvers invoking a messenger-field and compare them with (preconditioned) conjugate gradients approaches. We show that the messenger-field techniques correspond to fixed point iterations of an appropriately preconditioned initial system of linear equations. We then argue that a conjugate gradient solver applied to the same preconditioned system, or equivalently a preconditioned conjugate gradient solver using the same preconditioner and applied to the original system, will in general ensure at least a comparable and typically better performance in terms of the number of iterations to convergence and time-to-solution. We illustrate our conclusions on two common examples drawn from the Cosmic Microwave Background data analysis: Wiener filtering and map-making. In addition, and contrary to the standard lore in the CMB field, we show that the performance of the preconditioned conjugate gradient solver can depend importantly on the starting vector. This observation seems of particular importance in the cases of map-making of high signal-to-noise sky maps and therefore should be of relevance for the next generation of CMB experiments.

7.9. Low rank approximation of a sparse matrix based on LU factorization with column and row tournament pivoting

In [23] 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.

7.10. ALORA: affine low-rank approximations

In [17] we introduce the concept of affine low-rank approximation for an $m \times n$ matrix, consisting in fitting its columns into an affine subspace of dimension at most $k << \min(m, n)$. We show that the optimal affine approximation can be obtained by applying an orthogonal projection to the matrix before constructing its best approximation. Moreover, we present the algorithm ALORA that constructs an affine approximation by slightly modifying the application of any low-rank approximation method. We focus on approximations created with the classical QRCP and subspace iteration algorithms. For the former, we present a detailed analysis of the existing pivoting techniques and furthermore, we provide a bound for the error when an arbitrary pivoting technique is used. For the case of subspace iteration, we prove a result on the convergence of singular vectors, showing a bound that is in agreement with the one for convergence of singular values proved recently. Finally, we present numerical experiences using challenging matrices taken from different fields, showing good performance and validating the theoretical framework.

7.11. Linear-time CUR approximation of BEM matrices

In [33] we propose linear-time CUR approximation algorithms for admissible matrices obtained from the hierarchical form of Boundary Element matrices. We propose a new approach called geometric sampling to obtain indices of most significant rows and columns using information from the domains where the problem is posed. Our strategy is tailored to Boundary Element Methods (BEM) since it uses directly and explicitly the cluster tree containing information from the problem geometry. Our CUR algorithm has precision comparable with low-rank approximations created with the truncated QR factorization with column pivoting (QRCP) and the Adaptive Cross Approximation (ACA) with full pivoting, which are quadratic-cost methods. When compared to the well-known linear-time algorithm ACA with partial pivoting, we show that our algorithm improves, in general, the convergence error and overcomes some cases where ACA fails. We provide a general relative error bound for CUR approximations created with geometrical sampling. Finally, we evaluate the performance of our algorithms on traditional BEM problems defined over different geometries.

7.12. Fractional decomposition of matrices and parallel computing

In [40] we are interested in the design of parallel numerical schemes for linear systems. We give an effective solution to this problem in the following case: the matrix A of the linear system is the product of p nonsingular matrices A_i^m with specific shape: $A_i = I - h_i X$ for a fixed matrix X and real numbers h_i. Although having the special form, these matrices A_i arise frequently in the discretization of evolutionary Partial Differential Equations. The idea is to express A^{-1} as a linear combination of elementary matrices A_i^{-k}. Hence the solution of the linear system with matrix A is a linear combination of the solutions of linear systems with matrices A_i^k. These systems are solved simultaneously on different processors.

DELYS Team

5. New Results

5.1. Distributed Algorithms for Dynamic Networks and Fault Tolerance

Participants: Luciana Bezerra Arantes [correspondent], Sébastien Bouchard, Marjorie Bournat, João Paulo de Araujo, Swan Dubois, Laurent Feuilloley, Denis Jeanneau, Jonathan Lejeune, Franck Petit [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.

In 2018, we obtained the following results.

5.1.1. Scheduling in uncertain environments

In [19], we consider scheduling with faults/errors and we introduce a new non-probabilistic model with explorable (query-able) uncertainty. Each unit-time error is characterized by an uncertainty area during which the error will occur, and it is possible to learn the exact slot at which it will appear by issuing a query operation of unit cost. We study two problems: (i) the error-query scheduling problem, whose aim is to reveal enough error-free slots with the minimum number of queries, and (ii) the lexicographic error-query scheduling problem where we seek the earliest error-free slots with the minimum number of queries. We consider both the off-line and the on-line versions of the above problems. In the former, the whole instance and its characteristics are known in advance and we give a polynomial-time algorithm for the error-query scheduling problem. In the latter, the adversary has the power to decide, in an on-line way, the time-slot of appearance for each error. We propose then both lower bounds and algorithms whose competitive ratios asymptotically match these lower bounds.

5.1.2. Failure detectors in dynamic systems

The failure detector abstraction was introduced as a way to circumvent the impossibility of solving consensus in asynchronous systems prone to crash failures. A failure detector is a local oracle that provides processes in the system with unreliable information on process failures. But a failure detector that is sufficient to solve a given problem in a static system is not necessarily sufficient to solve the same problem in a dynamic system. In [37], we adapt an existing failure detector for mutual exclusion and prove that it is the weakest failure detector to solve mutual exclusion in dynamic systems, which means that it is weaker than any other failure detector capable of solving mutual exclusion.

We also propose in [15] a new failure detector, called the Impact failure detector (FD), 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.

5.1.3. Causal information dissemination

A causal broadcast ensures that messages are delivered to all nodes (processes) preserving causal relation of the messages. In [33], we propose a new causal broadcast protocol for distributed systems whose nodes are logically organized in a virtual hypercube-like topology called VCube. Messages are broadcast by dynamically building spanning trees rooted in the message's source node. By using multiple trees, the contention bottleneck problem of a single root spanning tree approach is avoided. Furthermore , different trees can intersect at some node. Hence, by taking advantage of both the out-of-order reception of causally related messages at a node and these paths intersections, a node can delay to one or more of its children in the tree. Experimental evaluation conducted on top of PeerSim simulator confirms the communication effectiveness of our causal broadcast protocol in terms of latency and message traffic reduction

5.1.4. Graceful Degradation

Gracefully degrading algorithms was introduced by Biely *et al.*i. Such algorithms offer the desirable properties to circumvent impossibility results in dynamic systems by adapting themselves to the dynamics. Indeed, such a algorithms solve a given problem under some dynamics and, moreover, guarantees that a weaker (but related) problem is solved under a higher dynamics under which the original problem is impossible to solve. The underlying intuition is to solve the problem whenever possible but to provide some kind of quality of service if the dynamics become (unpredictably) higher.

In [36], we apply for the first time this approach to robot networks. We focus on the fundamental problem of gathering a squad of autonomous robots on an unknown location of a dynamic ring. In this goal, we introduce a set of weaker variants of this problem. Motivated by a set of impossibility results related to the dynamics of the ring, we propose a gracefully degrading gathering algorithm.

5.1.5. Unreliable Hints

In [23], we address the question of a mobile agent deterministically searching for a target in the Euclidean plane. We assume that the mobile agent is equipped with a compass and a measure of length has to find an inert treasure in the Euclidean plane. Both the agent and the treasure are modeled as points. In the beginning, the agent is at a distance at most $D > 0$ from the treasure, but knows neither the distance nor any bound on it. Finding the treasure means getting at distance at most 1 from it. The agent makes a series of moves. Each of them consists in moving straight in a chosen direction at a chosen distance. In the beginning and after each move the agent gets a hint consisting of a positive angle smaller than 2π whose vertex is at the current position of the agent and within which the treasure is contained. We investigate the problem of how these hints permit the agent to lower the cost of finding the treasure, using a deterministic algorithm, where the cost is the worst-case total length of the agent's trajectory. It is well known that without any hint the optimal (worst case) cost is $\Theta(D^2)$. We show that if all angles given as hints are at most π, then the cost can be lowered to $O(D)$, which is optimal. If all angles are at most β, where $\beta < 2\pi$ is a constant unknown to the agent, then the cost is at most $O(D^2 - \epsilon)$, for some $\epsilon > 0$. For both these positive results we present deterministic algorithms achieving the above costs. Finally, if angles given as hints can be arbitrary, smaller than 2π, then we show that cost $\Theta(D^2)$ cannot be beaten.

5.1.6. Gathering of Mobile Agents

Gathering a group of mobile agents is a fundamental task in the field of distributed and mobile systems. It consists of bringing agents that initially start from different positions to meet all together in finite time. In the case when there are only two agents, the gathering problem is often referred to as the rendezvous problem.

In [14] and [22], we consider these tasks from a deterministic point of view in networks modeled as undirected and anonymous graphs. An adversary chooses the initial nodes of the agents (the number of agents may be larger than the number of nodes) and assigns a different positive integer (called label) to each of them. Initially, each agent knows its label as well as some global knowledge shared by all the agents. The agents can communicate with each other only when located at the same node.

This task has been considered in the literature under two alternative scenarios: weak and strong. Under the weak scenario, agents may meet either at a node or inside an edge. Under the strong scenario, they have to meet at a node, and they do not even notice meetings inside an edge. Gathering and rendezvous algorithms under the strong scenario are known for synchronous agents. For asynchronous agents, gathering and rendezvous under the strong scenario are impossible even in the two-node graph, and hence only algorithms under the weak scenario were constructed.

In [14] we show that rendezvous under the strong scenario is possible for agents with asynchrony restricted in the following way: agents have the same measure of time but the adversary can impose, for each agent and each edge, the speed of traversing this edge by this agent. The speeds may be different for different edges and different agents but all traversals of a given edge by a given agent have to be at the same imposed speed. We construct a deterministic rendezvous algorithm for such agents, working in time polynomial in the size of the graph, in the length of the smaller label, and in the largest edge traversal time.

Gathering mobile agents can be made drastically more difficult to achieve when some agents are subject to faults, especially the Byzantine ones that are known as being the worst faults to handle. Byzantine means that the agent is subject to unpredictable and arbitrary faults. For instance, such an agent may choose to never stop or to never move. In [22] we study the task of Byzantine gathering among synchronous agents under the strong scenario: despite the presence of f Byzantine agents, all the other (correct) agents have to meet at the same node. In this respect, assuming that the agents are in a *strong team* i.e., a team in which the number of correct agents is at least some prescribed value that is quadratic in f, we show an algorithm that solves Byzantine gathering with all strong teams in all graphs of size at most n, for any integers n and f, in a time polynomial in n and the length $|l_{min}|$ of the binary representation of the smallest label of a good agent. The algorithm works using a global knowledge of size $\mathcal{O}(\log\log\log n)$, which we prove to be of optimal order of magnitude in our context to reach a time complexity that is polynomial in n and $|l_{min}|$.

5.1.7. Self-Stabilizing Minimum Diameter Spanning Tree

In [13], we present a self-stabilizing algorithm for the minimum diameter spanning tree construction problem in the state model. Our protocol has the following attractive features. It is the first algorithm for this problem that operates under the *unfair and distributed* adversary (or *daemon*). In other words, no restriction is made on the asynchronous behavior of the system. Second, our algorithm needs only $O(\log n)$ bits of memory per process (where n is the number of processes), that improves the previous result by a factor n. These features are not achieved to the detriment of the convergence time, which stays polynomial.

5.2. Large-scale data distribution

Participants: Saalik Hatia, Mesaac Makpangou, Sébastien Monnet, Sreeja Nair, Jonathan Sid-Otmane, Pierre Sens, Marc Shapiro, Alejandro Tomsic, Ilyas Toumlilt, Dimitrios Vasilas, Paolo Viotti.

5.2.1. Impossibility results for distributed transactional reads

We study the costs and trade-offs of providing transactional consistent reads in a distributed storage system. We identify the following dimensions: read consistency, read delay (latency), and data freshness. We show that there is a three-way trade-off between them, which can be summarised as follows: (i) it is not possible to ensure at the same time order-preserving (e.g., causally-consistent) or atomic reads, Minimal Delay, and maximal freshness; thus, reading data that is the most fresh without delay is possible only in a weakly-isolated mode; (ii) to ensure atomic or order-preserving reads at Minimal Delay imposes to read data from the past (not fresh); (iii) however, order-preserving minimal-delay reads can be fresher than atomic; (iv) reading atomic or order-preserving data at maximal freshness may block reads or writes indefinitely. Our impossibility results hold independently of other features of the database, such as update semantics (totally ordered or not) or data model (structured or unstructured). Guided by these results, we modify an existing protocol to ensure minimal-delay reads (at the cost of freshness) under atomic-visibility and causally-consistent semantics. Our experimental evaluation supports the theoretical results.

This work was published at Middleware 2018 [31].

5.2.2. Co-design and verification of an available file system

Distributed file systems play a vital role in large-scale enterprise services. However, the designer of a distributed file system faces a vexing choice between strong consistency and asynchronous replication. The former supports a standard sequential model by synchronising operations, but is slow and fragile. The latter is highly available and responsive, but exposes users to concurrency anomalies. We describe a rigorous and general approach to navigating this trade-off by leveraging static verification tools that allow to verify different file system designs. We show that common file system operations can run concurrently without synchronisation, while still retaining a semantics reasonably similar to Posix hierarchical structure. The one exception is the "move" operation, for which we prove that, unless synchronised, it will have an anomalous behaviour.

This work was published at VMCAI 2018 [28].

5.3. Resources management in system software

Participants: Michael Damien Carver, Jonathan Lejeune, Pierre Sens, Julien Sopena [correspondent], Gauthier Voron, Francis Laniel.

5.3.1. Multicore schedulers

In collaboration with WHISPER team, we have contributed to an analysis of the impact on application performance of the design and implementation choices made in two widely used open-source schedulers: ULE, the default FreeBSD scheduler, and CFS, the default Linux scheduler. In a paper published at USENIX ATC'18 [24], we compare ULE and CFS in otherwise identical circumstances. This work involves porting ULE to Linux, and using it to schedule all threads that are normally scheduled by CFS. We compare the performance of a large suite of applications on the modified kernel running ULE and on the standard Linux kernel running CFS. The observed performance differences are solely the result of scheduling decisions, and do not reflect differences in other subsystems between FreeBSD and Linux. We found that there is no overall winner. On many workloads the two schedulers perform similarly, but for some workloads there are significant and even surprising differences. ULE may cause starvation, even when executing a single application with identical threads, but this starvation may actually lead to better application performance for some workloads. The more complex load balancing mechanism of CFS reacts more quickly to workload changes, but ULE achieves better load balance in the long run.

DYOGENE Project-Team

6. New Results

6.1. Energy Trade-offs for end-to-end Communications in Urban Vehicular Networks exploiting an Hyperfractal Model

In [33] presented this year at MSWIM DIVANet we show results on the trade-offs between the end-to-end communication delay and energy spent for completing a transmission in vehicular communications in urban settings. This study exploits our innovative model called " hyperfractal " that captures the self-similarity of the topology and vehicle locations in cities. We enrich the model by incorporating roadside infrastructure. We use analytical tools to derive theoretical bounds for the end-to-end communication hop count under two different energy constraints: either total accumulated energy, or maximum energy per node. More precisely, we prove that the hop count is bounded by $O(n(1-\alpha)/(dm-1))$ where $\alpha < 1$ and $m > 2$ is the precise hyperfractal dimension. This proves that for both constraints the energy decreases as we allow to chose among paths of larger length. In fact the asymptotic limit of the energy becomes significantly small when the number of nodes becomes asymptotically large. A lower bound on the network throughput capacity with constraints on path energy is also given. The results are confirmed through exhaustive simulations using different hyperfractal dimensions and path loss coefficients.

6.2. Broadcast Speedup in Vehicular Networks via Information Teleportation

In [32] presented this year at LCN our goal is to increase our understanding of the fundamental communication properties in urban vehicle-to-vehicle mobile networks by exploiting the self-similarity and hierarchical organization of modern cities. We use an innovative model called "hyperfractal" that captures the self-similarities of both the traffic and vehicle locations, and yet avoids the extremes of regularity and randomness. We use analytical tools to derive matching theoretical upper and lower bounds for the information propagation speed in an urban delay tolerant network (i.e., a network that is disconnected at all time, and thus uses a store-carry-and-forward routing model). We prove that the average broadcast time behaves as $n(1-\delta)$ (times a slowly varying function), where δ depends on the precise fractal dimension. Furthermore, we show that the broadcast speedup is due in part to an interesting self-similar phenomenon, that we denote as information teleportation. This phenomenon arises as a consequence of the topology of the vehicle traffic, and triggers an acceleration of the broadcast time. We show that our model fits real cities where open traffic data sets are available. The study presents simulations that confirm the validity of the bounds in multiple realistic settings, including scenarios with variable speed.

6.3. Vehicle-to-Infrastructure Communications Design in Urban Hyperfractals

In [25] presented at SPAWC our goal is to increase the awareness about the communication opportunities that arise in urban vehicle networks when exploiting the self-similarity and hierarchical organization of modern cities. The work uses our innovative model called " hyperfractal " that captures the self-similarity of the urban vehicular networks as well as incorporating roadside infrastructure with its own self-similarity. We use analytical tools to provide achievable trade-offs in operating the roadside units under the constraint of minimum routing path delay while maintaining a reasonably balanced load. The models and results are supported by simulations with different city hyperfractal dimensions in two different routing scenarios: nearest neighbor routing with no collision and minimum delay routing model assuming slotted Aloha, signal to interference ratio (SIR) capture condition, power-path loss, Rayleigh fading.

6.4. Book on Stochastic Geometry Analysis of Cellular Networks

In 2018 we have published a monograph [30] in which we explain the very latest analytic techniques and results from stochastic geometry for modelling the signal-to-interference-plus-noise ratio (SINR) distribution in heterogeneous cellular networks. This book is supposed to help readers to understand the effects of combining different system deployment parameters on key performance indicators such as coverage and capacity, enabling the efficient allocation of simulation resources. In addition to covering results for network models based on the Poisson point process, this book presents recent results for when non-Poisson base station configurations appear Poisson, due to random propagation effects such as fading and shadowing, as well as non-Poisson models for base station configurations, with a focus on determinantal point processes and tractable approximation methods. Theoretical results are illustrated with practical Long-Term Evolution (LTE) applications and compared with real-world deployment results.

6.5. Gibbsian On-Line Distributed Content Caching Strategy for Cellular Networks

In [9], 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.

6.6. Location Aware Opportunistic Bandwidth Sharing between Static and Mobile Users with Stochastic Learning in Cellular Networks

In [7], 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.

6.7. Performance analysis of cellular networks with opportunistic scheduling using queueing theory and stochastic geometry

In [38] submitted this year, combining stochastic geometric approach with some classical results from queuing theory,we propose a comprehensive framework for the performance study of large cellular networks featuring opportunistic scheduling. Rapid and verifiable with respect to real data, our approach is particularly useful for network dimensioning and long term economic planning. It is based on a detailed network model combining

an information-theoretic representation of the link layer, a queuing-theoretic representation of the users' scheduler, and a stochastic-geometric representation of the signal propagation and the network cells. It allows one to evaluate principal characteristics of the individual cells, such as loads (defined as the fraction of time the cell is not empty), the mean number of served users in the steady state, and the user throughput. A simplified Gaussian approximate model is also proposed to facilitate study of the spatial distribution of these metrics across the network. The analysis of both models requires only simulations of the point process of base stations and the shadowing field to estimate the expectations of some stochastic-geometric functionals not admitting explicit expressions. A key observation of our approach , bridging spatial and temporal analysis, relates the SINR distribution of the typical user to the load of the typical cell of the network. The former is a static characteristic of the network related to its spectral efficiency while the latter characterizes the performance of the (generalized) processor sharing queue serving the dynamic population of users of this cell.

6.8. The Influence of Canyon Shadowing on Device-to-Device Connectivity in Urban Scenario

In [48] submitted this year, we use percolation theory to study the feasibility of large-scale connectivity of relay-augmented device-to-device (D2D) networks in an urban scenario, featuring a haphazard system of streets and canyon shadowing allowing only for line-of-sight (LOS) communications in a limited finite range. We use a homogeneous Poisson-Voronoi tessellation (PVT) model of streets with homogeneous Poisson users (devices) on its edges and independent Bernoulli relays on the vertices. Using this model, we demonstrated the existence of a minimal threshold for relays below which large-scale connectivity of the network is not possible, regardless of all other network parameters. Through simulations, we estimated this threshold to 71.3%. Moreover, if the mean street length is not larger than some threshold (predicted to 74.3% of the communication range; which might be the case in a typical urban scenario) then any (whatever small) density of users can be compensated by equipping more crossroads with relays. Above this latter threshold, good connectivity requires some minimal density of users, compensated by the relays in a way we make explicit. The existence of the above regimes brings interesting qualitative arguments to the discussion on the possible D2D deployment scenarios.

6.9. Determinantal thinning of point processes with network learning applications

In [39] submitted this year, a new type of dependent thinning for point processes in continuous space is proposed, which leverages the advantages of determinantal point processes defined on finite spaces and, as such, is particularly amenable to statistical, numerical, and simulation techniques. It gives a new point process that can serve as a network model exhibiting repulsion. The properties and functions of the new point process, such as moment measures, the Laplace functional, the void probabilities, as well as conditional (Palm) characteristics can be estimated accurately by simulating the underlying (non-thinned) point process, which can be taken, for example, to be Poisson. This is in contrast (and preference to) finite Gibbs point processes, which, instead of thinning, require weighting the Poisson realizations, involving usually intractable normalizing constants. Models based on determinantal point processes are also well suited for statistical (supervised) learning techniques, allowing the models to be fitted to observed network patterns with some particular geometric properties. We illustrate this approach by imitating with determinantal thinning the well-known Matérn II hard-core thinning, as well as a soft-core thinning depending on nearest-neighbour triangles. These two examples demonstrate how the proposed approach can lead to new, statistically optimized, probabilistic transmission scheduling schemes.

6.10. Analyzing LoRa long-range, low-power, wide-area networks using stochastic geometry

In [40] submitted this year, we present a simple, stochastic-geometric model of a wireless access network exploiting the LoRA (Long Range) protocol, which is a non-expensive technology allowing for long-range,

single-hop connectivity for the Internet of Things. We assume a space-time Poisson model of packets transmitted by LoRA nodes to a fixed base station. Following previous studies of the impact of interference, we assume that a given packet is successfully received when no interfering packet arrives with similar power before the given packet payload phase. This is as a consequence of LoRa using different transmission rates for different link budgets (transmissions with smaller received powers use larger spreading factors) and LoRa intra-technology interference treatment. Using our model, we study the scaling of the packet reception probabilities per link budget as a function of the spatial density of nodes and their rate of transmissions. We consider both the parameter values recommended by the LoRa provider, as well as proposing LoRa tuning to improve the equality of performance for all link budgets. We also consider spatially non-homogeneous distributions of LoRa nodes. We show also how a fair comparison to non-slotted Aloha can be made within the same framework.

6.11. Statistical learning of geometric characteristics of wireless networks

In [41] to appear in Proc. INFOCOM 2019, motivated by the prediction of cell loads in cellular networks, we formulate the following new, fundamental problem of statistical learning of geometric marks of point processes: An unknown marking function, depending on the geometry of point patterns, produces characteristics (marks) of the points. One aims at learning this function from the examples of marked point patterns in order to predict the marks of new point patterns. To approximate (interpolate) the marking function, in our baseline approach, we build a statistical regression model of the marks with respect some local point distance representation. In a more advanced approach, we use a global data representation via the scattering moments of random measures, which build informative and stable to deformations data representation, already proven useful in image analysis and related application domains. In this case, the regression of the scattering moments of the marked point patterns with respect to the non-marked ones is combined with the numerical solution of the inverse problem, where the marks are recovered from the estimated scattering moments. Considering some simple, generic marks, often appearing in the modeling of wireless networks, such as the shot-noise values, nearest neighbour distance, and some characteristics of the Voronoi cells, we show that the scattering moments can capture similar geometry information as the baseline approach, and can reach even better performance, especially for non-local marking functions. Our results motivate further development of statistical learning tools for stochastic geometry and analysis of wireless networks, in particular to predict cell loads in cellular networks from the locations of base stations and traffic demand.

6.12. Ressource allocation in bike sharing systems

Vehicle sharing systems are becoming an urban mode of transportation, and launched in many cities, as Velib' and Autolib' in Paris. Managing such systems is quite difficult. One of the major issues is the availability of the resources: vehicles or free slots. These systems became a hot topic in Operation Research and the importance of stochasticity on the system behavior leads us to propose mathematical stochastic models. The aim is to understand the system behavior and how to manage these systems in order to improve the allocation of both resources to users.

To improve BSS (bike-sharing systems), two types of policies can be deployed: incentives to the users to choose a better station, called *natural* or *green* regulation, or redistribution by trucks, called *active* regulation. In a simple mathematical model, we proved the efficiency of the 2-choice incentive policy for BSS (bike-sharing systems). The drawback of the model is that it ignores the geometry of the system, where the choice is only local. The purpose of this first work is to deal with this policy in real systems.

We use data trip data obtained from JCDecaux and reports on station status collected as open data, to test local choice policy. Indeed we designed and tested a new policy relying on a local small change in user behaviors, by adapting their trips to resource availability around their departure and arrival stations, based on 2-choice policy. Results show that, even with a small user collaboration, the proposed method increases significantly the global balance of the bike sharing system and therefore the user satisfaction. This is done using trip data sets and detecting spatial outliers, stations having a behavior significantly different from their spatial neighbors, in a context where neighbors are heavily correlated. For that we proposed an improved version of the well-known

Moran scatterplot method, using a robust distance metric called Gower similarity. Using this new version of Moran scatterplot, we show that, for the occupancy data set obtained by modifiying trips, the number of spatial outliers drastically decreases. We generalize this study with W. Ghanem and L. Massoulié testing incentive and redistribution policies on a simulator, where the tradeoff between the number of frustrated trips and the penalty for the users can be measured. We propose new versions of these policies including prediction.

6.13. Analyzing the choice of the least loaded queue between two neighboring queues

A model of N queues, with a local choice policy, is studied. Each one-server queue has a Poissonian arrival of customers. When a customer arrives at a queue, he joins the least loaded queue between this queue and the next one, ties solved at random. Service times have exponential distribution. The system is stable if the arrival-to-service rate ratio, also called load, is less than one. When the load tends to zero, we derive the first terms of the expansion in this parameter for the stationary probabilities that a queue has few customers. Then we provide explicit asymptotics, as the load tends to zero, for the stationary probabilities of the queue length. We used the analyticity of the stationary probabilities as a function of the load. It shows the behavior difference between this local choice policy and the 2-choice policy (*supermarket model*).

6.14. Optimal Content Replication and Request Matching in Large Caching Systems

We consider models of content delivery networks in which the servers are constrained by two main resources: memory and bandwidth. In such systems, the throughput crucially depends on how contents are replicated across servers and how the requests of specific contents are matched to servers storing those contents. In this paper, we first formulate the problem of computing the optimal replication policy which if combined with the optimal matching policy maximizes the throughput of the caching system in the stationary regime. It is shown that computing the optimal replication policy for a given system is an NP-hard problem. A greedy replication scheme is proposed and it is shown that the scheme provides a constant factor approximation guarantee. We then propose a simple randomized matching scheme which avoids the problem of interruption in service of the ongoing requests due to re-assignment or repacking of the existing requests in the optimal matching policy. The dynamics of the caching system is analyzed under the combination of proposed replication and matching schemes. We study a limiting regime, where the number of servers and the arrival rates of the contents are scaled proportionally, and show that the proposed policies achieve asymptotic optimality. Extensive simulation results are presented to evaluate the performance of different policies and study the behavior of the caching system under different service time distributions of the requests.

6.15. Statistical thresholds for Tensor PCA

This is a joint work with Aukosh Jagannath and Patrick Lopatto. We study the statistical limits of testing and estimation for a rank one deformation of a Gaussian random tensor. We compute the sharp thresholds for hypothesis testing and estimation by maximum likelihood and show that they are the same. Furthermore, we find that the maximum likelihood estimator achieves the maximal correlation with the planted vector among measurable estimators above the estimation threshold. In this setting, the maximum likelihood estimator exhibits a discontinuous BBP-type transition: below the critical threshold the estimator is orthogonal to the planted vector, but above the critical threshold, it achieves positive correlation which is uniformly bounded away from zero.

6.16. The distribution of the Lasso: Uniform control over sparse balls and adaptive parameter tuning

This is a joint work with Andrea Montanari. The Lasso is a popular regression method for high-dimensional problems in which the number of parameters $\theta_1, \cdots, \theta_N$, is larger than the number n of samples: $N > n$. A

useful heuristics relates the statistical properties of the Lasso estimator to that of a simple soft-thresholding denoiser, in a denoising problem in which the parameters $(\theta_i)_{i\le N}$ are observed in Gaussian noise, with a carefully tuned variance. Earlier work confirmed this picture in the limit $n, N \to \infty$, pointwise in the parameters θ, and in the value of the regularization parameter.

Here, we consider a standard random design model and prove exponential concentration of its empirical distribution around the prediction provided by the Gaussian denoising model. Crucially, our results are uniform with respect to θ belonging to ℓ_q balls, $q \in [0, 1]$, and with respect to the regularization parameter. This allows to derive sharp results for the performances of various data-driven procedures to tune the regularization.

Our proofs make use of Gaussian comparison inequalities, and in particular of a version of Gordon's minimax theorem developed by Thrampoulidis, Oymak, and Hassibi, which controls the optimum value of the Lasso optimization problem. Crucially, we prove a stability property of the minimizer in Wasserstein distance, that allows to characterize properties of the minimizer itself.

6.17. Phase transitions in spiked matrix estimation: information-theoretic analysis

We study here the so-called spiked Wigner and Wishart models, where one observes a low-rank matrix perturbed by some Gaussian noise. These models encompass many classical statistical tasks such as sparse PCA, submatrix localization, community detection or Gaussian mixture clustering. The goal of these notes is to present in a unified manner recent results (as well as new developments) on the information-theoretic limits of these spiked matrix/tensor models. We compute the minimal mean squared error for the estimation of the low-rank signal and compare it to the performance of spectral estimators and message passing algorithms. Phase transition phenomena are observed: depending on the noise level it is either impossible, easy (i.e. using polynomial-time estimators) or hard (information-theoretically possible, but no efficient algorithm is known to succeed) to recover the signal.

6.18. Accelerated decentralized optimization with local updates for smooth and strongly convex objectives

We study the problem of minimizing a sum of smooth and strongly convex functions split over the nodes of a network in a decentralized fashion. We propose the algorithm $ESDACD$, a decentralized accelerated algorithm that only requires local synchrony. Its rate depends on the condition number κ of the local functions as well as the network topology and delays. Under mild assumptions on the topology of the graph, $ESDACD$ takes a time $O((\tau_{\max} + \Delta_{\max})\sqrt{\kappa/\gamma}\ln(\epsilon^{-1}))$ to reach a precision ϵ where γ is the spectral gap of the graph, $\tau_{\max}$ the maximum communication delay and $\Delta_{\max}$ the maximum computation time. Therefore, it matches the rate of $SSDA$, which is optimal when $\tau_{\max} = \Omega(\Delta_{\max})$. Applying $ESDACD$ to quadratic local functions leads to an accelerated randomized gossip algorithm of rate $O(\sqrt{\theta_{\text{gossip}}/n})$ where θ_{gossip} is the rate of the standard randomized gossip. To the best of our knowledge, it is the first asynchronous gossip algorithm with a provably improved rate of convergence of the second moment of the error. We illustrate these results with experiments in idealized settings.

6.19. Group synchronization on grids

Group synchronization requires to estimate unknown elements $(\theta_v)_{v\in V}$ of a compact group $\mathbb{G}$ associated to the vertices of a graph $G = (V, E)$, using noisy observations of the group differences associated to the edges. This model is relevant to a variety of applications ranging from structure from motion in computer vision to graph localization and positioning, to certain families of community detection problems.

We focus on the case in which the graph G is the d-dimensional grid. Since the unknowns θ_v are only determined up to a global action of the group, we consider the following weak recovery question. Can we determine the group difference $\theta_u^{-1}\theta_v$ between far apart vertices u, v better than by random guessing? We prove that weak recovery is possible (provided the noise is small enough) for $d \ge 3$ and, for certain finite groups, for $d \ge 2$. Vice-versa, for some continuous groups, we prove that weak recovery is impossible for $d = 2$. Finally, for strong enough noise, weak recovery is always impossible.

6.20. An Impossibility Result for Reconstruction in a Degree-Corrected Planted-Partition Model

We consider the Degree-Corrected Stochastic Block Model (DC-SBM): a random graph on n nodes, having i.i.d. weights $(\phi_u)_{u=1}^n$ (possibly heavy-tailed), partitioned into $q \geq 2$ asymptotically equal-sized clusters. The model parameters are two constants $a, b > 0$ and the finite second moment of the weights $\Phi^{(2)}$. Vertices u and v are connected by an edge with probability $(\phi_u\phi_v/n)a$ when they are in the same class and with probability $(\phi_u\phi_v/n)b$ otherwise. We prove that it is information-theoretically impossible to estimate the clusters in a way positively correlated with the true community structure when $(a-b)2\Phi^{(2)} \leq q(a+b)$. As by-products of our proof we obtain (1) a precise coupling result for local neighbourhoods in DC-SBM's, that we use in a follow up paper [Gulikers et al., 2017] to establish a law of large numbers for local-functionals and (2) that long-range interactions are weak in (power-law) DC-SBM's.

6.21. On the capacity of information processing systems

We propose and analyze a family of information processing systems, where a finite set of experts or servers are employed to extract information about a stream of incoming jobs. Each job is associated with a hidden label drawn from some prior distribution. An inspection by an expert produces a noisy outcome that depends both on the job's hidden label and the type of the expert, and occupies the expert for a finite time duration. A decision maker's task is to dynamically assign inspections so that the resulting outcomes can be used to accurately recover the labels of all jobs, while keeping the system stable. Among our chief motivations are applications in crowd-sourcing, diagnostics, and experiment designs, where one wishes to efficiently learn the nature of a large number of items, using a finite pool of computational resources or human agents. We focus on the capacity of such an information processing system. Given a level of accuracy guarantee, we ask how many experts are needed in order to stabilize the system, and through what inspection architecture. Our main result provides an adaptive inspection policy that is asymptotically optimal in the following sense: the ratio between the required number of experts under our policy and the theoretical optimal converges to one, as the probability of error in label recovery tends to zero.

6.22. Optimal Algorithms for Non-Smooth Distributed Optimization in Networks

In this work, we consider the distributed optimization of non-smooth convex functions using a network of computing units. We investigate this problem under two regularity assumptions: (1) the Lipschitz continuity of the global objective function, and (2) the Lipschitz continuity of local individual functions. Under the local regularity assumption, we provide the first optimal first-order decentralized algorithm called multi-step primal-dual (MSPD) and its corresponding optimal convergence rate. A notable aspect of this result is that, for non-smooth functions, while the dominant term of the error is in $O(1/\sqrt{t})$, the structure of the communication network only impacts a second-order term in $O(1/t)$, where t is time. In other words, the error due to limits in communication resources decreases at a fast rate even in the case of non-strongly-convex objective functions. Under the global regularity assumption, we provide a simple yet efficient algorithm called distributed randomized smoothing (DRS) based on a local smoothing of the objective function, and show that DRS is within a $d^{1/4}$ multiplicative factor of the optimal convergence rate, where d is the underlying dimension.

6.23. Zap Meets Momentum: Stochastic Approximation Algorithms with Optimal Convergence Rate

There are two well known Stochastic Approximation techniques that are known to have optimal rate of convergence (measured in terms of asymptotic variance): the Ruppert-Polyak averaging technique, and stochastic Newton-Raphson (SNR)(a matrix gain algorithm that resembles the deterministic Newton-Raphson method). The Zap algorithms, introduced by Devraj and Meyn in 2017, are a version of SNR designed

to behave more closely like their deterministic cousin. It is found that estimates from the Zap Q-learning algorithm converge remarkably quickly, but the per-iteration complexity can be high. In [43], we introduce a new class of stochastic approximation algorithms based on matrix momentum. For a special choice of the matrix momentum and gain sequences, it is found in simulations that the parameter estimates obtained from the algorithm couple with those obtained from the more complex stochastic Newton-Raphson algorithm. Conditions under which coupling is guaranteed are established for a class of linear recursions. Optimal finite-n error bounds are also obtained.

6.24. Ergodic theory for controlled Markov chains with stationary inputs

Consider a stochastic process $\mathbf{X}$ on a finite state space $X = \{1, \cdots, d\}$. It is conditionally Markov, given a real-valued 'input process' ζ. This is assumed to be small, which is modeled through the scaling, $\zeta_t = \varepsilon\zeta^1_t$, $0 \le \varepsilon \le 1$, where ζ^1 is a bounded stationary process. The following conclusions are obtained, subject to smoothness assumptions on the controlled transition matrix and a mixing condition on ζ:

- A stationary version of the process is constructed, that is coupled with a stationary version of the Markov chain $\mathbf{X}^\bullet$ obtained with $\zeta \equiv 0$. The triple $(\mathbf{X}, \mathbf{X}^\bullet, \zeta)$ is a jointly stationary process satisfying $\mathsf{P}\{X(t) \ne X^\bullet(t)\} = O(\varepsilon)$. Moreover, a second-order Taylor-series approximation is obtained:

$$\mathsf{P}\{X(t) = i\} = \mathsf{P}\{X^\bullet(t) = i\} + \varepsilon^2 \varrho(i) + o(\varepsilon^2), \quad 1 \le i \le d,$$

  with an explicit formula for the vector $\varrho \in \mathfrak{R}^d$.

- For any $m \ge 1$ and any function $f : \{1, \cdots, d\} \times \mathfrak{R} \to \mathfrak{R}^m$, the stationary stochastic process $Y(t) = f(X(t), \zeta(t))$ has a power spectral density S_f that admits a second order Taylor series expansion: A function $\mathrm{S}^{(2)}{}_f : [-\pi, \pi] \to C^{m\times m}$ is constructed such that

$$\mathrm{S}_f(\theta) = \mathrm{S}^\bullet_f(\theta) + \varepsilon^2 \mathrm{S}^{(2)}{}_f(\theta) + o(\varepsilon^2), \quad \theta \in [-\pi, \pi].$$

  An explicit formula for the function $\mathrm{S}^{(2)}{}_f$ is obtained, based in part on the bounds in (i).

The results are illustrated using a version of the timing channel of Anantharam and Verdu.

6.25. Ordinary Differential Equation Methods for Markov Decision Processes and Application to Kullback–Leibler Control Cost

A new approach to computation of optimal policies for MDP (Markov decision process) models is introduced in [5], published in SICON this year. The main idea is to solve not one, but an entire family of MDPs, parameterized by a scalar ζ that appears in the one-step reward function. For an MDP with d states, the family of relative value functions $\{h^*_\zeta : \zeta \in \mathbb{R}\}$ is the solution to an ODE, $\frac{d}{d\zeta} h^*_\zeta = \mathcal{V}(h^*_\zeta)$, where the vector field $\mathcal{V} : R^d \to R^d$ has a simple form, based on a matrix inverse. Two general applications are presented: Brockett's quadratic-cost MDP model, and a generalization of the "linearly solvable" MDP framework of Todorov in which the one-step reward function is defined by Kullback–Leibler divergence with respect to nominal dynamics. This technique was introduced by Todorov in 2007, where it was shown under general conditions that the solution to the average-reward optimality equations reduce to a simple eigenvector problem. Since then many authors have sought to apply this technique to control problems and models of bounded rationality in economics. A crucial assumption is that the input process is essentially unconstrained. For example, if the nominal dynamics include randomness from nature (eg, the impact of wind on a moving vehicle), then the optimal control solution does not respect the exogenous nature of this disturbance. In [16] we introduce a technique to solve a more general class of action-constrained MDPs.

6.26. Distributed control design for balancing the grid using flexible loads

Inexpensive energy from the wind and the sun comes with unwanted volatility, such as ramps with the setting sun or a gust of wind. Controllable generators manage supply-demand balance of power today, but this is becoming increasingly costly with increasing penetration of renewable energy. It has been argued since the 1980s that consumers should be put in the loop: "demand response" will help to create needed supply-demand balance. However, consumers use power for a reason and expect that the quality of service (QoS) they receive will lie within reasonable bounds. Moreover, the behavior of some consumers is unpredictable, while the grid operator requires predictable controllable resources to maintain reliability.

The goal of the book chapter [31] is to describe an emerging science for demand dispatch that will create virtual energy storage from flexible loads. By design, the grid-level services from flexible loads will be as controllable and predictable as a generator or fleet of batteries. Strict bounds on QoS will be maintained in all cases. The potential economic impact of these new resources is enormous. California plans to spend billions of dollars on batteries that will provide only a small fraction of the balancing services that can be obtained using demand dispatch. The potential impact on society is enormous: a sustainable energy future is possible with the right mix of infrastructure and control systems.

In [17], presented at IEEE CDC 2018, a natural notion of *energy capacity* is proposed for the special case of thermostatically controlled loads (TCLs). It is shown that this quantity is closely approximated by thermal energy capacity, which is a component of the "leaky battery model" introduced in prior work. Simulation experiments in a distributed control setting show that these energy limits, and accompanying power capacity limits, are reliable indicators of online capacity, even for a heterogeneous population of loads. A feedforward/feedback control scheme is proposed for a large collection of heterogeneous loads. At the local level, control loops are used to create cooperative responses from each load in a given class of homogeneous loads. This simplifies control of the aggregate based on two pieces of information: aggregate power consumption from each class of loads and the *state of charge* surrogate that is a part of the leaky battery model. This information is required at a slow time-scale (say, 5 minute sampling).

In [18], we study the problem of coordination of a collection of on/off thermostatically controlled loads (TCLs) to act as a "virtual battery". Virtual Energy Storage (VES) is provided by the collection by either consuming more (charging) or less (discharging) power than the baseline. VES can be an inexpensive alternative to batteries when a large share of the electricity comes from volatile sources such as solar and wind. Almost all prior work has assumed that the outside weather - which significantly effects a TCLs behavior - is constant. We combine the above distributed load control design with a grid level MPC (model predictive control) that uses predictions of disturbances (weather) over a planning horizon. Additionally, irrespective of the choice of control architecture, there is a fundamental limit to the power and energy capacity of the collection of TCLs. We partially address this issue by scaling the reference signal by a function of the outside air temperature.

6.27. Estimation and control of quality of service in demand dispatch

Flexibility of energy consumption can be harnessed for the purposes of grid-level ancillary services. In particular, through distributed control of a collection of loads, a balancing authority regulation signal can be tracked accurately, while ensuring that the quality of service (QoS) for each load is acceptable on average. Subject to distributed control approaches advocated in recent research, the histogram of QoS is approximately Gaussian, and consequently, each load will eventually receive poor service. In [11], published this year in IEEE Transactions on Smart Grid, statistical techniques are developed to estimate the mean and variance of QoS as a function of the power spectral density of the regulation signal. It is also shown that additional local control can eliminate risk. The histogram of QoS is truncated through this local control, so that strict bounds on service quality are guaranteed. While there is a tradeoff between the grid-level tracking performance (capacity and accuracy) and the bounds imposed on QoS, it is found that the loss of capacity is minor in typical cases.

The previous designs for distributed control of TCLs ensure that the indoor temperature remains within a pre-specified bound, but other QoS metrics, especially the frequency of turning on and off was not limited. In [19], presented at ACM BuildSys 2018, we propose a more advanced control architecture that reduces the

cycling rate of TCLs. We show through simulations that the proposed controller is able to reduce the cycling of individual TCLs compared to the previous designs with little loss in tracking of the grid-supplied reference signal.

6.28. Optimal control of energy storage

Energy storage revenue estimation is essential for analyzing financial feasibility of investment in batteries. In [22], we quantify the cycles of operation considering depth-of-discharge (DoD) of operational cycles and provide an algorithm to calculate equivalent 100% DoD cycles. This facilitates in comparing cycles of different DoDs. The battery life is frequently defined as a combination of cycle and calendar life. We propose a battery capacity degradation model based on the cycle and the calendar life and operational cycles. Using equivalent 100% DoD cycles and revenue generated, we calculate the dollars per cycle revenue of storage performing electricity price based arbitrage and ancillary services for load balancing in real time. Using PJM's (a regional transmission organization in the United States) real data we calculate short term and long term financial potential for the year of 2017. We observe that participating in ancillary services is significantly more beneficial for storage owners compared to participating in energy arbitrage.

Battery life is often described a combination of cycle life and calendar life. In [21], we propose a mechanism to limit the number of cycles of operation over a time horizon in an optimal arbitrage algorithm proposed in our previous work. The cycles of operation have to be tuned based on price volatility to maximize the battery life and arbitrage gains.

In [23], we analyze the effect of real time electricity price (RTP) on the amount of ancillary services required for load balancing in presence of responsive users, information asymmetry and forecast errors in demand and renewable energy sources (RES) generation. We consider a RTP that is determined by the forecasted generation and ramping cost. A community choice aggregator manages the load of all the consumers by setting the price. The consumer's objective is to minimize their overall cost of consumption. Ancillary services are called upon to balance the load in real time. With zero RES in the power network and a high degree of load flexibility, the proposed RTP flattens and the volatility in demand vanishes. However, in presence of RES the volatility in price and demand is reduced up to an extent and ancillary services are required for load balancing. The amount of ancillary services required increases with forecast errors. We also propose a real time algorithm that approximates the optimal consumer behavior under the complete information setting. Extensive numerical simulations are provided using real data from Pecan Street and Elia Belgium.

6.29. Dynamic matching models

The model of First Come First Served infinite bipartite matching was introduced in Caldentey, Kaplan and Weiss, 2009. In this model, there is a sequence of items that are chosen i.i.d. from a finite set $\mathcal{C}$ and an independent sequence of items that are chosen i.i.d. from a finite set $\mathcal{S}$, and a bipartite compatibility graph G between $\mathcal{C}$ and $\mathcal{S}$. Items of the two sequences are matched according to the compatibility graph, and the matching is FCFS, meaning that each item in the one sequence is matched to the earliest compatible unmatched item in the other sequence. In Adan and Weiss, 2012, a Markov chain associated with the matching was analyzed, a condition for stability was derived, and a product form stationary distribution was obtained. In [2], we present several new results that unveil the fundamental structure of the model. First, we provide a pathwise Loynes' type construction which enables to prove the existence of a unique matching for the model defined over all the integers. Second, we prove that the model is dynamically reversible: we define an exchange transformation in which we interchange the positions of each matched pair, and show that the items in the resulting permuted sequences are again independent and i.i.d., and the matching between them is FCFS in reversed time. Third, we obtain product form stationary distributions of several new Markov chains associated with the model. As a by-product, we compute useful performance measures, for instance the link lengths between matched items.

In [51], we propose an explicit construction of the stationary state of Extended Bipartite Matching (EBM) models, as defined in (Busic et. al., 2013). We use a Loynes-type backwards scheme similar in flavor to that in (Moyal et al., 2017), allowing to show the existence and uniqueness of a bi-infinite perfect matching under various conditions, for a large class of matching policies and of bipartite matching structures. The key algebraic element of our construction is the sub-additivity of a suitable stochastic recursive representation of the model, satisfied under most usual matching policies. By doing so, we also derive stability conditions for the system under general stationary ergodic assumptions, subsuming the classical markovian settings.

In [42], we consider holding costs for the items that are waiting to be matched. We model this problem as an MDP (Markov decision process) and study the discounted cost and the average cost case. We first consider a model with two types of supply and two types of demand items with an N matching graph. For linear cost function, we prove that an optimal matching policy gives priority to the end edges of the matching graph and is of threshold type for the diagonal edge. In addition, for the average cost problem, we compute the optimal threshold value. According to our preliminary numerical experiments, threshold-type policies performs also very well for more general bipartite graphs.

EVA Project-Team

7. New Results

7.1. From SmartMarina to Falco

Participants: Keoma Brun-Laguna, Thomas Watteyne.

SmartMarina project (http://smartmarina.org/) was a technical project in 2017 to study the feasibility of using the wireless technology developed at Inria-EVA for marina management. In 2018, the Wattson Elements company was born, which now commercializes the Falco solution (https://wefalco.fr/).

Figure 1. Screenshot of the Falco promotional video, https://youtu.be/35HdoFLrCf0.

7.2. 6TiSCH Standardization

Participants: Malisa Vucinic, Jonathan Muñoz, Tengfei Chang, Yasuyuki Tanaka, Thomas Watteyne.

The standardization work at 6TiSCH remains a strong federator of the work done in the team. In 2018, the working group published the specification of the 6TiSCH Operation Sublayer (6top) Protocol, RFC8480. Work is also ongoing in the fragment forwarding space, where we are working on how ot efficient forward long IPv6 packets which are fragmented to fit in short IEEE 802.15.4 frames.

7.3. 6TiSCH Security

Participants: Malisa Vucinic, Thomas Watteyne.

The security work of Inria-EVA revolves around 6TiSCH networks and is a continuation of the efforts started during the H2020 ARMOUR project. The work focused on stabilizing the "Minimal Security" solution that has now passed the working group last call in the IETF and is pending finals reviews before being published as an RFC. The solution that is standardized enables secure network access and configuration of 6TiSCH devices under the assumption that they have been provisioned with a secret key. Ongoing work extends this solution to support true zero-configuration network setup, under the assumption that the devices have been provisioned with certificates at manufacturing time.

7.4. 6TiSCH Benchmarking

Participants: Malisa Vucinic, Tengfei Chang, Yasuyuki Tanaka, Thomas Watteyne.

With the pure 6TiSCH standardizes coming to an end, the focus of the group is moving towards benchmarking how well it works. This has results in the following action. Although seemingly different, they all contribute to the overall goal of better understand (the performance of) 6TiSCH.

We have built and put online the OpenTestbed, a collection of 80 OpenMote B boards deployed in 20 "pods". These allow us to test the performance of the OpenWSN firmware in a realistic setting. The testbed is depicted in Fig. 2 . You can access its management interface at http://testbed.openwsn.org/.

Figure 2. The OpenTestbed deployed in Inria Paris since July 2018.

A tool complementary to the testbed is the 6TiSCH simulator (https://bitbucket.org/6tisch/simulator) which Yatsuyiki Tanaka is leading. The simulator now represents exactly the behavior of the 6TiSCH protocol stack, and has been a catalyst for benchmarking activities around 6TiSCH.

Beyond Inria, the benchmarking activity around 6TiSCH is a hot topic, with projects such as the 6TiSCH Open Data Action (SODA, http://www.soda.ucg.ac.me/), the IoT Benchmarks Initiative (https://www.iotbench.ethz.ch/), and the Computer and Networking Experimental Research using Testbeds (CNERT) workshop at INFOCOM, all of which Inria-EVA is very involved in.

7.5. IoT and Wireless Sensor Networks

More than 50 billions of devices will be connected in 2020. This huge infrastructure of devices, which is managed by highly developed technologies, is called Internet of Things (IoT). The latter provides advanced services, and brings economical and societal benefits. This is the reason why engineers and researchers of both industry and scientific communities are interested in this area. The Internet of Things enables the interconnection of smart physical and virtual objects, managed by highly developed technologies. WSN (Wireless Sensor Network), is an essential part of this paradigm. The WSN uses smart, autonomous and usually limited capacity devices in order to sense and monitor their environment.

7.5.1. Distributed Scheduling for IEEE 802.15.4e TSCH networks

Participants: Yasuyuki Tanaka, Pascale Minet, Thomas Watteyne.

Since the scheduling algorithm is not standardized for IEEE 802.15.4e TSCH networks, many scheduling algorithms have been proposed. Most of them are centralized, few are distributed. Among the distributed scheduling algorithms, many rely on assumptions that may be violated by real deployments. This violation usually leads to conflicting transmissions of application data, decreasing the reliability and increasing the latency of data delivery. Others require a processing complexity that cannot be provided by sensor nodes of limited capabilities. Still others are unable to adapt quickly to traffic or topology changes, or are valid only for small traffic loads.

In the study funded by the Inria ADT DASMU (Action de Developement Technologique Distributed Adaptive Scheduling for MUltichannel wireless sensor networks), we focus on a distributed scheduling algorithm that relies on realistic assumptions, does not require complex computation, is valid for any traffic load, is adaptive and compliant with the standardized protocols used in the 6TiSCH working group at IETF.

First results have been obtained and an intensive simulation campaign made with the 6TiSCH simulator has provided comparative performance results. Our proposal outperforms MSF, the 6TiSCH Minimal Scheduling Function, in terms of end-to-end latency and end-to-end packet delivery ratio. More evaluations are needed to improve the proposal (e.g. less packet drops during transient situations, less overhead) in terms of scheduled cells).

7.5.2. IoT and IEEE 802.15.4e TSCH networks

Participants: Pascale Minet, Ines Khoufi, Zied Soua.

In 2018, we focus on how an IEEE 802.15.4e is autonomously built and how nodes join the network.

To join the TSCH network, a device randomly selects a physical channel used by this network and listens to a beacon advertising this network. Since the physical channel on which the beacon is broadcast changes at each beacon slot due to channel hopping, the joining device will eventually hear a beacon sent by one of its neighbors. Upon receipt of a valid beacon, this device gets synchronized with the TSCH network.

In this study, we focus on the time needed by a node to detect a beacon sent by a TSCH network, as well as on the time needed to build a TSCH network. These times are important for industrial applications where new nodes are inserted progressively, or when failed nodes are replaced. Both times highly depend on the beacon advertisement policy, policy that is not specified in the standard and is under the responsibility of a layer upper than the MAC one. Since beacons are broadcast, they are lost in case of collisions: the vital information they carry is lost. The main problem is how to avoid collisions between two devices that are not neighbors.

That is why we propose the Enhanced Deterministic Beacon Advertising algorithm, called EDBA, that ensures a collision-free advertising of beacons. Since the beacon cells are fairly distributed in the slotframe, the average joining time is minimized. The behavior of a joining node has been modeled by a Markov chain from which the average joining time is computed, taking into account the reliability of wireless links. An intensive performance evaluation based on NS3 simulations allows us to validate this model and conclude on the very good performance of EDBA, even when compared with MBS, considered as the best advertising algorithm in the literature. These results have been published in the Annals of Telecommunications, [10].

7.5.3. UAV-based Data Gathering

Participants: Nadjib Achir (Paris 13), Tounsia Djamah, Paul Muhlethaler, Celia Tazibt (Paris 13).

The recent advances in wireless sensors and Unmanned Aerial Vehicles have created new opportunities for environmental control and low cost aerial data gathering. We propose to use an Unmanned Aerial Vehicle (UAV) for data gathering [36]. Basically, we have proposed a method for UAV path planning based on virtual forces and potential fields. In addition, and more importantly, we present a new approach to compute the attractive forces of the potential field.

We use as our starting point the idea used by Pereira of using a potential field approach. However, we extend this work by considering that each cell in the area apply an attractive force on the drone, not only the deployed sensors. We compared our results with those obtained with Pereira's method and we obtained better performance in terms of data collection time. In other words, for the same period of time our method collect more data. The second advantage of our approach is that it leads to a significant reduction in the distance that the drone must travel.

7.5.4. Towards evaluating Named Data Networking for the IoT: A framework for OMNeT++

Participants: Amar Abane, Samia Bouzefrane (Cnam), Paul Muhlethaler.

Named Data Networking is a promising architecture for emerging Internet applications such as the Internet of Things (IoT). Many studies have already investigated how NDN can be an alternative for IP in future IoT deployments. However, NDN-IoT propositions need accurate evaluation at network level and system level as well. We introduce an NDN framework for OMNeT++ [29]. Designed for low-end devices and gateways of the IoT, the framework is capable of simulating NDN scenarios at the boundary of the network and the system. The framework implementation is presented and used to study a typical aspect of NDN integration in IoT devices.

7.5.5. Evaluation of LORA with stochastic geometry

Participants: Bartek Blaszczyszyn (Dyogen), Paul Muhlethaler.

We present a simple, stochastic-geometric model of a wireless access network exploiting the LoRA (Long Range) protocol, which is a non-expensive technology allowing for long-range, single-hop connectivity for the Internet of Things. We assume a space-time Poisson model of packets transmitted by LoRA nodes to a fixed base station. Following previous studies of the impact of interference, we assume that a given packet is successfully received when no interfering packet arrives with similar power before the given packet payload phase, see [39]. This is as a consequence of LoRa using different transmission rates for different link budgets (transmissions with smaller received powers use larger spreading factors) and LoRa intra-technology interference treatment. Using our model, we study the scaling of the packet reception probabilities per link budget as a function of the spatial density of nodes and their rate of transmissions. We consider both the parameter values recommended by the LoRa provider, as well as proposing LoRa tuning to improve the equality of performance for all link budgets. We also consider spatially non-homogeneous distributions of LoRa nodes. We show also how a fair comparison to non-slotted Aloha can be made within the same framework.

7.5.6. Position Certainty Propagation: A location service for MANETs

Participants: Abdallah Sobehy, Paul Muhlethaler, Eric Renault (Telecom Sud-Paris).

Localization in Mobile Ad-hoc Networks (MANETs) and Wireless Sensor Networks (WSNs) is an issue of great interest, especially in applications such as the IoT and VANETs. We propose a solution that overcomes two limiting characteristics of these types of networks. The first is the high cost of nodes with a location sensor (such as GPS) which we will refer to as anchor nodes. The second is the low computational capability of nodes in the network. The proposed algorithm [28] addresses two issues; self-localization where each non-anchor node should discover its own position, and global localization where a node establishes knowledge of the position of all the nodes in the network. We address the problem as a graph where vertices are nodes in the network and edges indicate connectivity between nodes. The weights of edges represent the Euclidean distance between the nodes. Given a graph with at least three anchor nodes and knowing the maximum communication range for each node, we are able to localize nodes using fairly simple computations in a moderately dense graph.

7.6. Industry 4.0 and Low-Power Wireless Meshed Networks

7.6.1. Deterministic Networking for the Industrial Internet of Things (IIoT)

Participants: Keoma Brun-Laguna, Thomas Watteyne, Pascale Minet.

The Internet of Things (IoT) connects tiny electronic devices able to measure a physical value (temperature, humidity, etc.) and/or to actuate on the physical world (pump, valve, etc). Due to their cost and ease of deployment, battery-powered wireless IoT networks are rapidly being adopted.

The promise of wireless communication is to offer wire-like connectivity. Major improvements have been made in that sense, but many challenges remain as industrial application have strong operational requirements. This section of the IoT application is called Industrial IoT (IIoT).

The main IIoT requirement is reliability. Every bit of information that is transmitted in the network must not be lost. Current off-the-shelf solutions offer over 99.999% reliability.

Then come latency and energy-efficiency requirements. As devices are battery-powered, they need to consume as little as possible to be able to operate during years. The next step for the IoT is to target time-critical applications.

Industrial IoT technologies are now adopted by companies over the world, and are now a proven solution. Yet, challenges remain and some of the limits of the technologies are still not fully understood. In his PhD Thesis, Keoma Brun-Laguna addresses TSCH-based Wireless Sensor Networks and studies their latency and lifetime limits under real-world conditions.

We gathered 3M network statistics 32M sensor measurements on 11 datasets with a total of 170,037 mote hours in real-world and testbeds deployments. We assembled what we believed to be the largest dataset available to the networking community.

Based on those datasets and on insights we learned from deploying networks in real-world conditions, we study the limits and trade-offs of TSCH-based Wireless Sensor Networks. We provide methods and tools to estimate the network performances of such networks in various scenarios. We highlight the trade-off between short latency and long network lifetime. We believe we assembled the right tools for protocol designer to build deterministic networking to the Industrial IoT.

7.6.2. Industry 4.0 and IEEE 802.15.4e TSCH networks

Participants: Pascale Minet, Ines Khoufi, Zied Soua.

By the year 2020, it is expected that the number of connected objects will exceed several billions devices. These objects will be present in everyday life for a smarter home and city as well as in future smart factories that will revolutionize the industry organization. This is actually the expected fourth industrial revolution, more known as Industry 4.0. In which, the Internet of Things (IoT) is considered as a key enabler for this major transformation. IoT will allow more intelligent monitoring and self-organizing capabilities than traditional factories. As a consequence, the production process will be more efficient and flexible with products of higher quality.

To produce better quality products and improve monitoring in Industry 4.0, strong requirements in terms of latency, robustness and power autonomy have to be met by the networks supporting the Industry 4.0 applications. The wireless TSCH (Time Slotted Channel Hopping) network specified in the e amendment of the IEEE 802.15.4 standard has many appealing properties. Its schedule of multichannel slotted data transmissions ensures the absence of collisions. Because there is no retransmission due to collisions, communication is faster. Since the devices save energy each time they do not take part in a transmission, the power autonomy of nodes is prolonged. Furthermore, channel hopping enables to mitigate multipath fading and interferences.

To increase the flexibility and the self-organizing capacities required by Industry 4.0, the networks have to be able to adapt to changes. These changes may concern the application itself, the network topology by adding or removing devices, the traffic generated by increasing or decreasing the device sampling frequency, for instance. That is why the flexibility of the schedule ruling all network communications is needed.

In 2018, we show how a TSCH network can adapt to such changes. More precisely, we propose a solution ranging from network construction to data gathering. We show how a TSCH network is autonomously built, supports data gathering and is able to adapt to changes in network topology, traffic and application requirements.

The solution proposed preserves the merits of TSCH network, that can be listed hereafter. The time-slotted multichannel medium access enables parallel transmissions on several channels, leading to shorter latency and higher throughputs. In addition, channel hopping mitigates interference and multipath effects. Furthermore, since transmissions are scheduled, a conflict-free schedule is computed by the network coordinator (i.e. the CPAN). Hence, no collision occurs during data gathering. The absence of collision leads to a higher throughput, because there is no retransmission due to collisions. It also preserves nodes power autonomy.

This simple solution is based on the coexistence of several periodic slotframes. We distinguish three slotframes, which are the Beacon Slotframe, the Data Slotframe and the Shared Slotframe. The network schedule corresponds to the superposition of the three schedules given by each slotframe, where the slotframe with the highest priority wins.

This solution ensures a collision-free dissemination over the whole network. Beacons are broadcast in sequence by increasing depth of devices. This broadcast is also used to disseminate Data Schedules (new schedule or update).

In addition, this solution is adaptive. Topology, traffic or application changes are notified to the CPAN. Depending on the changes notified, the CPAN updates the current schedule or recomputes a new one. Shared slots are used to cope with unexpected events.

We compute the theoretical bounds with regard to key performance indicators and compare them with the values obtained by NS3 simulation. Simulation results confirm the theoretical upper bounds computed for network construction and data gathering. Hence, TSCH networks are able to adapt to traffic or topology changes in a reasonable time which is a strong requirement of Industry 4.0 applications. These results have been presented at the PEMWN 2018 conference in [26]. In some further work, we will study how to improve this delay to support the most demanding applications.

7.7. Machine Learning for an efficient and dynamic management of data centers

7.7.1. Data Analysis in Data Centers

Participants: Eric Renault (Telecom Sud-Paris), Selma Boumerdassi (Cnam), Pascale Minet, Ines Khoufi.

In High Performance Computing (HPC), it is assumed that all machines are homogeneous in terms of CPU and memory capacities, and that the tasks making up the jobs have similar resource requests. It has been shown that this homogeneity relating both to machine capacity and workload, although generally valid for HPC, does no longer apply to data centers. This explains why the publication of data gathered in an operational Google data center over 29 days has aroused great interest among researchers.

It is crucial to have real traces of a Google data center publicly available that are representative of the functioning of real data centers. Our goal is to analyze the data collected and to draw useful conclusions about machines, jobs and tasks as well as resource usage. Our main results have been published in [25], [24] and can be summarized as follows:

- Although 92% of machines have a CPU capacity of 0.5, there are 10 machine configurations in the data center, each configuration is characterized by a pair $(CPU\ capacity, memory\ capacity)$. The most frequent configuration is supported by only 53% of machines.
- Over the 29 days, all the machines in the data center that were removed, were restarted later after an off-period. 50% of these periods have a duration less than or equal to 1000 seconds (i.e. 16.66 minutes), suggesting a maintenance operation.
- The distribution of jobs per category reveals only one job, representing 0.002%, for the Infrastructure, 0.13% of jobs for Monitoring, 9.91% of jobs for Production, 56.30% of jobs for Other, and 33.63% of jobs for Free. 92.05% of jobs have a single task. 95.75% have fewer than 10 tasks. But 12 jobs have 5000 tasks and 114 jobs have around 1000 tasks.
- With regard to resource requests, 0.11% of jobs have a memory request and a CPU request higher than or equal to 10%.
- 94.25% of jobs wait less than 10 seconds before being scheduled. However, some of them wait for more than 1000 seconds. Such large values could be explained by the existence of placement constraints for the jobs, making them harder to place and schedule. 49% of jobs have an execution time less than 100 seconds.

Such results are needed to validate or invalidate some simplifying assumptions that are usually made when reasoning about models, and make the models more accurate for jobs and tasks as well as for available machines. Having validated these models on real data centers, they can then be used for extensive evaluation of placement and scheduling algorithms and more generally for resource allocation (i.e. CPU and memory). These algorithms can then be applied in real data centers.

Another possible use of this data set is to consider it as a learning set in order to predict some feature of the data center, such as the workload of hosts or the next arrival of jobs.

7.7.2. Machine Learning for an Energy-Efficient Management of Data Centers

Participants: Ruben Milocco (University Of Camahue, Argentina), Pascale Minet, Eric Renault (Telecom Sud-Paris), Selma Boumerdassi (Cnam).

To limit global warming, all industrial sectors must make effort to reduce their carbon footprint. Information and Communication Technologies (ICTs) alone generate 2% of global CO2 emissions every year. Due to the rapid growth in Internet services, data centers have the largest carbon footprint of all ICTs. According to ARCEP (the French telecommunications regulator), Internet data traffic multiplied by 4.5 between 2011 and 2016. In order to support such a growth and maintain this traffic, data centers'energy consumption needs to be optimized. The problem of managing Data Centers (DC) and clouds optimally, in the sense that the demand is met with a minimal energy cost, remains a major issue. In this research, we evaluate the maximum energy saving that can be obtained in DCs by means of a proactive management of resources. The proposed management is based on models that predict resource requests.

Diverse approaches to obtain predictive models of DCs have been studied recently. Among the most popular methods with the comparatively lowest prediction errors are the predictive models of the ARMAX family. Hence, we study the predictive model given by the ARMAX family. We compare its performance with that of the Last Value (LV) model which predicts that the next value will be equal to the current one. To the best of our knowledge, there are no studies relating to the performance bounds that can be achieved using these models. In this research, we study the limits of the improvement in terms of energy cost that can be obtained using proactive strategies for DC management based on predictive models.

Using the Google dataset collected over a period of 29 days and made publicly available, we evaluate the largest benefit that can be obtained with those two predictors.

7.8. Protocols and Models for Wireless Networks - Application to VANETs

7.8.1. Predicting Vehicles Positions using Roadside Units: a Machine-Learning Approach

Participants: Samia Bouzefrane (Cnam), Soumya Banerjee (Birla Institute Of Technology, Mesra), Paul Mühlethaler, Mamoudou Sangare.

We study positioning systems using Vehicular Ad Hoc Networks (VANETs) to predict the position of vehicles [35]. We use the reception power of the packets received by the Road Side Units (RSUs) and sent by the vehicles on the roads. In fact, the reception power is strongly influenced by the distance between a vehicle and a RSU. To predict the position of vehicles in this context, we adopt the machine learning methodology. As a pre-requisite, the vehicles know their positions and the vehicles send their positions in the packets. The positioning system can thus perform a training sequence and build a model. The system is then able to handle a prediction request. In this request, a vehicle without external positioning will request its position from the neighboring RSUs. The RSUs which receive this request message from the vehicle will know the power at which the message was received and will study the positioning request using the training set. In this study, we use and compare three widely recognized techniques : K Nearest Neighbors (KNN), Support Vector Machine (SVM) and Random Forest. We study these techniques in various configurations and discuss their respective advantages and drawbacks. Our results show that these three techniques provide very good results in terms of position predictions when the error on the transmission power is small.

7.8.2. Predicting transmission success with Machine-Learning and Support Vector Machine in VANETs

Participants: Samia Bouzefrane (Cnam), Soumya Banerjee (Birla Institute Of Technology, Mesra), Paul Mühlethaler, Mamoudou Sangare.

We study the use of the Support Vector Machine technique to estimate the probability of the reception of a given transmission in a Vehicular Ad hoc NETwork (VANET). The transmission takes place between a vehicle and a RoadSide Unit (RSU) at a given distance and with a given transmission rate. The RSU computes the statistics of the receptions and is able to compute the percentage of successful transmissions versus the distance between the vehicle and the RSU and the transmission rate. Starting from this statistic, a Support Vector Machine (SVM) scheme can produce a model. Then, given a transmission rate and a distance between the vehicle and the RSU, the SVM technique can estimate the probability of a succcessful reception. This probability can be used to build an adaptive technique which optimizes the expected throughput between the vehicle and the RSU. Instead of using transmission values of a real experiment, we use the results of an analytical model of CSMA that is customized for 1D VANETs. The model we adopt to perform this task uses a Matern selection process to mimic the transmission in a CSMA IEEE 802.11p VANET. With this model we obtain a closed formula for the probability of successful transmissions. Thus with these results we can train an SVM model and predict other values for other couples : distance, transmission rate. The numerical results we obtain show that SVM seems very suitable to predict the reception probability in a VANET.

7.8.3. TDMA scheduling strategies for vehicular ad hoc networks: from a distributed to a centralized approach

Participants: Mohammed Hadded, Anis Laouiti (Telecom Sud-Paris, Paul Mühlethaler.

We focus on vehicular safety applications based on the Dedicated Short Range Communication (DSRC) standard. We propose a new mechanism to alleviate channel congestion by reducing the beacons load while maintaining an accurate awareness level. Our scheme is based on the collective perception concept which consists in sharing perceived status information collected by vehicles equipped with different types of sensors (radars, lidars, cameras, etc.). To achieve our goal, we propose two main schemes [30]. The first one consists in implementing the collective perception capability on vehicles and adding a new category of status messages to share locally collected sensor data in order to reduce channels load and enhance vehicles' awareness. The second scheme concerns the accuracy level of the received information from the collective perception enabled vehicles by fixing a prior error threshold on the position. The method proposed is validated by simulations and

the results obtained are compared to those of an application based on the traditional beaconing scheme of the IEEE802.11p standard. The simulations show that the proposed scheme is able to significantly reduce the load on the control channel incurred by the beacons and the packet error ratio for different network densities and built-in sensors characteristics.

7.8.4. A Collaborative Environment Perception Approach for Vehicular Ad hoc Networks

Participants: Sadia Ingrachen, Nadjib Achir (Paris 13), Paul Mühlethaler, Tounsia Djamah (Paris 13), Amine Berqia (Paris 13).

We focus on vehicular safety applications based on the Dedicated Short Range Communication (DSRC) standard. We propose a new mechanism to alleviate channel congestion by reducing the beacons load while maintaining an accurate awareness level. Our scheme is based on the collective perception concept which consists in sharing perceived status information collected by vehicles equipped with different types of sensors (radars, lidars, cameras, etc.). To achieve our goal, we propose two main schemes [31]. The first one consists in implementing the collective perception capability on vehicles and adding a new category of status messages to share locally collected sensor data in order to reduce channels load and enhance vehicles' awareness. The second scheme concerns the accuracy level of the received information from the collective perception enabled vehicles by fixing a prior error threshold on the position. The method proposed is validated by simulations and the results obtained are compared to those of an application based on the traditional beaconing scheme of the IEEE802.11p standard. The simulations show that the proposed scheme is able to significantly reduce the load on the control channel incurred by the beacons and the packet error ratio for different network densities and built-in sensors characteristics.

GANG Project-Team

7. New Results

7.1. Graph and Combinatorial Algorithms

7.1.1. Random Walks with Multiple Step Lengths

In nature, search processes that use randomly oriented steps of different lengths have been observed at both the microscopic and the macroscopic scales. Physicists have analyzed in depth two such processes on grid topologies: *Intermittent Search*, which uses two step lengths, and *Lévy Walk*, which uses many. Taking a computational perspective, in [26] we consider the number of distinct step lengths k as a *complexity measure* of the considered process. Our goal is to understand what is the optimal achievable time needed to cover the whole terrain, for any given value of k. Attention is restricted to dimension one, since on higher dimensions, the simple random walk already displays a quasi linear cover time.

We say X is a *k-intermittent search* on the one dimensional n-node cycle if there exists a probability distribution $\mathbf{p} = (p_i)_{i=1}^{k}$, and integers $L_1, L_2, ..., L_k$, such that on each step X makes a jump $\pm L_i$ with probability p_i, where the direction of the jump ($+$ or $-$) is chosen independently with probability $1/2$. When performing a jump of length L_i, the process consumes time L_i, and is only considered to visit the last point reached by the jump (and not any other intermediate nodes). This assumption is consistent with biological evidence, in which entities do not search while moving ballistically. We provide upper and lower bounds for the cover time achievable by k-intermittent searches for any integer k. In particular, we prove that in order to reduce the cover time $\Theta(n^2)$ of a simple random walk to $\widetilde{\Theta}(n)$, roughly $\frac{\log n}{\log \log n}$ step lengths are both necessary and sufficient, and we provide an example where the lengths form an exponential sequence.

In addition, inspired by the notion of intermittent search, we introduce the *Walk or Probe* problem, which can be defined with respect to arbitrary graphs. Here, it is assumed that querying (probing) a node takes significantly more time than moving to a random neighbor. Hence, to efficiently probe all nodes, the goal is to balance the time spent walking randomly and the time spent probing. We provide preliminary results for connected graphs and regular graphs.

7.1.2. Searching a Tree with Permanently Noisy Advice

In [16], we consider a search problem on trees using unreliable guiding instructions. Specifically, an agent starts a search at the root of a tree aiming to find a treasure hidden at one of the nodes by an adversary. Each visited node holds information, called *advice*, regarding the most promising neighbor to continue the search. However, the memory holding this information may be unreliable. Modeling this scenario, we focus on a probabilistic setting. That is, the advice at a node is a pointer to one of its neighbors. With probability q each node is *faulty*, independently of other nodes, in which case its advice points at an arbitrary neighbor, chosen uniformly at random. Otherwise, the node is *sound* and points at the correct neighbor. Crucially, the advice is *permanent*, in the sense that querying a node several times would yield the same answer. We evaluate efficiency by two measures: The *move complexity* denotes the expected number of edge traversals, and the *query complexity* denotes the expected number of queries.

Let Δ denote the maximal degree. Roughly speaking, the main message of this paper is that a phase transition occurs when the *noise parameterq* is roughly $1/\sqrt{\Delta}$. More precisely, we prove that above the threshold, every search algorithm has query complexity (and move complexity) which is both exponential in the depth d of the treasure and polynomial in the number of nodes n. Conversely, below the threshold, there exists an algorithm with move complexity $O(d\sqrt{\Delta})$, and an algorithm with query complexity $O(\sqrt{\Delta}\log\Delta\log^2 n)$. Moreover, for the case of regular trees, we obtain an algorithm with query complexity $O(\sqrt{\Delta}\log n\log\log n)$. For q that is below but close to the threshold, the bound for the move complexity is tight, and the bounds for the query complexity are not far from the lower bound of $\Omega(\sqrt{\Delta}\log_\Delta n)$.

In addition, we also consider a *semi-adversarial* variant, in which faulty nodes are still chosen at random, but an adversary chooses (beforehand) the advice of such nodes. For this variant, the threshold for efficient moving algorithms happens when the noise parameter is roughly $1/\Delta$. In fact, above this threshold a simple protocol that follows each advice with a fixed probability already achieves optimal move complexity.

7.1.3. Patterns on 3 vertices

In [31] we deal with graph classes characterization and recognition. A popular way to characterize a graph class is to list a minimal set of forbidden induced subgraphs. Unfortunately this strategy usually does not lead to an efficient recognition algorithm. On the other hand, many graph classes can be efficiently recognized by techniques based on some interesting orderings of the nodes, such as the ones given by traversals.

We study specifically graph classes that have an ordering avoiding some ordered structures. More precisely, we consider what we call *patterns on three nodes*, and the recognition complexity of the associated classes. In this domain, there are two key previous works. Damashke started the study of the classes defined by forbidden patterns, a set that contains interval, chordal and bipartite graphs among others. On the algorithmic side, Hell, Mohar and Rafiey proved that any class defined by a set of forbidden patterns can be recognized in polynomial time. We improve on these two works, by characterizing systematically all the classes defined sets of forbidden patterns (on three nodes), and proving that among the 23 different classes (up to complementation) that we find, 21 can actually be recognized in linear time.

Beyond this result, we consider that this type of characterization is very useful, leads to a rich structure of classes, and generates a lot of open questions worth investigating.

7.1.4. The Dependent Doors Problem: An Investigation into Sequential Decisions without Feedback

In [13], we introduce the *dependent doors problem* as an abstraction for situations in which one must perform a sequence of dependent decisions, without receiving feedback information on the effectiveness of previously made actions. Informally, the problem considers a set of d doors that are initially closed, and the aim is to open all of them as fast as possible. To open a door, the algorithm knocks on it and it might open or not according to some probability distribution. This distribution may depend on which other doors are currently open, as well as on which other doors were open during each of the previous knocks on that door. The algorithm aims to minimize the expected time until all doors open. Crucially, it must act at any time without knowing whether or which other doors have already opened. In this work, we focus on scenarios where dependencies between doors are both positively correlated and acyclic.

The fundamental distribution of a door describes the probability it opens in the best of conditions (with respect to other doors being open or closed). We show that if in two configurations of d doors corresponding doors share the same fundamental distribution, then these configurations have the same optimal running time up to a universal constant, no matter what are the dependencies between doors and what are the distributions. We also identify algorithms that are optimal up to a universal constant factor. For the case in which all doors share the same fundamental distribution we additionally provide a simpler algorithm, and a formula to calculate its running time. We furthermore analyse the price of lacking feedback for several configurations governed by standard fundamental distributions. In particular, we show that the price is logarithmic in d for memoryless doors, but can potentially grow to be linear in d for other distributions.

We then turn our attention to investigate precise bounds. Even for the case of two doors, identifying the optimal sequence is an intriguing combinatorial question. Here, we study the case of two cascading memoryless doors. That is, the first door opens on each knock independently with probability p_1. The second door can only open if the first door is open, in which case it will open on each knock independently with probability p_2. We solve this problem almost completely by identifying algorithms that are optimal up to an additive term of 1.

7.1.5. Finding maximum cliques in disk and unit ball graphs

In an *intersection graph*, the vertices are geometric objects with an edge between any pair of intersecting objects. Intersection graphs have been studied for many different families of objects due to their practical

applications and their rich structural properties. Among the most studied ones are *disk graphs*, which are intersection graphs of closed disks in the plane, and their special case, *unit disk graphs*, where all the radii are equal. Their applications range from sensor networks to map labeling, and many standard optimization problems have been studied on disk graphs. Most of the hard optimization and decision problems remain NP-hard on disk graphs and even unit disk graphs. For instance, disk graphs contain planar graphs on which several of those problems are intractable.

The complexity of MAXIMUM CLIQUE on general disk graphs is a notorious open question in computational geometry. On the one hand, no polynomial-time algorithm is known, even when the geometric representation is given. On the other hand, the NP-hardness of the problem has not been established, even when only the graph is given as input.

Recently, Bonnet *et al.* showed that the disjoint union of two odd cycles is not the complement of a disk graph. From this result, they obtained a subexponential algorithm running in time $2^{\widetilde{O}(n^{2/3})}$ for MAXIMUM CLIQUE on disk graphs, based on a win-win approach. They also got a QPTAS by calling a PTAS for MAXIMUM INDEPENDENT SET on graphs with sublinear odd cycle packing number due to Bock *et al.*, or branching on a low-degree vertex.

In [17], our main contributions are twofold. The first is a randomized EPTAS (Efficient Polynomial-Time Approximation Scheme, that is, a PTAS in time $f(\varepsilon)n^{O(1)}$) for MAXIMUM INDEPENDENT SET on graphs of $\mathcal{X}(d, \beta, 1)$. The class $\mathcal{X}(d, \beta, 1)$ denotes the class of graphs whose neighborhood hypergraph has VC-dimension at most d, independence number at least βn, and no disjoint union of two odd cycles as an induced subgraph. Using the forbidden induced subgraph result of Bonnet *et al.*, it is then easy to reduce MAXIMUM CLIQUE on disk graphs to MAXIMUM INDEPENDENT SET on $\mathcal{X}(4, \beta, 1)$ for some constant β. We therefore obtain a randomized EPTAS (and a PTAS) for MAXIMUM CLIQUE on disk graphs, settling almost [0] completely the approximability of this problem.

The second contribution is to show the same forbidden induced subgraph for unit ball graphs as the one obtained for disk graphs : their complement cannot have a disjoint union of two odd cycles as an induced subgraph. The proofs are radically different and the classes are incomparable. So the fact that the same obstruction applies for disk graphs and unit ball graphs might be somewhat accidental. And again we therefore obtain a randomized EPTAS in time $2^{\widetilde{O}(1/\varepsilon^3)}n^{O(1)}$ for MAXIMUM CLIQUE on unit ball graphs, even without the geometric representation.

Before that result, the best approximation factor was 2.553, due to Afshani and Chan. In particular, even getting a 2-approximation algorithm (as for disk graphs) was open.

Finally we show that such an approximation scheme, even in subexponential time, is unlikely for ball graphs (that is, 3-dimensional disk graphs with arbitrary radii), and unit 4-dimensional disk graphs. Our lower bounds also imply NP-hardness. To the best of our knowledge, the NP-hardness of MAXIMUM CLIQUE on unit d-dimensional disk graphs was only known when d is superconstant ($d = \Omega(\log n)$).

7.1.6. δ-hyperbolicity

In [19], we show that the eccentricities (and thus the centrality indices) of all vertices of a δ-hyperbolic graph $G = (V, E)$ can be computed in linear time with an additive one-sided error of at most $c\delta$, i.e., after a linear time preprocessing, for every vertex v of G one can compute in $O(1)$ time an estimate $\widehat{e}(v)$ of its eccentricity $ecc_G(v)$ such that $ecc_G(v) \leq \widehat{e}(v) \leq ecc_G(v) + c\delta$ for a small constant c. We prove that every δ-hyperbolic graph G has a shortest path tree, constructible in linear time, such that for every vertex v of G, $ecc_G(v) \leq ecc_T(v) \leq ecc_G(v) + c\delta$. These results are based on an interesting monotonicity property of the eccentricity function of hyperbolic graphs: the closer a vertex is to the center of G, the smaller its eccentricity is. We also show that the distance matrix of G with an additive one-sided error of at most $c'\delta$ can be computed in $O(|V|^2 \log^2 |V|)$ time, where $c' < c$ is a small constant. Recent empirical studies show that many real-world graphs (including Internet application networks, web networks, collaboration networks, social networks, biological networks, and others) have small hyperbolicity. So, we analyze the performance of our algorithms

[0] The NP-hardness, ruling out a 1-approximation, is still to show.

for approximating centrality and distance matrix on a number of real-world networks. Our experimental results show that the obtained estimates are even better than the theoretical bounds.

7.1.7. Graph searches and geometric convexities in graphs

In an attempt to understand graph searching on cocomparability graphs has been so successful, one quickly notices that the orderings produced by these traversals are precisely words of some antimatroids or convex geometries. The notion of antimatroids and convex geometries have appeared in the literature under various settings; in this work, we focus on the graph searching setting, where we discuss some known geometries on cocomparability graphs, and then present new structural properties on AT-free graphs in the hope of exploring whether the algorithms on cocomparability graphs can be lifted to this larger graph class. A first version of this work in collaboration with Feodor Dragan and Lalla Mouatadib was presented at ICGT Lyon, july 2018.

7.2. Distributed Computing

7.2.1. On the Limits of Noise in Distributed Computing

Biological systems can share and collectively process information to yield emergent effects, despite inherent noise in communication. While man-made systems often employ intricate structural solutions to overcome noise, the structure of many biological systems is more amorphous. It is not well understood how communication noise may affect the computational repertoire of such groups. To approach this question we consider in [9], [15] the basic collective task of rumor spreading, in which information from few knowledgeable sources must reliably flow into the rest of the population. We study the effect of communication noise on the ability of groups that lack stable structures to efficiently solve this task. We present an impossibility result which strongly restricts reliable rumor spreading in such groups. Namely, we prove that, in the presence of even moderate levels of noise that affect all facets of the communication, no scheme can significantly outperform the trivial one in which agents have to wait until directly interacting with the sources—a process which requires linear time in the population size. Our results imply that in order to achieve efficient rumor spread a system must exhibit either some degree of structural stability or, alternatively, some facet of the communication which is immune to noise. We then corroborate this claim by providing new analyses of experimental data regarding recruitment in Cataglyphis niger desert ants. Finally, in light of our theoretical results, we discuss strategies to overcome noise in other biological systems.

7.2.2. Minimizing message size in stochastic communication patterns: fast self-stabilizing protocols with 3 bits

In [8], we consider the basic PULL model of communication, in which in each round, each agent extracts information from few randomly chosen agents. We seek to identify the smallest amount of information revealed in each interaction (message size) that nevertheless allows for efficient and robust computations of fundamental information dissemination tasks. We focus on the *Majority Bit Dissemination* problem that considers a population of n agents, with a designated subset of *source agents*. Each source agent holds an *input bit* and each agent holds an *output bit*. The goal is to let all agents converge their output bits on the most frequent input bit of the sources (the *majority bit*). Note that the particular case of a single source agent corresponds to the classical problem of *Broadcast* (also termed *Rumor Spreading*). We concentrate on the severe fault-tolerant context of *self-stabilization*, in which a correct configuration must be reached eventually, despite all agents starting the execution with arbitrary initial states. In particular, the specification of who is a source and what is its initial input bit may be set by an adversary.

We first design a general compiler which can essentially transform any self-stabilizing algorithm with a certain property (called "the *the bitwise-independence property*") that uses ℓ-bits messages to one that uses only $\log \ell$-bits messages, while paying only a small penalty in the running time. By applying this compiler recursively we then obtain a self-stabilizing *Clock Synchronization* protocol, in which agents synchronize their clocks modulo some given integer T, within $\widetilde{O}(\log n \log T)$ rounds w.h.p., and using messages that contain 3 bits only. We then employ the new Clock Synchronization tool to obtain a self-stabilizing Majority Bit Dissemination protocol which converges in $\widetilde{O}(\log n)$ time, w.h.p., on every initial configuration, provided that the ratio of

sources supporting the minority opinion is bounded away from half. Moreover, this protocol also uses only 3 bits per interaction.

7.2.3. Intense Competition can Drive Selfish Explorers to Optimize Coverage

In [30], we consider a game-theoretic setting in which selfish individuals compete over resources of varying quality. The motivating example is a group of animals that disperse over patches of food of different abundances. In such scenarios, individuals are biased towards selecting the higher quality patches, while, at the same time, aiming to avoid costly collisions or overlaps. Our goal is to investigate the impact of collision costs on the parallel coverage of resources by the whole group.

Consider M *sites*, where a site x has *value* $f(x)$. We think of $f(x)$ as the reward associated with site x, and assume that if a single individual visits x exclusively, it receives this exact reward. Typically, we assume that if $\ell > 1$ individuals visit x then each receives at most $f(x)/\ell$. In particular, when competition costs are high, each individual might receive an amount strictly less than $f(x)/\ell$, which could even be negative. Conversely, modeling cooperation at a site, we also consider cases where each one gets more than $f(x)/\ell$. There are k identical *players* that compete over the rewards. They independently act in parallel, in a one-shot scenario, each specifying a single site to visit, without knowing which sites are explored by others. The group performance is evaluated by the expected *coverage*, defined as the sum of $f(x)$ over all sites that are explored by at least one player. Since we assume that players cannot coordinate before choosing their site we focus on symmetric strategies.

The main takeaway message of this paper is that the optimal symmetric coverage is expected to emerge when collision costs are relatively high, so that the following "Judgment of Solomon" type of rule holds: If a single player explores a site x then it gains its full reward $f(x)$, but if several players explore it, then neither one receives any reward. Under this policy, it turns out that there exists a unique symmetric Nash Equilibrium strategy, which is, in fact, evolutionary stable. Moreover, this strategy yields the best possible coverage among all symmetric strategies. Viewing the coverage measure as the social welfare, this policy thus enjoys a *(Symmetric) Price of Anarchy* of precisely 1, whereas, in fact, any other congestion policy has a price strictly greater than 1.

Our model falls within the scope of mechanism design, and more precisely in the area of incentivizing exploration. It finds relevance in evolutionary ecology, and further connects to studies on Bayesian parallel search algorithms.

7.2.4. Universal Protocols for Information Dissemination Using Emergent Signals

In [23], we consider a population of n agents which communicate with each other in a decentralized manner, through random pairwise interactions. One or more agents in the population may act as authoritative sources of information, and the objective of the remaining agents is to obtain information from or about these source agents. We study two basic tasks: broadcasting, in which the agents are to learn the bit-state of an authoritative source which is present in the population, and source detection, in which the agents are required to decide if at least one source agent is present in the population or not.

We focus on designing protocols which meet two natural conditions: (1) universality, i.e., independence of population size, and (2) rapid convergence to a correct global state after a reconfiguration, such as a change in the state of a source agent. Our main positive result is to show that both of these constraints can be met. For both the broadcasting problem and the source detection problem, we obtain solutions with a convergence time of $O(\log^2 n)$ rounds, w.h.p., from any starting configuration. The solution to broadcasting is exact, which means that all agents reach the state broadcast by the source, while the solution to source detection admits one-sided error on a ε-fraction of the population (which is unavoidable for this problem). Both protocols are easy to implement in practice and have a compact formulation.

Our protocols exploit the properties of self-organizing oscillatory dynamics. On the hardness side, our main structural insight is to prove that any protocol which meets the constraints of universality and of rapid convergence after reconfiguration must display a form of non-stationary behavior (of which oscillatory dynamics are an example). We also observe that the periodicity of the oscillatory behavior of the protocol,

when present, must necessarily depend on the number $^{\#}X$ of source agents present in the population. For instance, our protocols inherently rely on the emergence of a signal passing through the population, whose period is $\Theta(\log \frac{n}{\#X})$ rounds for most starting configurations. The design of clocks with tunable frequency may be of independent interest, notably in modeling biological networks.

7.2.5. Ergodic Effects in Token Circulation

In [25], we consider a dynamical process in a network which distributes all particles (tokens) located at a node among its neighbors, in a round-robin manner.

We show that in the recurrent state of this dynamics (i.e., disregarding a polynomially long initialization phase of the system), the number of particles located on a given edge, averaged over an interval of time, is tightly concentrated around the average particle density in the system. Formally, for a system of k particles in a graph of m edges, during any interval of length T, this time-averaged value is $k/m \pm \widetilde{O}(1/T)$, whenever $gcd(m,k) = \widetilde{O}(1)$ (and so, e.g., whenever m is a prime number). To achieve these bounds, we link the behavior of the studied dynamics to ergodic properties of traversals based on Eulerian circuits on a symmetric directed graph. These results are proved through sum set methods and are likely to be of independent interest.

As a corollary, we also obtain bounds on the *idleness* of the studied dynamics, i.e., on the longest possible time between two consecutive appearances of a token on an edge, taken over all edges. Designing trajectories for k tokens in a way which minimizes idleness is fundamental to the study of the patrolling problem in networks. Our results immediately imply a bound of $\widetilde{O}(m/k)$ on the idleness of the studied process, showing that it is a distributed $\widetilde{O}(1)$-competitive solution to the patrolling task, for all of the covered cases. Our work also provides some further insights that may be interesting in load-balancing applications.

7.2.6. Improved Analysis of Deterministic Load-Balancing Schemes

In [7], we consider the problem of deterministic load balancing of tokens in the discrete model. A set of n processors is connected into a d-regular undirected network. In every time step, each processor exchanges some of its tokens with each of its neighbors in the network. The goal is to minimize the discrepancy between the number of tokens on the most-loaded and the least-loaded processor as quickly as possible.

Rabani et al. (1998) present a general technique for the analysis of a wide class of discrete load balancing algorithms. Their approach is to characterize the deviation between the actual loads of a discrete balancing algorithm with the distribution generated by a related Markov chain. The Markov chain can also be regarded as the underlying model of a continuous diffusion algorithm. Rabani et al. showed that after time $T = O(\log(Kn)/\mu)$, any algorithm of their class achieves a discrepancy of $O(d \log n/\mu)$, where μ is the spectral gap of the transition matrix of the graph, and K is the initial load discrepancy in the system.

In this work we identify some natural additional conditions on deterministic balancing algorithms, resulting in a class of algorithms reaching a smaller discrepancy. This class contains well-known algorithms, eg., the Rotor-Router. Specifically, we introduce the notion of cumulatively fair load-balancing algorithms where in any interval of consecutive time steps, the total number of tokens sent out over an edge by a node is the same (up to constants) for all adjacent edges. We prove that algorithms which are cumulatively fair and where every node retains a sufficient part of its load in each step, achieve a discrepancy of $O(\min\{d\sqrt{\log n/\mu}, d\sqrt{n}\})$ in time $O(T)$. We also show that in general neither of these assumptions may be omitted without increasing discrepancy. We then show by a combinatorial potential reduction argument that any cumulatively fair scheme satisfying some additional assumptions achieves a discrepancy of $O(d)$ almost as quickly as the continuous diffusion process. This positive result applies to some of the simplest and most natural discrete load balancing schemes.

7.2.7. The assignment problem

In the allocation problem, asynchronous processors must partition a set of items so that each processor leave knowing all items exclusively allocated to it. In [21], we introduce a new variant of the allocation problem called the assignment problem, in which processors might leave having only partial knowledge of their assigned items. The missing items in a processor's assignment must eventually be announced by other processors.

While allocation has consensus power 2, we show that the assignment problem is solvable read-write wait-free when k processors compete for at least $2k-1$ items. Moreover, we propose a long-lived read-write wait-free assignment algorithm which is fair, allocating no more than 2 items per processor, and in which a slow processor may delay the assignment of at most n items, where n is the number of processors.

The assignment problem and its read-write solution may be of practical interest for implementing resource allocators and work queues, which are pervasive concurrent programming patterns, as well as stream-processing systems.

7.2.8. A Characterization of t-Resilient Colorless Task Anonymous Solvability

One of the central questions in distributed computability is characterizing the tasks that are solvable in a given system model. In the anonymous case, where processes have no identifiers and communicate through multi-writer/multi-reader registers, there is a recent topological characterization (Yanagisawa 2017) of the colorless tasks that are solvable when any number of asynchronous processes may crash. In [22], we consider the case where at most t asynchronous processes may crash, where $1 \leq t < n$. We prove that a colorless task is t-resilient solvable anonymously if and only if it is t-resilient solvable non-anonymously. We obtain our results through various reductions and simulations that explore how to extend techniques for non-anonymous computation to anonymous one.

7.2.9. Implementing Snapshot Objects on Top of Crash-Prone Asynchronous Message-Passing Systems

In asynchronous crash-prone read/write shared-memory systems there is the notion of a snapshot object, which simulates the behavior of an array of single-writer/multi-reader (SWMR) shared registers that can be read atomically. Processes in the system can access the object invoking (any number of times) two operations, denoted write() and snapshot(). A process invokes write() to update the value of its register in the array. When it invokes snapshot(), the process obtains the values of all registers, as if it read them simultaneously. It is known that a snapshot object can be implemented on top of SWMR registers, tolerating any number of process failures. Snapshot objects provide a level of abstraction higher than individual SWMR registers, and they simplify the design of applications. Building a snapshot object on an asynchronous crash-prone message-passing system has similar benefits. The object can be implemented by using the known simulations of a SWMR shared memory on top of an asynchronous message-passing system (if less than half the processes can crash), and then build a snapshot object on top of the simulated SWMR memory. [10] presents an algorithm that implements a snapshot object directly on top of the message-passing system, without building an intermediate layer of a SWMR shared memory. To the authors knowledge, the proposed algorithm is the first providing such a direct construction. The algorithm is more efficient than the indirect solution, yet relatively simple.

7.2.10. Distributed decision

We have carried out our study of distributed decision, either for its potential application to the design of fault-tolerant distributed algorithm, or for the purpose of designing a complexity/computanility theory for distributed network computing.

In the framework of *distributed network computing*, it is known that not all Turing-decidable predicates on labeled networks can be decided *locally* whenever the computing entities are Turing machines (TM), and this holds even if nodes are running *non-deterministic* Turing machines (NTM). In contrast, we show in [6] that every Turing-decidable predicate on labeled networks can be decided locally if nodes are running *alternating* Turing machines (ATM). More specifically, we show that, for every such predicate, there is a local algorithm for ATMs, with at most two alternations, that decides whether the actual labeled network satisfies that predicate. To this aim, we define a hierarchy of classes of decision tasks, where the lowest level contains tasks solvable with TMs, the first level those solvable with NTMs, and the level $k > 1$ contains those tasks solvable with ATMs with $k-1$ alternations. We characterize the entire hierarchy, and show that it collapses in the second level. In addition, we show separation results between the classes of network predicates that are locally decidable with TMs, NTMs, and ATMs, and we establish the existence of completeness results for

each of these classes, using novel notions of *local reduction*. We complete these results by a study of the local decision hierarchy when certificates are bounded to be of logarithmic size.

Distributed proofs are mechanisms enabling the nodes of a network to collectively and efficiently check the correctness of Boolean predicates on the structure of the network (e.g. having a specific diameter), or on data structures distributed over the nodes (e.g. a spanning tree). In [24], we consider well known mechanisms consisting of two components: a *prover* that assigns a *certificate* to each node, and a distributed algorithm called *verifier* that is in charge of verifying the distributed proof formed by the collection of all certificates. We show that many network predicates have distributed proofs offering a high level of redundancy, explicitly or implicitly. We use this remarkable property of distributed proofs to establish perfect tradeoffs between the *size of the certificate* stored at every node, and the *number of rounds* of the verification protocol.

The role of unique node identifiers in network computing is well understood as far as *symmetry breaking* is concerned. However, the unique identifiers also *leak information* about the computing environment—in particular, they provide some nodes with information related to the size of the network. It was recently proved that in the context of *local decision*, there are some decision problems that cannot be solved without unique identifiers, but unique identifiers leak a *sufficient* amount of information such that the problem becomes solvable (PODC 2013). In [11], we give a complete picture of what is the *minimal* amount of information that we need to leak from the environment to the nodes in order to solve local decision problems. Our key results are related to *scalar oracles* f that, for any given n, provide a multiset $f(n)$ of n labels; then the adversary assigns the labels to the n nodes in the network. This is a direct generalisation of the usual assumption of unique node identifiers. We give a complete characterisation of the *weakest oracle* that leaks at least as much information as the unique identifiers. Our main result is the following dichotomy: we classify scalar oracles as *large* and *small*, depending on their asymptotic behaviour, and show that (1) any large oracle is at least as powerful as the unique identifiers in the context of local decision problems, while (2) for any small oracle there are local decision problems that still benefit from unique identifiers.

7.3. Models and Algorithms for Networks

7.3.1. Revisiting Radius, Diameter, and all Eccentricity Computation in Graphs through Certificates

In [28], we introduce notions of certificates allowing to bound eccentricities in a graph. In particular , we revisit radius (minimum eccentricity) and diameter (maximum eccentricity) computation and explain the efficiency of practical radius and diameter algorithms by the existence of small certificates for radius and diameter plus few additional properties. We show how such computation is related to covering a graph with certain balls or complementary of balls. We introduce several new algorithmic techniques related to eccentricity computation and propose algorithms for radius, diameter and all eccentricities with theoretical guarantees with respect to certain graph parameters. This is complemented by experimental results on various real-world graphs showing that these parameters appear to be low in practice. We also obtain refined results in the case where the input graph has low doubling dimension, has low hyperbolicity, or is chordal.

7.3.2. Efficient Loop Detection in Forwarding Networks and Representing Atoms in a Field of Sets

In [29], we consider the problem of detecting loops in a forwarding network which is known to be NP-complete when general rules such as wildcard expressions are used. Yet, network analyzer tools such as Netplumber (Kazemian et al., NSDI'13) or Veriflow (Khurshid et al., NSDI'13) efficiently solve this problem in networks with thousands of forwarding rules. In this paper, we complement such experimental validation of practical heuristics with the first provably efficient algorithm in the context of general rules. Our main tool is a canonical representation of the atoms (i.e. the minimal non-empty sets) of the field of sets generated by a collection of sets. This tool is particularly suited when the intersection of two sets can be efficiently computed and represented. In the case of forwarding networks, each forwarding rule is associated with the set of packet headers it matches. The atoms then correspond to classes of headers with same behavior in the network. We

propose an algorithm for atom computation and provide the first polynomial time algorithm for loop detection in terms of number of classes (which can be exponential in general). This contrasts with previous methods that can be exponential, even in simple cases with linear number of classes. Second, we introduce a notion of network dimension captured by the overlapping degree of forwarding rules. The values of this measure appear to be very low in practice and constant overlapping degree ensures polynomial number of header classes. Forwarding loop detection is thus polynomial in forwarding networks with constant overlapping degree.

7.3.3. Exact Distance Oracles Using Hopsets

In [33], we consider for fixed $h \geq 2$ the task of adding to a graph G a set of weighted shortcut edges on the same vertex set, such that the length of a shortest h-hop path between any pair of vertices in the augmented graph is exactly the same as the original distance between these vertices in G. A set of shortcut edges with this property is called an exact h-hopset and may be applied in processing distance queries on graph G. In particular, a 2-hopset directly corresponds to a distributed distance oracle known as a hub labeling. In this work, we explore centralized distance oracles based on 3-hopsets and display their advantages in several practical scenarios. In particular, for graphs of constant highway dimension, and more generally for graphs of constant skeleton dimension, we show that 3-hopsets require exponentially fewer shortcuts per node than any previously described distance oracle while incurring only a quadratic increase in the query decoding time, and actually offer a speedup when compared to simple oracles based on a direct application of 2-hopsets. Finally, we consider the problem of computing minimum-size h-hopset (for any $h \geq 2$) for a given graph G, showing a polylogarithmic-factor approximation for the case of unique shortest path graphs. When $h = 3$, for a given bound on the space used by the distance oracle, we provide a construction of hopsets achieving polylog approximation both for space and query time compared to the optimal 3-hopset oracle given the space bound.

7.3.4. Game Theory in Networks

Two notable contributions to game theory applied to networks are worth being mentioned.

In [14], we show that the Preferential Attachment rule naturally emerges in the context of evolutionary network formation, as the *unique* Nash equilibrium of a simple social network game. To demonstrate this result, we start from the fact that each node of a social network aims at maximizing its degree in the future, as this degree is representing its social capital in the "society" formed by the nodes and their connections. We show that, to maximize the node degree in the future, the unique Nash equilibrium consists in playing the Preferential Attachment rule when each node connects to the network. This result provides additional formal support to the commonly used Preferential Attachment model, initially designed to capture the "rich get richer" aphorism. In the process of establishing our result, we expose new connections between Preferential Attachment, random walks, and Young's Lattice.

In [20], we notice that distributed tasks such as constructing a maximal independent set (MIS) in a network, or properly coloring the nodes or the edges of a network with reasonably few colors, are known to admit efficient distributed randomized algorithms. Those algorithms essentially proceed according to some simple generic rules, by letting each node choosing a temptative value at random, and checking whether this choice is consistent with the choices of the nodes in its vicinity. If this is the case, then the node outputs the chosen value, else it repeats the same process. However, although such algorithms are, with high probability, running in a polylogarithmic number of rounds, they are *not robust* against actions performed by rational but selfish nodes. Indeed, such nodes may prefer specific individual outputs over others, e.g., because the formers suit better with some individual constraints. For instance, a node may prefer not being placed in a MIS as it is not willing to serve as a relay node. Similarly, a node may prefer not being assigned some radio frequencies (i.e., colors) as these frequencies would interfere with other devices running at that node. We show that the probability distribution governing the choices of the output values in the generic algorithm can be tuned such that no nodes will rationally deviate from this distribution. More formally, and more generally, we prove that the large class of so-called LCL tasks, including MIS and coloring, admit simple "Luby's style" algorithms where the probability distribution governing the individual choices of the output values forms a Nash equilibrium. In fact, we establish the existence of a stronger form of equilibria, called symmetric trembling-hand perfect equilibria for those games.

MIMOVE Project-Team

7. New Results

7.1. Ontology categorization for IoT semantics

Participants: Rachit Agarwal, Nikolaos Georgantas, Valérie Issarny.

IoT systems are now being deployed worldwide to sense phenomena of interest. The existing IoT systems are often independent which limits the use of sensor data to only one application. Semantic solutions have been proposed to support reuse of sensor data across IoT systems and applications. This allows integration of IoT systems for increased productivity by solving challenges associated with their interoperability and heterogeneity. Several ontologies have been proposed to handle different aspects of sensor data collection in IoT systems, ranging from sensor discovery to applying reasoning on collected sensor data for drawing inferences. In this work, we study and categorise the existing ontologies based on the fundamental ontological concepts (e.g., sensors, context, location, and more) required for annotating different aspects of data collection and data access in an IoT application. We identify these fundamental concepts by answering the 4Ws (What, When, Who, Where) and 1H (How) identified using the 4W1H methodology.

7.2. Massively-Parallel Feature Selection for Big Data

Participant: Vassilis Christophides.

We present the Parallel, Forward-Backward with Pruning (PFBP) algorithm for feature selection (FS) in Big Data settings (high dimensionality and/or sample size). To tackle the challenges of Big Data FS, PFBP partitions the data matrix both in terms of rows (samples, training examples) as well as columns (features). By employing the concepts of p-values of conditional independence tests and meta-analysis techniques, PFBP manages to rely only on computations local to a partition while minimizing communication costs. Then, it employs powerful and safe (asymptotically sound) heuristics to make early, approximate decisions, such as Early Dropping of features from consideration in subsequent iterations, Early Stopping of consideration of features within the same iteration, or Early Return of the winner in each iteration. PFBP provides asymptotic guarantees of optimality for data distributions faithfully representable by a causal network (Bayesian network or maximal ancestral graph). Our empirical analysis confirms a superlinear speedup of the algorithm with increasing sample size, linear scalability with respect to the number of features and processing cores, while dominating other competitive algorithms in its class.

7.3. Universal Social Network Bus

Participants: Ehsan Ahvar, Shohreh Ahvar, Rafael Angarita, Nikolaos Georgantas, Valérie Issarny, Bruno Lefèvre.

Online social network services (OSNSs) are changing the fabric of our society, impacting almost every aspect of it. Over the last decades, the aggressive market rivalry has led to the emergence of multiple competing, "closed" OSNSs. As a result, users are trapped in the walled gardens of their OSNS, encountering restrictions about what they can do with their personal data, the people they can interact with and the information they get access to. As an alternative to the platform lock-in, "open" OSNSs promote the adoption of open, standardized APIs. However, users still massively adopt closed OSNSs to benefit from the services' advanced functionalities and/or follow their "friends", although the users' virtual social sphere is ultimately limited by the OSNSs they join. Our work aims at overcoming such a limitation by enabling users to meet and interact beyond the boundary of their OSNSs, including reaching out to "friends" of distinct closed OSNSs. We specifically introduce USNB -*Universal Social Network Bus*, which revisits the "service bus" paradigm that enables interoperability across computing systems, to address the requirements of "*social interoperability*". USNB features *synthetic profiles* and *personae* for interaction across the boundaries of –closed and open–, –profile- and non-profile-based– OSNSs through a *reference social interaction service*.

USNB enables users to reach out to their social peers independently of the communication service (and especially underlying platform) each one uses in the virtual world. The success and massive adoption of OSNSs -as magnified by the success of Facebook- shows that online social communication is an essential tool for people. This further paves the way for collective and collaborative actions at the Internet scale. However, existing online collaborative tools come along with their communication platform, which is either a proprietary solution or a third-party OSNS. We argue that USNB contributes to enabling participatory systems at a larger inclusive scale by overcoming the technical boundaries set by existing online communication platforms. In that direction, we investigate the customization of USNB for specific applications and more specifically: participatory systems and massive open online courses.

7.4. Middleware for Mobile Crowdsensing

Participants: Yifan Du, Valérie Issarny, Bruno Lefèvre, Françoise Sailhan.

Mobile Phone Sensing (MPS) offers a great opportunity toward the large scale monitoring of urban phenomena, such as the exposition of the population to environmental pollution. Indeed, mobile crowdsensing empowers ordinary citizens to contribute (whether pro-actively or passively) data sensed or generated from their mobile devices. It allows acquiring hyperlocal knowledge at scale, thanks to the proliferation of mobile devices and the ubiquity of wireless broadband connection. On-demand mobile crowdsensing is in particular a cost-effective service model for smart cities. Numerous sensor types embedded in today's smartphones contribute valuable quantitative observations about the urban environment (e.g., noise, temperature, atmospheric pressure, humidity, light, magnetism). The observations further come along with the related spatial and temporal data, which allows for the analysis of hyper-local environmental knowledge. However, mobile crowdsensing brings valuable knowledge only if a sufficiently large crowd contributes and if we overcome the relatively low accuracy of the gathered data. This is the focus of our research.

We have in particular studied how to reduce the gap between the need for the massive collection of relevant data, and the quantity and accuracy of the measurements that are actually gathered. We specifically carried out an iterative research process to tackle this challenge, which combines technological innovation and social design. We have been developing a number of social tools to study the motivations and usages of MPS-based smart city apps, with the Ambiciti app serving as our use case. Our study has been taking into account the cultural and societal contexts that the usages of Ambiciti could feed, spanning health, environment, education, and urban policies. We carried out an online survey together with interviews with users and local actors in Europe, i.e., France, Belgium, and Finland. The research results contribute to a better understanding of why and how people use mobile phone sensing applications; the results also inform how to best leverage mobile crowd-sensing in the development of smart cities and how it may serve addressing urban challenges related to, e.g., public health or urban planning.

The quality of the contributed measurements challenges the aggregation of relevant knowledge from crowdsensed observations. The measurements quality depends on the *accuracy* of the contributing sensors and the adequacy of the *sensing context*. Addressing the former relies on the sensor calibration for which we study both micro- and macro-level solutions. Addressing the latter requires a supporting inference mechanism, for which we introduce a *personalized hierarchical inference* of all the context elements that are relevant to the phenomenon that is monitored through crowdsensing, and under which the crowdsensor operates. This enables accounting for the specific behavior of the contributing end-user across time, as well as for all the features -and only those- that are relevant and locally available, while reducing the feedback required from the user for the personalization.

7.5. QoS-Aware Resource Allocation for Mobile IoT Pub/Sub Systems

Participants: Georgios Bouloukakis, Nikolaos Georgantas.

IoT applications are usually characterized by large-scale demand and the widespread use of mobile devices. Similarly, performing interaction among application and system components in a decoupled and elastic way, and enforcing Quality of Service (QoS) usually also become issues. Hence, paradigms such as pub/sub on top of cloud resources represent a suitable strategy for application development. However, management of QoS-aware resource allocation for pub/sub systems remains challenging, especially when system peers connect in an intermittent way. In this work, we propose a new approach for resource allocation focusing on end-to-end performance in face of peers' disconnections. We evaluate and demonstrate the benefits of our approach using simulations. QoS enforcement was achieved in almost all scenarios, and we have shown that our approach can help reasoning about efficient resource allocation.

7.6. Queueing Network Modeling Patterns for Reliable & Unreliable Pub/Sub Protocols

Participants: Georgios Bouloukakis, Nikolaos Georgantas, Patient Ntumba, Valérie Issarny.

Mobile Internet of Things (IoT) applications are typically deployed on resource-constrained devices with intermittent network connectivity. To support the deployment of such applications, the Publish/Subscribe (pub/sub) interaction paradigm is often employed, as it decouples mobile peers in time and space. Furthermore, pub/sub middleware protocols and APIs consider the Things' hardware limitations and support the development of effective applications by providing Quality of Service (QoS) features. These features aim to enable developers to tune an application by switching different levels of response times and delivery success rates. However, the profusion of pub/sub middleware protocols coupled with intermittent network connectivity result in non-trivial application tuning. In this work, we model the performance of middleware protocols found in IoT, which are classified within the pub/sub interaction paradigm – both reliable and unreliable underlying network layers are considered. We model reliable and unreliable protocols, by considering QoS semantics for data validity, buffer capacities, as well as the intermittent availability of peers. To this end, we rely on queueing network models, which offer a simple modeling environment that can be used to represent IoT interactions by combining multiple queueing model types. Based on these models, we perform statistical analysis by varying the QoS semantics, demonstrating their significant effect on response times and on the rate of successful interactions. We showcase the application of our analysis in concrete scenarios relating to Traffic Information Management systems, that integrate both reliable and unreliable participants. The consequent PerfMP performance modeling pattern may be tailored for a variety of deployments, in order to control fine-grained QoS policies.

7.7. Lightweight, General Inference of Streaming Video Quality from Encrypted Traffic

Participants: Francesco Bronzino, Sara Ayoubi, Renata Teixeira, Sarah Wasserman.

Accurately monitoring application performance is becoming more important for Internet Service Providers (ISPs), as users increasingly expect their networks to consistently deliver acceptable application quality. At the same time, the rise of end-to-end encryption makes it difficult for network operators to determine video stream quality–including metrics such as startup delay, resolution, rebuffering, and resolution changes–directly from the traffic stream. This work develops general methods to infer streaming video quality metrics from encrypted traffic using lightweight features. Our evaluation shows that our models are not only as accurate as previous approaches, but they also generalize across multiple popular video services, including Netflix, YouTube, Amazon Instant Video, and Twitch. The ability of our models to rely on lightweight features points to promising future possibilities for implementing such models at a variety of network locations along the end-to-end network path, from the edge to the core.

7.8. Service traceroute: Tracing Paths of Application Flows

Participants: Ivan Morandi, Francesco Bronzino, Renata Teixeira.

Traceroute is often used to help diagnose when users experience issues with Internet applications or services. Unfortunately, probes issued by classic traceroute tools differ from application traffic and hence can be treated differently by middleboxes within the network. This work proposes a new traceroute tool, called Service traceroute. Service traceroute leverages the idea from paratrace, which passively listens to application traffic to then issue traceroute probes that pretend to be part of the application flow. We extend this idea to work for modern Internet services with support for identifying the flows to probe automatically, for tracing of multiple concurrent flows, and for UDP flows. We implement command-line and library versions of Service traceroute, which we release as open source. This paper also presents an evaluation of Service traceroute when tracing paths traversed by Web downloads from the top-1000 Alexa websites and by video sessions from Twitch and Youtube. Our evaluation shows that Service traceroute has no negative effect on application flows. Our comparison with Paris traceroute shows that a typical traceroute tool that launches a new flow to the same destination discovers different paths than when embedding probes in the application flow in a significant fraction of experiments (from 40% to 50% of our experiments in PlanetLab Europe).

WHISPER Project-Team

7. New Results

7.1. Software engineering for infrastructure software

The most visible tool developed in the Whisper team is Coccinelle, which this year marked the 10th anniversary of its release in open source. The paper "Coccinelle: 10 Years of Automated Evolution in the Linux Kernel," published at USENIX ATC'18 [14], traced the history of Coccinelle, its underlying design decisions and impact. The Coccinelle C-code matching and transformation tool was first released in 2008 to facilitate specification and automation in the evolution of Linux kernel code. The novel contribution of Coccinelle was to allow software developers to write code manipulation rules in terms of the code structure itself, via a generalization of the patch syntax. Over the years, Coccinelle has been extensively used in Linux kernel development, resulting in over 6000 commits to the Linux kernel, and has found its place as part of the Linux kernel development process. The USENIX ATC paper studies the impact of Coccinelle on Linux kernel development and the features of Coccinelle that have made it possible. It provides guidance on how other research-based tools can achieve practical impact in the open-source development community. This work was also presented to Linux kernel developers at Kernel Recipes and Open Source Summit Europe, and at the 8th Inria/Technicolor Workshop On Systems.

In a modern OS, kernel modules often use spinlocks and interrupt handlers to monopolize a CPU core to execute concurrent code in atomic context. In this situation, if the kernel module performs an operation that can sleep at runtime, a system hang may occur. We refer to this kind of concurrency bug as a sleep-in-atomic-context (SAC) bug. In practice, SAC bugs have received insufficient attention and are hard to find, as they do not always cause problems in real executions. In a paper published at USENIX ATC'18 [12], we propose a practical static approach named DSAC, to effectively detect SAC bugs and automatically recommend patches to help fix them. DSAC uses four key techniques: (1) a hybrid of flow-sensitive and -insensitive analysis to perform accurate and efficient code analysis; (2) a heuristics-based method to accurately extract kernel interfaces that can sleep at runtime; (3) a path-check method to effectively filter out repeated reports and false bugs; (4) a pattern-based method to automatically generate recommended patches to help fix the bugs. We evaluate DSAC on kernel modules (drivers, file systems, and network modules) of the Linux kernel, and on the FreeBSD and NetBSD kernels, and in total find 401 new real bugs. 272 of these bugs have been confirmed by the relevant kernel maintainers, and 43 patches generated by DSAC have been applied by kernel maintainers.

7.2. Trustworthy domain-specific compilers

To achieve safety and composability, we believe that an 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 [10]. It combines ideas from gradual typing – through which types are compiled into run-time assertions – and the theory of ornaments [31] – 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 the result of a collaboration with Eric Tanter, from the University of Chile, and Nicolas Tabareau, from the Gallinette Inria project-team.

7.3. High-performance domain-specific compilers

As part of Darius Mercadier's PhD project, we are developing a synchronous dataflow language targeting high-performance (and, eventually, verified) implementations of bitsliced algorithms, with application to cryptographical algorithms [33]. Using our Usuba language, cryptographers can specify a block cipher at a very high level as a set of dataflow equations. From such a description, our `usubac` compiler is able to generate efficient, vectorized code exploiting the SIMD instruction sets of the underlying architecture. We have demonstrated that our generated code performs on par with hand-tuned assembly programs while, at the same time, being able to target multiple CPU architectures as well as multiple generations of SIMD instruction sets on each architecture. This project illustrates perfectly our methodology: the design of Usuba is driven by semantic considerations (bitslicing is only meaningful for bit parallel operations) that are then structured using types and subsequently reifed into syntactic artefacts. Our preliminary results [15], published in an international workshop, are encouraging.

7.4. Multicore schedulers

As a side-effect of our work on verification of schedulers [48], we have contributed to an analysis of the impact on application performance of the design and implementation choices made in two widely used open-source schedulers: ULE, the default FreeBSD scheduler, and CFS, the default Linux scheduler. In a paper published at USENIX ATC'18 [13], we compare ULE and CFS in otherwise identical circumstances. This work involves porting ULE to Linux, and using it to schedule all threads that are normally scheduled by CFS. We compare the performance of a large suite of applications on the modified kernel running ULE and on the standard Linux kernel running CFS. The observed performance differences are solely the result of scheduling decisions, and do not reflect differences in other subsystems between FreeBSD and Linux. We found that there is no overall winner. On many workloads the two schedulers perform similarly, but for some workloads there are significant and even surprising differences. ULE may cause starvation, even when executing a single application with identical threads, but this starvation may actually lead to better application performance for some workloads. The more complex load balancing mechanism of CFS reacts more quickly to workload changes, but ULE achieves better load balance in the long run.

ALMAnaCH Team

6. New Results

6.1. Syntax modelling and treebank development

Participants: Djamé Seddah, Benoît Sagot, Éric Villemonte de La Clergerie, Emilia Verzeni, Wigdan Abbas Mekki Medeni, Elias Benaïssa, Farah Essaidi, Amal Fethi.

- In 2018, members of ALMAnaCH have finalised a conversion of the biggest annotated data set for French, the French Treebank, to Universal Dependencies 2.3, the now *de facto* standard for syntactic annotations [27]. The same group was also deeply involved in a proposal co-written with others leaders of the field [25], aiming at representing morpho-syntactic ambiguities from user-generated content and morphologically-rich languages. This proposal was implemented via the development of language specific analysers and data-driven normalised lexica [26].
- As part of the ANR Parsiti project, the development of gold standards for North-African dialectal Arabic has seen great progresses and is coming to a pre-release date in the first semester of 2019. This work involved more than 24 man.months over the last 12 months and will culminate with a multi-layered corpus of about 2000 sentences that is made of user-generated content with a highly variable dialect that contains up to 36% of French words and mixed syntax with Arabic. In order to assess the quality of the translation produced by the Parsiti project, we also included a translation layer (North-African Arabic-French) as well as all expected morpho-syntactic and syntactic annotations, following the state-of-the-art in terms of annotations. Papers are currently being written and will target the main NLP conferences of early 2019.
- In parallel to the last item, we also translated to English half of the French Social Media Bank which was developed in our previous project [92]. A morpho-syntactic annotation layer was added. The crucial difficulty was to maintain a symmetry in term of style and level of languages between French user-generated content and its English counterpart. This data set is currently being used in the Parsiti project in order to evaluate the MT models currently being developed by the LIMSI partner.

6.2. Modeling of language variability via diachronic embeddings and extra-linguistic contextual features

Participants: Djamé Seddah, Benjamin Muller, Ganesh Jawahar, Benoît Sagot, Éric Villemonte de La Clergerie.

Following ALMAnaCH's participation in the 2017 CoNLL shared task on heavily multilingual dependency parsing in the *Universal Dependency* (hereafter UD) framework (we ranked 3rd/33 on part-of-speech tagging and 6th/33 on parsing), the team has taken part in the 2018 edition of the shared task. This year, most of the work was carried out by junior members of the team, for whom it was an interesting opportunity to gain experience on the development of NLP architectures and their deployment in the context of a shared task. It was also the opportunities to test new ideas.

We developed a neural dependency parser and a neural part-of-speech tagger, which we called 'ELMoLex' [21]. We augmented the deep Biaffine (BiAF) parser [64] with novel features to perform competitively: we utilize an in-domain version of ELMo features [77], which provide context-dependent word representations; we utilised disambiguated, embedded, morphosyntactic features extracted from our UD-compatible lexicons [26], which complements the existing feature set. In addition to incorporating character embeddings, ELMoLex leverages pre-trained word vectors, ELMo and morphosyntactic features (whenever available) to correctly handle rare or unknown words which are prevalent in languages with complex morphology. ELMoLex ranked 11th in terms of the Labeled Attachment Score metrics (70.64%) and the Morphology-aware LAS metrics (55.74%), and ranked 9th in terms of Bilexical dependency metric (60.70%). In an extrinsic evaluation setup, ELMoLex ranked 7th for Event Extraction, Negation Resolution tasks and 11th for Opinion Analysis task in terms of F1 score.

6.3. Modelling of language variability via diachronic embeddings and extra-linguistic contextual features

Participants: Djamé Seddah, Ganesh Jawahar, Éric Villemonte de La Clergerie, Benoît Sagot.

As part of the ANR SoSweet and the PHC Maimonide projects (in collaboration with Bar Ilan University for the latter), ALMAnaCH has invested a lot of efforts in 2018 into studying language variability (i.e. how the language evolve over time and how this evolution is tied to socio-demographic and dynamic network variables). Taking advantages of the SoSweet corpus (220 millions tweet) and of the Bar Ilan Hebrew Tweets (180M tweets) both collected over the last 5 years, we have been addressing the problem of studying semantic changes. We devised a novel attentional model, based on Bernouilli word embeddings, that are conditioned on contextual extra-linguistic (social) features such as network, spatial and socio-economic variables, which are associated with Twitter users, as well as topic-based features. We posit that these social features provide an inductive bias that is susceptible to helping our model to overcome the narrow time-span regime problem. Our extensive experiments reveal that, as a result of being less biased towards frequency cues, our proposed model was able to capture subtle semantic shifts and therefore benefits from the inclusion of a reduced set of contextual features. Our model thus fit the data better than current state-of-the-art dynamic word embedding models and therefore is a promising tool to study diachronic semantic changes over small time periods. A paper on this work is currently under review.

6.4. Standardisation of Natural Language data

Participants: Laurent Romary, Jack Bowers, Charles Riondet, Mohamed Khemakhem, Benoît Sagot, Loïc Grobol.

One essential aspect of working with human traces as they occur in digital humanities at large and in natural language processing in particular, is to be able to re-use any kind of primary content and further enrichments thereof. The central aspect of re-using such content is the development and applications of reference standards that reflect the best state of the art in the corresponding domains. In this respect, our team is particularly attentive to the existing standardisation background when both producing language resources or devloping NLP components. Furthermore, our specific leading roles in the domain of standardisation in both the Parthenos and EHRI EU projects as well as in related initiatives (TEI consortium, ISO committee TC 37, DARIAH lexical working group) has allowed to make progress along the following lines:

- Contributing to the revision of the ISO 24613 standard (Lexical Markup Framework) in the form of a multipart standard covering, for the time being, the core model (ISO 24613-1), machine readable dictionaries (ISO 24613-2), etymology (ISO 24613-3) and a TEI based serialisation (ISO 24613-4). Several members of the team have been particularly active as experts in the definition of the first two parts, which are now at publication and DIS stage respectively [0] and are co-editors of parts 3 and 4;
- Proposal for a reference TEI subset for integrating dictionary content: in the context of the DARIAH working group on lexical resources, a first release of the *TEI Lex 0*[0] was issued in September 2018 integrating the continuous work of the group over the the 2016-2018 period and already taken up by the infrastructure project ELEXIS [0] as its reference back-office format. This work is also the basis for the output format of Grobid-Dictionaries [71];
- Finalisation of the ISO proposal on reference annotation (ISO 24617-9): the team has been leading the work on the definition of the Reference Annotation Framework (RAF) [0] which is now at DIS ballot stage and already implemented in several concrete annotation projects[19], [43]. The standard is feature complete from a linguistic point of view (from simple co-reference to complex bridging anaphora phenomena) and compliant with the TEI stand-off annotation module [59] from the point of view of its implementation [66];

[0] See the ISO/TC 37/SC 4 work current work program under https://www.iso.org/committee/297592/x/catalogue/p/0/u/1/w/0/d/0
[0] https://github.com/DARIAH-ERIC/lexicalresources
[0] https://elex.is
[0] https://www.iso.org/standard/69658.html

- Large-scale implementation of international standard for the documentation of the Mixtepec-Mixtec language (see section 6.11);
- Proposing a customisation architecture for the EAD international standard: EAD (Encoding Archival Description [0]) is used worldwide in cultural heritage institution to describe and exchange collection level information. In the context of the EHRI project, where we had to design a mechanism for integrating heterogeneous implementations of EAD-based data, we used the TEI ODD specification language to re-design and subset the international EAD specification to precisely provide interoperability conditions within the project[14];
- Release of the SSK (Standardisation Survival Kit), a generic environment for describing standards-based digital humanities research scenarios: the SSK is an online platform for describing research scenarios developed within the Parthenos project[40] and now deployed as a service hosted by the French national Huma-Num infrastructure [0]. The SSK has been developed as a completely open project [0], where the scenarios are themselves described as TEI-based representations[51], [35], [50].

6.5. Entity-fishing: a generic named entity recognition and disambiguation for digital humanities projects

Participants: Marie Puren, Charles Riondet, Laurent Romary, Luca Foppiano, Tanti Kristanti.

Since several years (starting at the beginning of the EU Cendari project in 2012 [75]) we have been working on the provision of a generic named-entity recognition and disambiguation module (NERD) called *entity-fishing*[18] as a stable on-line service. The work we have achieved demonstrates the possible delivery of sustainable technical services as part of the development of research infrastructures for the humanities in Europe. In particular, our results contribute not only to DARIAH, the European digital research infrastructure for the arts and humanities, but also to OPERAS, the European research infrastructure for the development of open scholarly communication in the social sciences and humanities. Deployed as part of the French national infrastructure Huma-Num, the service provides an efficient state-of-the-art implementation coupled with standardised interfaces allowing easy deployment in a variety of potential digital humanities contexts. In 2018, we have specifically integrated *entity-fishing* within the H2020 HIRMEOS project where several open access publishers have used the service in their collections of published monographs as a means to enhance retrieval and access.

To this end, we have set up a common layer of services on top of several existing e-publishing platforms for Open Access monographs. The *entity extraction* task was deployed over a corpus of monographs provided by the HIRMEOS partners, with the following coverage:

- 4000 books in English and French from Open Edition Books
- 2000 titles in English and German from OAPEN
- 162 books in English from Ubiquity Press
- 765 books (606 in German, 159 in English) from the University of Göttingen

The introduction of *entity-fishing* has undergone different levels of integration. The majority of the participating publishers provided additional features in their user interface, using the data generated by *entity-fishing*, for example, as search facets for persons and locations to help users narrow down their searches and obtain more precise results.

entity-fishing has been developed in Java and it has been designed for fast processing on text and PDF, with relatively limited memory and to offer relatively close to state-of-the-art accuracy (as compared with other NERD systems). The accuracy f-score for disambiguation is currently between 76.5 and 89.1 on standard datasets (ACE2004, AIDA-CONLL-testb, AQUAINT, MSNBC) (Table 1) [74].

[0] https://en.wikipedia.org/wiki/Encoded_Archival_Description
[0] http://ssk.huma-num.fr
[0] https://github.com/ParthenosWP4/SSK

Table 1. Accuracy measures

	ACE 2004	AIDA CONLL-testb	AQUAINT	MSNBC
Priors	83.1	66.1	80.3	71.1
entity-fishing	83.5	76.5	**89.1**	86.7
Wikifier	83.4	77.7	86.2	85.1
DoSeR	**90.7**	78.4	84.2	91.1
AIDA	81.5	77.4	53.2	78.2
Spotlight	71.3	59.3	71.3	51.1
Babelfy	56.1	59.2	65.2	60.7
WAT	80.0	84.3	76.8	77.7
(Ganea & Hofmann, 2017)	88.5	**92.2**	88.5	**93.7**

The objective, however, is to provide a generic service that has a steady throughput of 500-1000 words per second or one PDF page of a scientific article in 1-2 seconds on a medium range (4CPU, 3Gb Ram) Linux server.

From the point of view of the technical deployment itself, we have provided all the necessary components of a sustainable service:

- release and publish *entity-fishing* as open source software [0];
- deploy the service in the DARIAH infrastructure through HUMA-NUM [0];
- produce evaluation data and metrics for content validation.

6.6. From GROBID to GROBID-Dictionaries

Participants: Luca Foppiano, Mohamed Khemakhem, Laurent Romary, Pedro Ortiz Suárez, Alba Marina Malaga Sabogal.

GROBID is an open source software suite initiated in 2007 by Patrice Lopez with the purpose of extracting metadata automatically from scholarly papers available in PDF. Over the years, it has developed into a rich information extraction environment, and deployed in many Inria projects, but also national and international services, such as HAL (front-end meta-data extraction from uploaded scholarly publications). It is a central piece for our information extraction activities and we have been particularly active in 2018 in the following domains:

- General contributions to GROBID [0]:
  - Major refactoring and design improvements
  - fixes, tests, documentation and update of the pdf2xml fork for Windows
  - added and improved several models in collaboration with CERN (e.g. for the recognition of arXiv identifier)
  - Further tests on the specific case of bibliographic documents[32]
- Contribution to GROBID-Dictionaries [0]: the lexical GROBID extension has been implemented and tested on modern and multilingual dictionaries[23]. In the context of several collaborative activities, GROBID-Dictionaries has been applied on several documentary sources:
  - Early editions of the The Petit Larrousse Illustré in the context of the Nénufar project[45], [29]

[0] http://github.com/kermitt2/nerd
[0] http://nerd.huma-num.fr/nerd/
[0] https://github.com/kermitt2/grobid
[0] https://github.com/MedKhem/grobid-dictionaries

- Further experiments on etymological dictionaries from the Berlin Brandenburg Academy of Sciences
- Experiments on entry-based documents such as manuscript catalogues (with University of Neuchâtel)[16] and the French address Directory Bottin from the end of the XIXth Century[22]

These various experiments have been accompanied by an intense training and hand-on activity in the context in particular of the French research network CAHIERS (Huma-Num consortium), the Lexical Data Master Class and a series of workshop organised in South Africa under the auspices of a national linguistic documentation program. Finally, further alignments with the ongoing standardisation activities around TEI Lex0 and ISO 24613 (LMF) has been carried out to ensure a proper standards compliance of the generated output

The experience gained in the develoment and application of GROBID-Dictionaries has been the basis for the recently accepted ANR BASNUM project which aims at automatically structuring and enriching of the Dictionnaire universel (DU) by Antoine Furetière, in its 1701 edition rewritten by Basnage de Beauval and the doctoral work of Pedro Ortiz.

6.7. Resources, models and tools for coreference resolution

Participants: Loïc Grobol, Éric Villemonte de La Clergerie.

This year we performed many experiments, some of them detailed in [28], targeting end-to-end coreference systems for spontaneous oral French. More precisely, for several mention-pair coreference detection models, we tried to assess their sensibility to various features of coreference chains and their viability for end-to-end systems, compared to the more recent antecedent scoring models.

Also, one of our objective being to assess the usefulness of syntactic features for coreference detection, we enriched the coreference annotations of the ANCOR corpus with both automatically produced dependency syntax annotations and improved speech transcription. All these annotation were wrapped in a TEI-compliant XML format as described in [20] (see also 6.4).

Finally, we have been working on neural architectures for coreference detection, building upon some recent state of the art techniques. They are based on embeddings for general text span and we try to make them more scalable through efficient uses of the local context but also more tunable to different document types and language variation. The base idea is to complete pre-training by training on related lower-level tasks such as entity-mention detection.

6.8. Computational history through information extraction from archive texts

Participants: Éric Villemonte de La Clergerie, Marie Puren, Charles Riondet, Alix Chagué, Marie-Laurence Bonhomme.

From two different DH projects emerged some interesting research questions related to the extraction of information from archival documents, in particular the management of the diversity of document types and structures and on the contrary the acquisition of detailed information from a regular visual structure.

In the context of the ANR TIME-US, whose goal is to reconstruct the "time-budgets" of textile workers in France (18th - early 20th centuries), we worked on the creation of a digitization workflow to acquire structured textual data from a wide range of printed and handwritten materials: professional court records (like *Prud'Hommes*), Police reports on strikes or early sociological studies such as the *Monographies de Le Play* This workflow has been presented at the ADHO DH conference in Mexico (see the presentation here: [34].The set up of this workflow is a prerequisite for further experiments and processing to extract information that can be exploited by historians, such as the relation between working tasks, the time spent by workers to perform them and the price they are paid for this time.

Another project was intiated in collaboration with the EPHE and the French National Archives, in the framework of the convention signed between Inria and the Ministry of Culture. This project is called LECTAUREP (for *LECTure AUtomatique de REPertoires*, and is aimed at extracting the information recorded in the registries of Parisian notaries, held by the National Archives. This project it at the intersection of NLP and Computer Vision because one of the main objectives is to extract information from the physical layout of the documents, presented as tables. Another issue is to be able to recognize with accuracy an important diversity of handwritten scripts. The final goal of LECTAUREP is to give access to researchers the information contained in these records, in particular the name of the persons involved in cases recorded by notaries, their addresses and the nature of the case (wills, powers of attorney, wedding contracts, etc.). An initial report has been produced (see [39]), and the project will continue in 2019 with the release of the extracted information (named entities, geolocation, typology, etc) into a structured database.

6.9. Discovering correlations between parser features and neurological observations

Participants: Éric Villemonte de La Clergerie, Murielle Fabre, Pauline Brunet.

In the context of the CRCNS international network, the ANR-NSF NCM-ML project (dubbed "*Petit Prince* project") aims to discover and explore correlations between features (or predictors) provided by NLP tools such as parsers, and fMRI data resulting from listening of the novel Le Petit Prince.

In 2018, Pauline Brunet, during her Master thesis, has worked on developing the infrastructure (scripts and formats) for the integration of the features, and the use of these features for computing correlations with fMRI data. A first set of features has been identified and collected from the novel and from its processing by ALMAnaCH tools (namely FRMG as an instance of a symbolic TAG-based parser and Dyalog-SR, as an instance of an hybrid feature-based neural-based dependency parser). A first dataset of fMRI scan was received to assess the infrastructure and get some preliminary results.

The work is now being continued with the arrival as a post-doc of Murielle Fabre (November 2018). With the expected arrival of the second half of the scans, she will explore more features, use her expertise to interpret the correlations, and guide the choice of new features to be tested. Since her arrival, she has in particular focused on Multi-Word Expressions (MWEs), in particular to be comparable with results published on the English side of the project. We have also identified several kinds of parsing architectures to test, in relation with various complexity parameters: (1) LSTM (two layers), (2) RNNG (with a partile filter), (3) Dyalog-SR et (4) FRMG (TAG).

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

Both Éric de La Clergerie and Murielle Fabre attended the annual meeting of the CRCNS network (Berkeley, June 2018).

6.10. Evaluating the quality of text simplification

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

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

In 2018, our collaboration on text simplification with the Facebook Artificial Intelligence Research lab in Paris (in particular with Antoine Bordes) has started in practice. It has taken the form of a CIFRE PhD. In this context, in 2018, we dedicated important efforts to the problem of the evaluation of text simplification (TS) systems, which remains an open challenge. As the task has common points with machine translation (MT), TS is often evaluated using MT metrics such as BLEU. However, such metrics require high quality reference data, which is rarely available for TS. TS has the advantage over MT of being a monolingual task, which allows for direct comparisons to be made between the simplified text and its original version.

We compared multiple approaches to reference-less quality estimation of sentence-level TS systems, based on the dataset used for the QATS 2016 shared task. We distinguished three different dimensions: grammaticality, meaning preservation and simplicity. We have shown that n-gram-based MT metrics such as BLEU and METEOR correlate the most with human judgment of grammaticality and meaning preservation, whereas simplicity is best evaluated by basic length-based metrics [24].

6.11. Advances in descriptive, computational and historical linguistics

Participants: Benoît Sagot, Laurent Romary, Jack Bowers, Rebecca Blevins.

ALMAnaCH members have resumed their work in descriptive, computational and historical linguistics, an important way to ensure that NLP models and tools are robust to the diversity of world languages, as well as a way to apply NLP models and tools for contributing to research in linguistics. Three of 2018 advances in this regard are the following:

- In the context of the doctoral work of Jack Bowers, a first release of a global documentation of the Mixtepec-Mixtec language has been released which covers, multilayered annotated spoken and written resources as well as a reference lexical resource covering both basic word descriptions and elaborate semantic and etymological (word formation) content [13];
- Work on language description and computational morphology for Romansh Tuatschin in collaboration with Géraldine Walther (Universität Zürich) was pursued, following the work published in 2017 [99]. A new interest in the quantitative, corpus-based study of code switching in this language has emerged in collaboration with Claudia Cathomas (Universität Zürich), leading to preliminary results to be published in 2019;
- We resumed our work in (classical) etymology in collaboration with Romain Garnier (Université de Limoges, Institut Universitaire de France), with a focus not only on (Ancient) Greek and its substrates, but also, more specifically, on Anatolian languages that could be amongst said substrates. In particular, we proposed that Lydian could be the source language for a number of Greek words lacking a good etymology in the literature [31], which motivated Rebecca Blevins's internship on the development of a lexicon of the Lydian language. We also published new etymological results at the (Proto-)Indo-European level [37].

6.12. Language resources and NLP tools for Medieval French

Participants: Éric Villemonte de La Clergerie, Mathilde Regnault, Benoît Sagot.

The main objectives of the ANR project "Profiterole" are to automatically annotate a large corpus of medieval French (9th-15th centuries) in dependency syntax and to provide a methodology for dealing with heterogeneous data like such a corpus (because of diachronic, dialectal, geographic, stylistic and genre-based variation, among other types of linguistic variation). To this end, we have continued previous experiments in morpho-syntactic tagging by trying to determine which parameters and which training sets are the best ones to use when annotating a new text. We explored two approaches for syntactic annotation (i.e. parsing). On the one hand, an ongoing thesis aims at adapting the FRMG metagrammar to medieval French, notably by changing the constraints on certain syntactic phenomena and relaxing the order of words. The development of the OFrLex lexicon has started within the Alexina framework, following the Lefff lexicon for contemporary French [5]. It already allowed for preliminary experiments. On the other hand, we conducted parsing experiments with neural models (DyALog's SRNN models). Note that members of the ALMAnaCH team participated in the CoNLL dependency parsing Shared Task 2018, which included an Old French dataset (see section 6.2).

COML Team

6. New Results

6.1. Speech and Audio Processing from the Raw Waveform

State-of-the-art speech technology systems (e.g., ASR and TTS) rely on fixed, hand-crafted features such as mel-filterbanks to preprocess the waveform before the training pipeline. This is at odds with recent work in machine vision where hand-crafter features (SIFT, etc) have been succesfully replaced by features derived from raw pixels trained jointly with a downstream task. In this line of work, we explored how a similar approach could be undertaken for audio and speech processing.

- In [24], we train a bank of complex filters that operates at the level of the raw speech signal and feeds into a convolutional neural network for phone recognition. These time-domain filterbanks (TD-filterbanks) are initialized as an approximation of MFSC, and then fine-tuned jointly with the remaining convolutional network. We perform phone recognition experiments on TIMIT and show that for several architectures, models trained on TD-filterbanks consistently out-perform their counterparts trained on comparable MFSC. We get our best performance by learning all front-end steps, from pre-emphasis up to averaging. Finally, we observe that the filters at convergence have an asymmetric impulse response while preserving some analyticity.
- In [25], we study end-to-end systems trained directly from the raw waveform, building on two alternatives for trainable replacements of mel-filterbanks that use a convolutional architecture. The first one is inspired by gammatone filterbanks [4], [9], and the second one by the scattering transform [24]. We propose two modifications to these architectures and systematically compare them to mel-filterbanks, on the Wall Street Journal dataset. The first modification is the addition of an instance normalization layer, which greatly improves on the gammatone-based trainable filterbanks and speeds up the training of the scattering-based filterbanks. The second one relates to the low-pass filter used in these approaches. These modifications consistently improve performances for both approaches, and remove the need for a careful initialization in scattering-based trainable filterbanks. In particular, we show a consistent improvement in word error rate of the trainable filterbanks relatively to comparable mel-filterbanks. It is the first time end-to-end models trained from the raw signal significantly outperform mel-filterbanks on a large vocabulary task under clean recording conditions.
- Recent progress in deep learning for audio synthesis opens the way to models that directly produce the waveform, shifting away from the traditional paradigm of relying on vocoders or MIDI synthesizers. Despite their successes, current state-of-the-art neural audio synthesizers such as WaveNet and SampleRNN [12], [8] suffer from prohibitive training and inference times because they are based on autoregressive models that generate audio samples one at a time at a rate of 16kHz. In this work [26], we study the more computationally efficient alternative of generating the waveform frame-by-frame with large strides. We present SING, a lightweight neural audio synthesizer for the original task of generating musical notes given desired instrument, pitch and velocity. Our model is trained end-to-end to generate notes from nearly 1000 instruments with a single decoder, thanks to a new loss function that minimizes the distances between the log spectrograms of the generated and target waveforms. On the generalization task of synthesizing notes for pairs of pitch and instrument not seen during training, SING produces audio with significantly improved perceptual quality compared to a state-of-the-art autoencoder based on WaveNet [4] as measured by a Mean Opinion Score (MOS), and is about 32 times faster for training and 2,500 times faster for inference.

6.2. Development of cognitively inspired algorithms

Speech and language processing in humans infants and adults is particularly efficient. We use these as sources of inspiration for developing novel machine learning and speech technology algorithms. In this area, our results are as follows:

- In [22], we summarize the accomplishments of a multi-disciplinary 6-weeks workshop organized by E. Dupoux (PI) at Carnegy Mellon Univerrsity (Pittsburgh), funded through the Jelinek Memorial Summer Workshop Program of Johns Hopkins University. The workshop explored the computational and scientific issues surrounding the discovery of linguistic units (subwords and words) in a language without orthography. We studied the replacement of orthographic transcriptions by images and/or translated text in a well-resourced language to help unsupervised discovery from raw speech.
- Developing speech technologies for low-resource languages has become a very active research field over the last decade. Among others, Bayesian models have shown some promising results on artificial examples but still lack of in situ experiments. In [20] , we apply state-of-the-art Bayesian models to unsupervised Acoustic Unit Discovery (AUD) in a real low-resource language scenario. We also show that Bayesian models can naturally integrate information from other resourceful languages by means of informative prior leading to more consistent discovered units. Finally, discovered acoustic units are used, either as the 1-best sequence or as a lattice, to perform word segmentation. Word segmentation results show that this Bayesian approach clearly outperforms a Segmental-DTW baseline on the same corpus.
- Fixed-length embeddings of words are very useful for a variety of tasks in speech and language processing. In [19], we systematically explore two methods of computing fixed-length embeddings for variable-length sequences. We evaluate their susceptibility to phonetic and speaker-specific variability on English, a high resource language, and Xitsonga, a low resource language, using two evaluation metrics: ABX word discrimination and ROC-AUC on same-different phoneme n-grams. We show that a simple downsampling method supplemented with length information can be competitive with the variable-length input feature representation on both evaluations. Recurrent autoencoders trained without supervision can yield even better results at the expense of increased computational complexity.
- Recent studies have investigated siamese network architectures for learning invariant speech representations using same-different side information at the word level. In [21], we investigate systematically an often ignored component of siamese networks: the sampling procedure (how pairs of same vs. different tokens are selected). We show that sampling strategies taking into account Zipf's Law, the distribution of speakers and the proportions of same and different pairs of words significantly impact the performance of the network. In particular, we show that word frequency compression improves learning across a large range of variations in number of training pairs. This effect does not apply to the same extent to the fully unsupervised setting, where the pairs of same-different words are obtained by spoken term discovery. We apply these results to pairs of words discovered using an unsupervised algorithm and show an improvement on state-of-the-art in unsupervised representation learning using siamese networks.
- Unsupervised spoken term discovery is the task of finding recurrent acoustic patterns in speech without any annotations. Current approaches consists of two steps: (1) discovering similar patterns in speech, and (2) partitioning those pairs of acoustic tokens using graph clustering methods. In, [23] we propose a new approach for the first step. Previous systems used various approximation algorithms to make the search tractable on large amounts of data. Our approach is based on an optimized k-nearest neighbours (KNN) search coupled with a fixed word embedding algorithm. The results show that the KNN algorithm is robust across languages, consistently outperforms the DTW-based baseline, and is competitive with current state-of-the-art spoken term discovery systems.

6.3. Test of the psychological validity of AI algorithms.

In this section, we focus on the utilisation of machine learning algorithms of speech and language processing to derive testable quantitative predictions in humans (adults or infants).

- Two PhDs were defended this year. In [14], Adriana Guavara Rukoz presented a computational model of the perception of non-native speech contrasts based on standard ASR pipelines is presented. An adaptation of the model is proposed to account for forced-choice classification psycholinguistic experiments and directly reproduced classical results. The general finding is that, suprisingly, the acoustic model part of a phone recognizer is sufficient to account for experimental data, even those apparently related to phonotactic properties of the native language. The 'language model' part does not improve the correlation with adult data (if anything, it degrades it). Yet the match between model and human is not perfect, and is was hypothetized that improvement in the acoustic model could help. In [13], Julia Maria Carbajal presented a study of the effect of multilingual exposure on language acquisition. She used a computational model of language separation based on i-vectors to reproduce some of the known effects of phonological distance on language discrimination in infants.
- In [16], we investigate whether infant-directed speech (IDS) facilitates lexical learning when compared to adult-directed speech (ADS). To study this, we compare the distinctiveness of the lexicon at two levels, acoustic and phonological, using a large database of spontaneous speech in Japanese. At the acoustic level we show that, as has been documented before for phonemes, the realizations of words are more variable and less discriminable in IDS. At the phonological level, we find that despite a slight increase in the number of phonological neighbors, the IDS lexicon contains more distinctive words (such as onomatopeias). Combining the acoustic and phonological metrics together in a global discrimination score, the two effects cancel each other out and the IDS lexicon winds up being as discriminable as its ADS counterpart. We discuss the implication of these findings for the view of IDS as hyperspeech, i.e., a register whose purpose is to facilitate language acquisition.
- Existing theories of cross-linguistic phonetic category perception agree that listeners perceive foreign sounds by mapping them onto their native phonetic categories. Yet, none of the available theories specify a way to compute this mapping. As a result, they cannot provide systematic quantitative predictions and remain mainly descriptive. Here [17], Automatic Speech Recognition (ASR) systems are used to provide a fully specified mapping between foreign and native sounds. This is shown to provide a quantitative model that can account for several empirically attested effects in human cross-linguistic phonetic category perception.
- Spectacular progress in the information processing sciences (machine learning, wearable sensors) promises to revolutionize the study of cognitive development. In [15], we analyse the conditions under which 'reverse engineering' language development, i.e., building an effective system that mimics infant's achievements, can contribute to our scientific understanding of early language development. We argue that, on the computational side, it is important to move from toy problems to the full complexity of the learning situation, and take as input as faithful reconstructions of the sensory signals available to infants as possible. On the data side, accessible but privacy-preserving repositories of home data have to be setup. On the psycholinguistic side, specific tests have to be constructed to benchmark humans and machines at different linguistic levels. We discuss the feasibility of this approach and present an overview of current results.

6.4. Applications and tools for researchers

Some of CoMLs' activity is to produce speech and language technology tools that facilitate research into language development or clinical applications.

- In [18], we present BabyCloud, a platform for capturing, storing and analyzing daylong audio recordings and photographs of children's linguistic environments, for the purpose of studying infant's cognitive and linguistic development and interactions with the environment. The proposed platform connects two communities of users: families and academics, with strong innovation potential for each type of users. For families, the platform offers a novel functionality: the ability for parents to follow the development of their child on a daily basis through language and cognitive

metrics (growth curves in number of words, verbal complexity, social skills, etc). For academic research, the platform provides a novel means for studying language and cognitive development at an unprecedented scale and level of detail. They will submit algorithms to the secure server which will only output anonymized aggregate statistics. Ultimately, BabyCloud aims at creating an ecosystem of third parties (public and private research labs...) gravitating around developmental data, entirely controlled by the party whose data originate from, i.e. families.

RITS Project-Team

7. New Results

7.1. Deep Reinforcement Learning for end-to-end driving

Participants: Raoul de Charette, Maximilian Jaritz, Fawzi Nashashibi.

Following the work initiated in 2017, we continued the work on end-to-end driving using with asynchronous reinforcement learning directly. The network learns to map low level control directly with RGB images. To continue previous works initiated, we have applied recent domain adaptation and evaluated our reinforcement learning (learn in a realistic car game) in open-loop on real video footage, showing promising adaptation results. New outcome also include tests on real data (web footage). This led to a publication in ICRA [25]. This research was partially funded by Valeo.

7.2. Convolutional neural networks for Semantic and Completion with Sparse and Dense Data

Participants: Raoul de Charette, Maximilian Jaritz, Fawzi Nashashibi.

Deep convolutional networks have outperform all previous techniques on most vision tasks. This is because they are able to utilize dense data and extract relationship between local information such as gradients, or high level features. However, convolutional neural networks (CNNs) require dense data and are known to fail when data is sparse. Here, we address the research problem and proposed a solution. Instead of using a sparse convolution methodology, we show that using the right architecture with a proper training strategy the network can learn sparsity invariant feature while remaining stable when dense data are present. Our architecture uses an encoder-decoder version of Mobile NasNet with skip connections. The results show that we can accomplish both data completion or semantic segmentation changing only the last layer of the network. Performance obtained were published on Kitti Benchmark and ranks among the first ones, and the methodology was published in 3DV [26]. This research was partially funded by Valeo.

7.3. Realistic Weather Augmentation for Evaluation of Bad Weather in Computer Vision

Participants: Raoul de Charette, Shirsendu Halder.

Computer vision is evaluated on extensive databases that include large number of examples and allow the ranking of algorithms. However, all databases are acquired in clear weather conditions, where the atmosphere is a transparent medium. In rain/snow/fog, when the atmosphere is filled with particles the light is refracted/reflected/diffracted and the appearance is altered. Here we propose a new research that augment existing databases with new weather or arbitrary amount. We applied it on Kitti and Cityscapes. Our approach uses an accurate understanding of physical and optics models to generate realistic rain/fog and augment existing images or sequences. This allows us to evaluate state-of-the-art vision algorithms for both object detection and semantics and quantitatively measure the effect of rain or fog on them. This research was conducted in collaboration with Jean-Francois Lalonde from Université Laval and was supported by Samuel de Champlain Quebec-France collaboration program.

7.4. Perception for Cooperative Driving

Participants: Pierre Bourre, Raoul de Charette, Carlos Flores, Renaud Poncelet, Luis Roldao, Dinh-Van Nguyen.

In the context of multiple autonomous vehicles, sharing the perception of each other allows an enriched perception of the environment. For the PACV2x FUI project, we propose a mix of vision sensors and communication exchanges is used for merging, overtaking, and other risky situations that benefit from multi perception. A speed planning algorithm as well as low level control and lidar data clustering were developed to allow a small fleet of two to three vehicles to handle such scenarios. The vehicles use communication and GPS coordinates to closely follow a planned trajectory.

7.5. A Statistical Update of Grid Representations from Range Sensors

Participants: Luis Roldao, Raoul de Charette, Anne Verroust-Blondet.

An accurate 3D model of the surrounding environment is a fundamental feature for autonomous vehicles to perform different tasks such as obstacle detection, localization and mapping. While continuous representations are widely used in the literature, we prefer to use a three dimensional discrete grid representation in this work in order to reduce memory an computational complexity. In this case, each grid cell represents the occupancy state of a portion of the environment in a probabilistic manner.

By definition, a discretized representation inhibits a completely accurate reconstruction. Therefore, grid models are unable to create a perfect model of the surroundings. In the literature, it is usually considered that within a single scan, the state of each cell is binary (free or occupied). Hence, a cell is set occupied if at least one impact occurred within, and free if it has been traversed by any ray. The problem of such an approach is that the complete state of the cell is updated from a single partial observation, neglecting the contribution of multiple measurements and their validity. Moreover, the traversed distance of the rays within each cell is usually ignored.

Towards the goal of achieving a more accurate representation, we propose a different way to update the occupancy probability of each cell according to the observations; considering the traversed distance of the rays within each cell (ray-path information), the contribution of the complete set of observations within the cell, and the density of observations that can be obtained at such cell according to its distance from the sensor. Proposed method was evaluated in both simulated and real data. Reconstruction results show an improvement on the representation of the surroundings with less occupancy state errors in the cells of the grid. Future works will include the comparison against a continuous representation to test the accuracy along with the time and computation needs for both representations.
More details can be found in [38] and [30]. This research is partially funded by AKKA Technology.

7.6. Recognizing Pedestrians using Cross-Modal Convolutional Networks

Participants: Danut-Ovidiu Pop, Fawzi Nashashibi.

This year, we have continued our research, which is based on multi-modal image fusion schemes with deep learning classification methods. We propose four different learning patterns based on Cross-Modality deep learning of Convolutional Neural Networks:

(1) a Particular Cross-Modality Learning;

(2) a Separate Cross-Modality Learning;

(3) a Correlated Cross-Modality Learning and

(4) an Incremental CrossModality Learning model.
Moreover, we also design a new variation of a Lenet architecture, which improves the classification performance. Finally, we propose to learn this model with the incremental cross-modality approach using optimal learning settings, obtained with a K-fold Cross Validation pattern. This method outperforms the state-of-the-art classifier provided with Daimler datasets on both non-occluded and partially-occluded pedestrian tasks.

7.7. Vehicle Trajectory Prediction

Participants: Kaouther Messaoud, Itheri Yahiaoui, Anne Verroust-Blondet, Fawzi Nashashibi.

In order to enhance the road safety, the first and the most important step for an effective autonomous navigation is the environment perception and surrounding objects recognition. So, advanced sensing systems are mounted in vehicles to monitor the on-road environment. One of the most challenging tasks is to understand, analyze the driving situations and make a reasonable and safe navigation decisions accordingly. Human drivers make decisions while implicitly reasoning about how neighboring drivers will move in the future. In this context, we aim to predict the motion of drivers neighboring an autonomous vehicle based on data captured using deployed sensors.
This year, we studied the state of the art approaches for trajectory and maneuver prediction. We focused on general trajectory prediction representation while considering interactions between the neighboring drivers using different types of neural networks such as recurrent and convolutional neural networks.

7.8. WiFi Fingerprinting Localization for Intelligent Vehicles in Car Park

Participants: Dinh-Van Nguyen, Raoul de Charette, Fawzi Nashashibi.

A novel method of WiFi fingerprinting for localizing intelligent vehicles in GPS-denied area, such as car parks, has been proposed. Although the method itself is a popular approach for indoor localization application, adapting it to the speed of vehicles requires different treatment. By deploying an ensemble neural network for fingerprinting classification, the method shows a reasonable localization precision at car park speed. Furthermore, a Gaussian Mixture Model (GMM) Particle Filter is applied to increase localization frequency as well as accuracy. Experiments show promising results with average localization error of 0.6m (cf. [29]).

A more complete study on the use of Wifi fingerprinting for solving the localization problem for autonomous vehicles in GPS-denied environments is presented in the thesis manuscript entitled "Wireless Sensors Networks for Indoor Mapping and Accurate Localization for Low Speed Navigation in Smart Cities" (cf. [11]).

7.9. Enhancing the Accuracy of SLAM-based Localization Systems for Autonomous Driving

Participants: Zayed Alsayed, Anne Verroust-Blondet, Fawzi Nashashibi.

Computing a reliable and accurate pose for a vehicle in any situation is one of the challenges for Simultaneous Localization And Mapping methods (SLAM) methods. Based on the probabilistic form of the SLAM solution, SLAM methods suffer from systematic errors related to the linearization of the solution models. The accuracy of the SLAM method can be improved by estimating a correction to be applied to the SLAM output based on relevant information available from the SLAM algorithm. In [20] two approaches predicting corrections to be applied to SLAM estimations are proposed:

1) The first approach is designed for 2D SLAM methods, i.e. independently of the underlying SLAM process and sensor used, where we aim to reduce the errors due to the dynamical modeling during specific maneuvers.

2) The second method is designed to handle errors related to the probabilistic formulation of Maximum Likelihood SLAM approaches, and thus it is suitable for 2D Maximum Likelihood SLAM methods (i.e. no assumptions on the sensor used).
The validity of both approaches was proved through two experiments using different evaluation metrics and using different sensor characteristics.
More detail can be found in the thesis manuscript of Zayed Alsayed entitled "Characterizing the Robustness and Enhancing the Accuracy of SLAM-based Localization Systems for Autonomous Driving" (cf. [7]).

7.10. LIDAR-based lane marking detection for vehicle localization

Participants: Farouk Ghallabi, Fawzi Nashashibi.

Accurate self-vehicle localization is an important task for autonomous driving and ADAS. Current GNSS-based solutions do not provide better than 2-3 m in open-sky environments. In order to achieve lane-level accuracy, a lane marking detection system using a multilayer LIDAR (velodyne) and a map matching algorithm has been introduced. The perception system includes three different steps: road segmentation, image construction and line detection. Our road segmentation method purely relies on geometric analysis of each layer returns. Detected lane markings are matched to a prototype third party map which was built with absolute accuracy = 5cm. The map matching algorithm is a particle filtering process that achieves lane-level accuracy (20 cm). More details are in [23]. This work has been partially funded by Renault.

7.11. Motion Planning among Highly Dynamic Obstacles

Participants: Pierre de Beaucorps, Anne Verroust-Blondet, Renaud Poncelet, Fawzi Nashashibi.

Motion planning in a dynamic environment is of great importance in many robotics applications. In the context of an autonomous mobile robot, it requires computing a collision-free path from a start to a goal among moving and static obstacles. We have introduced a framework to integrate into a motion planning method the interaction zones of a moving robot with its future surroundings, the reachable interaction sets (RIS). It can handle highly dynamic scenarios when combined with path planning methods optimized for quasi-static environments. It has been integrated with an artificial potential field reactive method and with a Bézier curve path planning. Experimental evaluations show that this approach significantly improves dynamic path planning methods, especially when the speeds of the obstacles are higher than the one of the robot (cf. [32] for more detail). This work has been partially funded by Valeo.

7.12. Control Architecture for Adaptive and Cooperative Car-Following

Participants: Carlos Flores, Fawzi Nashashibi.

The general scope of this work deals with three open challenges in the state-of-the-art of cooperative car-following systems:

1) Deal with the impact of not only communication links delays, but also heterogeneity between vehicles' dynamics in the same string. This should be targeted ensuring the gap-regulation robustness without degrading the expected performance to keep car-following benefits (individual and string stability). In particular, when a heterogeneous string is formed, the differences between vehicles dynamics introduce disturbances in the closed loop system affecting the string stability. In [22] we presented an online Cooperative Adaptive Cruise Control (CACC) feedforward adaptation with a fractional-order feedback controller for stable heterogeneous strings of vehicles. Simulations demonstrate the advantages over conventional homogeneous structures as well as system's capability to both enhance stability and guarantee string stability regardless the vehicles distribution.

2) Design a modular architecture that permits to introduce cooperative string driving in urban environments, where interaction with vulnerable road users is highly probable. In this context, a cooperative car-following/emergency braking system with prediction-based pedestrian avoidance capabilities using vehicle-to-vehicle and vehicle-to-pedestrian communication links has been proposed in [14] and validated with RITS platforms.

3) Further extend the benefits of Adaptive Cruise Control (ACC) and Cooperative Adaptive Cruise Control (CACC) applications on traffic flow and safety, having strict $\mathcal{L}_2$ string stability as a hard constraint, employing different calculus techniques for the control design task. A fractional-order-based control algorithm is employed to enhance the car-following and string stability performance for both ACC and CACC vehicle strings, including communication temporal delay effects has been presented in [15]. Simulation and real experiments have been conducted for validating the approach.

The aforementioned contributions have been developed in the framework of the VALET project ANR-15-CE22-0013. They have been also implemented in the vehicle platforms of RITS team, for the sake of validation and further demonstration of the final VALET system.

This scientific work can be found as well in the thesis manuscript of Carlos Flores entitled "Control Architecture for Adaptive and Cooperative Car-Following" (cf. [8]).

7.13. Stability analysis for controller switching in autonomous vehicles

Participants: Francisco Navas, Imane Mahtout, Fawzi Nashashibi.

This work investigates the Youla-Kucera (YK) parameterization to provide stable responses for autonomous vehicles when dynamics or environmental changes occur. This work explores the use of the YK parameterization in dynamics systems such as vehicles, with special emphasis on stability when some dynamic change or the traffic situation demands controller reconfiguration:

- YK parameterization provides all stabilizing controllers for a given plant. This is used in order to perform stable controller reconfiguration. Different YK-based control structures are obtained for dealing with problems such order complexity, plant disconnection or matrix inversability. Stability properties are preserved even if different structures are employed, but transient behavior between controllers changes depending on the employed YK-based structure. One of the structures presents the best transient behavior without oscillations, a lower order controller complexity and no need to disconnect the initial controller, which would be important if the system shutdown is very expensive, or the initial controller is part of a safety circuit [28]. This structure is used together with CACC applications improving CACC state-of-the-art. An hybrid behavior between two CACC controllers with different time gaps is explored by means of the YK parameterization, in order to avoid ACC degradation when communication link with preceding vehicle is lost. The proposed system uses YK parameterization and communication with a vehicle ahead (different from the preceding one) providing stable responses and, more interestingly, reducing intervehicle distances in comparison with an ACC degradation. A similar idea of hybrid behavior between CACC controller with different time gap is developed for entering/exiting vehicles in the string. In that case, YK parameterization is able to ensure stability of these merging/splitting maneuvers.
- Dual YK parameterization provides all the plants stabilized by a controller. This is employed for solving CL identification problems, or adaptive control solutions, which integrate identification and controller reconfiguration processes. YK-based CL identification uses classical OL identification algorithms, providing better results than if it is used alone. Results in a CACC-equipped vehicle prove how CL nature of the data affects a classical OL identification algorithm, and how dual YK parameterization helps to mitigate these effects. Finally, an adaptive control application is developed by using MMAC. Longitudinal dynamics of two vehicles in a CACC string are estimated within a model set, so the good CACC sytem can be chosen even if a heterogeneous string of vehicles is considered. Dynamics estimation results much more faster than other estimation processes in the literature.
- Different types of controllers and structures are used throughout Francisco Navas thesis ([10]), proving the adaptability of the YK parameterization to any type of controller. Simulation and experimental results demonstrate real implementation of stable controller reconfiguration, CL identification and adaptive control solutions dealing with dynamics changes or different traffic situations. The author thinks that YK is a suitable control framework able to ensure responses in autonomous driving.
- In [27] a design and implementation of a novel lateral control approach is proposed within Imane Mahtout thesis work. The control strategy is based on Youla-Kucera parametrization to switch gradually between controllers that are designed separately for big and small lateral errors. The presented approach studies the critical problem of initial lateral error in line following. It ensures smooth and stable transitions between controllers and provides a smooth vehicle response regardless of the lateral error. For an initial validation the work was tested in simulation, implementing a dynamic bicycle model. It has also been tested in real platforms implementing an electric Renault ZOE, with good results when activating the system at different lateral errors. Current work is focused on using YK-parametrization in estimating lateral vehicle dynamics.

7.14. Belief propagation inference for traffic prediction

Participant: Jean-Marc Lasgouttes.

This work [45], [44], in collaboration with Cyril Furtlehner (TAU, Inria), deals with real-time prediction of traffic conditions in a urban setting with incomplete data. The main focus is on finding a good way to encode available information (flow, speed, counts,...) in a Markov Random Field, and to decode it in the form of real-time traffic reconstruction and prediction. Our approach relies in particular on the Gaussian belief propagation algorithm.

This year, continuing our collaboration with PTV Sistema, we improved our techniques and obtained extensive results on large-scale datasets containing 250 to 2000 detectors. The results show very good ability to predict flow variables and a reasonably good performance on speed or occupancy variables. Some element of understanding of the observed performance are given by a careful analysis of the model, allowing to some extent to disentangle modelling bias from intrinsic noise of the traffic phenomena and its measurement process [35].

7.15. Large scale simulation interfacing

Participant: Jean-Marc Lasgouttes.

The SINETIC FUI project aims to build a complete simulation environment handling both mobility and communication. 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 the IEEE 802.11p protocol. This simulation tool is 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 [31].

7.16. Platoons Formation for autonomous vehicles redistribution

Participants: Mohamed Hadded, Jean-Marc Lasgouttes, Fawzi Nashashibi, Ilias Xydias.

In this paper, we consider the problem of vehicle collection assisted by a fleet manager where parked vehicles are collected and guided by fleet managers. Each platoon follows a calculated and optimized route to collect and guide the parked vehicles to their final destinations. The Platoon Route Optimization for Picking up Automated Vehicles problem, called PROPAV, consists in minimizing the collection duration, the number of platoons and the total energy required by the platoon leaders. We propose a formal definition of PROPAV as an integer linear programming problem, and then we show how to use the Non-dominated Sorting Genetic Algorithm II (NSGA-II), to deal with this multi-criteria optimization problem. Results in various configurations are presented to demonstrate the capabilities of NSGA-II to provide well-distributed Pareto-front solutions.

This work has been presented at ITSC 2018 conference [24].

7.17. Prediction-based handover between VLC and IEEE 802.11p for vehicular environment

Participants: Mohammad Abualhoul, Fawzi Nashashibi.

Despite years of development and deployment, the standardized IEEE 802.11p communication for vehicular networks can be pushed toward insatiable performance demands for wireless network data access, with a remarkable increase of both latency and channel congestion levels when subjected to scenarios with a very high vehicle density.

In specific hard safety applications such as convoys, IEEE 802.11p could seriously fail to meet the fundamental vehicular safety requirements. On the other hand, the advent of LED technologies has opened up the possibility of leveraging the more robust Visible Light Communication (VLC) technology to assist IEEE 802.11p and provide seamless connectivity in dense vehicular scenarios.

In this particular research, we proposed and validated a Prediction-based Vertical handover (PVHO) between VLC and IEEE 802.11p meant to afford seamless switching and ensure the autonomous driving safety requirements [19].

Algorithm validation and platoon system performance were evaluated using a specially implemented IEEE 802.11p-VLC module in the NS3 Network Simulator. The simulation results showed a speed-based dynamic redundancy before and after VLC interruptions with seamless switching. Moreover, the deployment of VLC for platoon intra-communication can achieve a 10-25% PDR gain in high-density vehicular scenarios, where the work was published in the IEEE International Conference on Intelligent Transportation Systems 2018.

7.18. Lane-Centering to Ensure the Visible Light Communication (VLC) Connectivity for a Platoon of Autonomous Vehicles

Participants: Mohammad Abualhoul, Fawzi Nashashibi.

VLC technology limitations were defined and supported by different solutions proposals to enhance the crucial alignment and mobility limitations. In this research [17], we proposed the incorporation of the VLC technology and a Lane-Centering (LC) technique to assure the VLC-connectivity by keeping the autonomous vehicle aligned to the lane center using vision-based lane detection in a convoy-based formation. Such combination can ensure the optical communication connectivity. This contribution by RITS-Team won the best paper award during the ICVES conference.

The system performance and evaluation showed that ss soon as the road lanes are detectable, the evaluated results showed stable behavior independently from the inter-vehicle distances and without the need for any exchanged information of the remote vehicles. Further investigations are to be carried-out in this direction.

7.19. Cyberphysical Constructs for Next-Gen Vehicles and Autonomic Vehicular Networks

Participant: Gérard Le Lann.

Behaviors of Connected Automated Vehicles (CAVs) rest on robotics capabilities (sensors, motion control laws, actuators) and wireless radio communications. Reduction of non-harmful crashes and fatalities despite higher vehicular density (safety and efficiency properties) is a fundamental objective, whatever the SAE automated driving levels considered (use cases).

Based on "hard sciences", onboard robotics capabilities designed so far are satisfactory for numerous settings, to the exception of non-line-of-sight scenarios. That is the rationale for wireless radio communications. Over the years, a growing fraction of the scientific community has been questioning the adequacy of current IEEE and ETSI standards aimed at automotive wireless communications, herein referred to as wave protocols (wireless access in vehicular environment) for convenience.

Analyses based on well-known results in various areas such as life/safety-critical systems, distributed algorithms, dependable real-time computing, ad hoc mobile networking, and cyber-physics (to name a few) come to the conclusion that wave protocols do not meet essential requirements regarding safety, efficiency, privacy or cybersecurity (SPEC). These conclusions are based on scientific demonstrations. Notably, wave protocols rest on intuitive designs (no proofs, only simulations or experimental testing) that violate well-known impossibility results in asynchronous or synchronous systems. It follows that future vehicles shall be commanded and controlled by onboard robotics supplemented with wireless communication capabilities other than wave protocols. These vehicles are referred to as Next-Gen Vehicles (NGVs) in order to avoid confusion with CAVs.

That wave solutions are far from being convincing is at the core of the recommendations issued at the EU level (the latest WG29 resolution). Moreover, the important question of how to instantiate the EU GDPR directive in future CAVs is left unanswered, despite the fact that it is possible (proofs provided) to achieve safety and privacy jointly. Preliminary results for NGVs have appeared in [34].

The work reported herein, started in 2017 along with international researchers, aims at specifying solutions to the SPEC problem, considering self-organizing and self-healing Autonomic Vehicular Networks (AVNs) of NGVs. Parallel to this, risks of privacy breaches and cyberattacks proper to wave solutions have been exposed to the public via invited interventions and presentations.

An issue not very well addressed so far is to which extent robotics and computer science supplement each other. The cyber-physical perspective is essential to formulate a coherent vision. In cyber space and in physical space, safety has to do with resource sharing. Deadlock-free and fair resource sharing in systems of concurrent processes has been a major topic in computer science for more than 50 years. Asphalt (2D systems), asphalt and air space (3D systems) are the shared resources of interest in the physical space.

As is well known, there are three classes of algorithmic solutions: detection-and-recovery, prevention, avoidance. The former class is inapplicable (one cannot "roll back an accident"). Prevention is aimed at prohibiting the emergence of hazardous (no safety) or deadlock-prone (no safety, no efficiency) conditions. Solutions are the province of distributed algorithms (computer science). Avoidance is relied on for maintaining non-hazardous conditions while making progress (also, in case some of the assumptions that underlie prevention schemes would be violated). Solutions are the province of automation control (linear/non-linear dynamics).

Prevention and avoidance schemes are needed, put in action as follows: NGVs run (cyber) distributed agreement algorithms in order to preclude the emergence of hazardous conditions, prior to executing physical motions (collision-free trajectories), which motions are made feasible thanks to prevention schemes. This is how computer science and robotics can be "married" consistently: with prevention schemes, one achieves proactive safety, and with avoidance schemes, one achieves reactive safety (both types are needed).

NGVs and AVNs are life/safety-critical cyber-physical systems. Consequently, correct solutions to the SPEC problem are based on cyber-physical constructs endowed with appropriate intrinsic properties. We have devised the cell and the cohort constructs, which rest on the obvious observation according to which only vehicles sufficiently close to each other may experience a collision. Time-bounded ultra-fast message-passing and inter-vehicular coordination can be achieved within these constructs thanks to very short-range radio and optical communications, as well as deterministic protocols (MAC protocols in particular) and distributed algorithms (dissemination, approximate agreement, and consensus). Analytical expressions of upper bounds for message-passing and inter-vehicular coordination are established for worst-case conditions, such as contention and failures, message losses in particular. We have shown that these solutions can sustain message loss frequencies an order of magnitude higher than frequencies beyond which none of the wave protocols could work.

We have defined the concept of cyberphysical levels, which are orthogonal to SAE automated driving levels. Joining a cohort longitudinally or laterally (which implies a lane change) is conditioned on a number of criteria, such as cyberphysical levels, NGV sizes, and proof of authentication (requestor's name must be a certified pseudonym).

Naming raises open problems in spontaneous mobile open systems, such as AVNs. Privacy-preserving naming is even more complex. The "longitudinal privacy-preserving naming" problem is solved with the cohort construct. The "lateral privacy-preserving naming" problem which arises with NGVs members of a cell or that circulate in adjacent cohorts has solutions based on combined optical and radio communications.

Novel deterministic time-bounded MAC protocols at the core of distributed coordination algorithms are needed to solve the open problem of safe entrances into unsignalized intersections of arbitrary topologies (any number of arterials, any number of lanes per arterial) in the presence of noisy radio channels. This problem has been solved with CSMA-CD/DCR (deterministic collision resolution) MAC protocols.

7.20. Functional equations

Participant: Guy Fayolle.

The article [13] presents functional equations (involving one or two complex variables) as an Important analytic method in stochastic modelling and in combinatorics.

7.21. Optimization of test case generation for ADAS via Gibbs sampling algorithms

Participant: Guy Fayolle.

Validating Advanced Driver Assistance Systems (ADAS) is a strategic issue, since such systems are becoming increasingly widespread in the automotive field.

But ADAS validation is a complex issue, particularly for camera based systems, because these functions may be facing a very high number of situations that can be considered as infinite. Building at a low cost level a sufficiently detailed campaign is thus very difficult. Indeed, test case generation faces the crucial question of *inherent combinatorial explosion*. An important constraint is to generate *almost all* situations in the most economical way. This task, in general, can be considered from two points of view: deterministic via binary search trees, or stochastic via Markov chain Monte Carlo (MCMC) sampling. We choose the latter probabilistic approach described below, which in our opinion seems to be the most efficient one. Typically, the problem is to produce samples of large random vectors, the components of which are possibly dependent and take a finite number of values with some given probabilities.The following flowchart is proposed.

1. In a first step, starting from the simulation graph generated by the toolboxes of MATLAB, we construct a so-called *Markov Random Field (MRF)* . When the parameters are locally dependent, this can be achieved from the user's specifications and by a systematic application of Bayes' formula.
2. Then, to cope with the combinatorial explosion, test cases are produced by implementing (and comparing) various *Gibbs samplers*, which are fruitfully employed for large systems encountered in physics. In particular, we strive to make a compromise between the convergence rate toward equilibrium, the percentage of generated duplicates and the path coverage, recalling that the speed of convergence is exponential, a classical property deduced from the general theory of Markov chains.
3. The problem of generating rare events by mixing Gibbs samplers, Large Deviation Techniques (LDT) and cross-entropy method a work in progress.

The French car manufacturer *Groupe PSA* shows a great interest in these methods and has established a contractual collaboration involving ARMINES-Mines ParisTech (Guy Fayolle as associate researcher) and Can Tho University in Vietnam (Pr. Van Ly Tran).

7.22. Random walks in orthants and lattice path combinatorics

Participant: Guy Fayolle.

In the second edition of the book [2], original methods were proposed to determine the invariant measure of random walks in the quarter plane with small jumps (size 1), the general solution being obtained via reduction to boundary value problems. Among other things, an important quantity, the so-called *group of the walk*, allows to deduce theoretical features about the nature of the solutions. In particular, when the *order* of the group is finite and the underlying algebraic curve is of genus 0 or 1, necessary and sufficient conditions have been given for the solution to be rational, algebraic or D-finite (i.e. solution of a linear differential equation). In this framework, number of difficult open problems related to lattice path combinatorics are currently being explored, in collaboration with A. Bostan and F. Chyzak (project-team SPECFUN, Inria-Saclay), both from theoretical and computer algebra points of view: concrete computation of the criteria, utilization of differential Galois theory, genus greater than 1 (i.e. when some jumps are of size ≥ 2), etc. A recent topic of future research deals with the connections between simple product-form stochastic networks (so-called *Jackson networks*) and explicit solutions of functional equations for counting lattice walks.

VALDA Project-Team

6. New Results

6.1. Query Enumeration

Query enumeration is the problem of enumerating the results of a query over a database one by one; the goal is to obtain, after some initial low preprocessing time (e.g., linear in the data), one solution after the other with low delay (e.g., constant-time) in between.

In a first work [26], we consider the enumeration of MSO queries over strings under updates. For each MSO query we build an index structure enjoying the following properties: The index structure can be constructed in linear time, it can be updated in logarithmic time and it allows for constant delay time enumeration. This improves from the previous known index structures allowing for constant delay enumeration that would need to be reconstructed from scratch, hence in linear time, in the presence of updates. We allow relabeling updates, insertion of individual labels and removal of individual labels.

In a second work [29], we consider the evaluation of first-order queries over classes of databases that are nowhere dense. The notion of nowhere dense classes was introduced by Nešetřil and Ossona de Mendez as a formalization of classes of "sparse" graphs and generalizes many well-known classes of graphs, such as classes of bounded degree, bounded treewidth, or bounded expansion. It has recently been shown by Grohe, Kreutzer, and Siebertz that over nowhere dense classes of databases, first-order sentences can be evaluated in pseudo-linear time (pseudo-linear time means that for all ε there exists an algorithm working in time $O(n^{1+\varepsilon})$, where n is the size of the database). For first-order queries of higher arities, we show that over any nowhere dense class of databases, the set of their solutions can be enumerated with constant delay after a pseudo-linear time preprocessing. In the same context, we also show that after a pseudo-linear time preprocessing we can, on input of a tuple, test in constant time whether it is a solution to the query.

6.2. Provenance Circuits

We are interested in obtaining efficiently compact representation of the provenance of a query over a database.

In [28], we generalize three existing graph algorithms to compute the provenance of regular path queries over graph databases, in the framework of provenance semirings – algebraic structures that can capture different forms of provenance. Each algorithm yields a different trade-off between time complexity and generality, as each requires different properties over the semiring. Together, these algorithms cover a large class of semirings used for provenance (top-k, security, etc.). Experimental results suggest these approaches are complementary and practical for various kinds of provenance indications, even on a relatively large transport network.

In [16], we showcase ProvSQL, an open-source module for the PostgreSQL database management system that adds support for computation of provenance and probabilities of query results. A large range of provenance formalisms are supported, including all those captured by provenance semirings, provenance semirings with monus, as well as where-provenance. Probabilistic query evaluation is made possible through the use of knowledge compilation tools, in addition to standard approaches such as enumeration of possible worlds and Monte-Carlo sampling. ProvSQL supports a large subset of non-aggregate SQL queries.

Finally, in [20], [35], we focus on knowledge compilation, which can be used to obtain compact circuit-based representations of (Boolean) provenance. Some width parameters of the circuit, such as bounded treewidth or pathwidth, can be leveraged to convert the circuit to structured classes, e.g., deterministic structured NNFs (d-SDNNFs) or OBDDs. We show how to connect the width of circuits to the size of their structured representation, through upper and lower bounds. For the upper bound, we show how bounded-treewidth circuits can be converted to a d-SDNNF, in time linear in the circuit size. Our bound, unlike existing results, is constructive and only singly exponential in the treewidth. We show a related lower bound on monotone DNF or CNF formulas, assuming a constant bound on the arity (size of clauses) and degree (number of

occurrences of each variable). Specifically, any d-SDNNF (resp., SDNNF) for such a DNF (resp., CNF) must be of exponential size in its treewidth; and the same holds for pathwidth when compiling to OBDDs. Our lower bounds, in contrast with most previous work, apply to any formula of this class, not just a well-chosen family. Hence, for our language of DNF and CNF, pathwidth and treewidth respectively characterize the efficiency of compiling to OBDDs and (d-)SDNNFs, that is, compilation is singly exponential in the width parameter.

6.3. Exploiting Content from the Web

One of our main domain of application is that of Web content. We investigate methods to acquire and exploit content from the Web.

In [30], we analyze form-based websites to discover sequences of actions and values that result in a valid form submission. Rather than looking at the text or DOM structure of the form, our method is driven by solving constraints involving the underlying client-side JavaScript code. In order to deal with the complexity of client-side code, we adapt a method from program analysis and testing, concolic testing, which mixes concrete code execution, symbolic code tracing, and constraint solving to find values that lead to new code paths. While concolic testing is commonly used for detecting bugs in stand-alone code with developer support, we show how it can be applied to the very different problem of filling Web forms. We evaluate our system on a benchmark of both real and synthetic Web forms.

In [21], we investigate *focused crawling*: collecting as many Web pages relevant to a target topic as possible while avoiding irrelevant pages, reflecting limited resources available to a Web crawler. We improve on the efficiency of focused crawling by proposing an approach based on reinforcement learning. Our algorithm evaluates hyperlinks most profitable to follow over the long run, and selects the most promising link based on this estimation. To properly model the crawling environment as a Markov decision process, we propose new representations of states and actions considering both content information and the link structure. The size of the state-action space is reduced by a generalization process. Based on this generalization, we use a linear-function approximation to update value functions. We investigate the trade-off between synchronous and asynchronous methods. In experiments, we compare the performance of a crawling task with and without learning; crawlers based on reinforcement learning show better performance for various target topics.

Finally, in [23], [24] we propose a framework to follow the dynamics of vanished Web communities, based on the exploration of corpora of Web archives. To achieve this goal, we define a new unit of analysis called Web fragment: a semantic and syntactic subset of a given Web page, designed to increase historical accuracy. This contribution has practical value for those who conduct large-scale archive exploration (in terms of time range and volume) or are interested in computational approaches to Web history and social science.

6.4. Knowledge Bases

Knowledge bases are collection of semantic facts (typically of the form subject–predicate–object) along with possible logical rules (e.g., in the form of existential rules) that apply to these facts. We investigate querying, data integration, and inference in such knowledge bases.

In [27], we focus on autocompletion of SPARQL queries over knowledge bases. We analyze several autocompletion features proposed by the main editors, highlighting the needs currently not taken into account while met by a user community we work with, scientists. Second, we introduce the first (to our knowledge) autocompletion approach able to consider snippets (fragments of SPARQL query) based on queries expressed by previous users, enriching the user experience. Third, we introduce a usable, open and concrete solution able to consider a large panel of SPARQL autocompletion features that we have implemented in an editor. Last but not least, we demonstrate the interest of our approach on real biomedical queries involving services offered by the Wikidata collaborative knowledge base.

In [25], we introduce a novel open-source framework for integrating the data of a user from different sources into a single knowledge base. Our framework integrates data of different kinds into a coherent whole, starting with email messages, calendar, contacts, and location history. We show how event periods in the user's location data can be detected and how they can be aligned with events from the calendar. This allows users to query their personal information within and across different dimensions, and to perform analytics over their emails, events, and locations. Our system models data using RDF, extending the schema.org vocabulary and providing a SPARQL interface.

Finally, in [22], [32], we view knowledge bases as composed of an instance that contains incomplete data and a set of existential rules, and investigate ontology-based query answering: answers to queries are logically entailed from the knowledge base. This brings to light the fundamental chase tool, and its different variants that have been proposed in the literature. It is well-known that the problem of determining, given a chase variant and a set of existential rules, whether the chase will halt on a given instance / on any instance, is undecidable. Hence, a crucial issue is whether it becomes decidable for known subclasses of existential rules. We consider linear existential rules, a simple yet important subclass of existential rules. We study the decidability of the associated chase termination problem for different chase variants, with a novel approach based on a single graph and a single notion of forbidden pattern. Besides the theoretical interest of a unified approach, an original result is the decidability of the restricted chase termination for linear existential rules.

6.5. Transparency and Bias

In this last set of results, we investigate transparency and bias in data management.

Bias in online information has recently become a pressing issue, with search engines, social networks and recommendation services being accused of exhibiting some form of bias. In [15], we make the case for a systematic approach towards measuring bias. To this end, we discuss formal measures for quantifying the various types of bias, we outline the system components necessary for realizing them, and we highlight the related research challenges and open problems.

In [19], we pursue an investigation of data-driven collaborative work-flows. In the model, peers can access and update local data, causing side-effects on other peers' data. In this paper, we study means of explaining to a peer her local view of a global run, both at runtime and statically. We consider the notion of "scenario for a given peer" that is a subrun observationally equivalent to the original run for that peer. Because such a scenario can sometimes differ significantly from what happens in the actual run, thus providing a misleading explanation , we introduce and study a faithfulness requirement that ensures closer adherence to the global run. We show that there is a unique minimal faithful scenario, that explains what is happening in the global run by extracting only the portion relevant to the peer. With regard to static explanations, we consider the problem of synthesizing, for each peer, a "view program" whose runs generate exactly the peer's observations of the global runs. Assuming some conditions desirable in their own right, namely transparency and boundedness, we show that such a view program exists and can be synthesized. As an added benefit, the view program rules provide provenance information for the updates observed by the peer.

Finally, in two articles oriented towards applications and policy, we discuss bias and neutrality and their impact on regulation. In [18] we discuss the different forms of neutrality in the digital world, from the neutrality of networks to neutrality of content. In [17], we investigate the impact of bias and neutrality concerns on algorithms used by businesses.

WILLOW Project-Team

7. New Results

7.1. 3D object and scene modeling, analysis, and retrieval

7.1.1. Indoor Visual Localization with Dense Matching and View Synthesis

Participants: Hajime Taira, Masatoshi Okutomi, Torsten Sattler, Mircea Cimpoi, Marc Pollefeys, Josef Sivic, Tomas Pajdla, Akihiko Torii.

In [20], we seek to predict the 6 degree-of-freedom (6DoF) pose of a query photograph with respect to a large indoor 3D map. The contributions of this work are three-fold. First, we develop a new large-scale visual localization method targeted for indoor environments. The method proceeds along three steps: (i) efficient retrieval of candidate poses that ensures scalability to large-scale environments, (ii) pose estimation using dense matching rather than local features to deal with textureless indoor scenes, and (iii) pose verification by virtual view synthesis to cope with significant changes in viewpoint, scene layout, and occluders. Second, we collect a new dataset with reference 6DoF poses for large-scale indoor localization. Query photographs are captured by mobile phones at a different time than the reference 3D map, thus presenting a realistic indoor localization scenario. Third, we demonstrate that our method significantly outperforms current state-of-the-art indoor localization approaches on this new challenging data. Figure 1 presents some example results.

7.1.2. Benchmarking 6DOF Outdoor Visual Localization in Changing Conditions

Participants: Torsten Sattler, Will Maddern, Carl Toft, Akihiko Torii, Lars Hammarstrand, Erik Stenborg, Daniel Safari, Masatoshi Okutomi, Marc Pollefeys, Josef Sivic, Frederik Kahl, Tomas Pajdla.

Visual localization enables autonomous vehicles to navigate in their surroundings and augmented reality applications to link virtual to real worlds. Practical visual localization approaches need to be robust to a wide variety of viewing condition, including day-night changes, as well as weather and seasonal variations, while providing highly accurate 6 degree-of-freedom (6DOF) camera pose estimates. In [19], we introduce the first benchmark datasets specifically designed for analyzing the impact of such factors on visual localization. Using carefully created ground truth poses for query images taken under a wide variety of conditions, we evaluate the impact of various factors on 6DOF camera pose estimation accuracy through extensive experiments with state-of-the-art localization approaches. Based on our results, we draw conclusions about the difficulty of different conditions, showing that long-term localization is far from solved, and propose promising avenues for future work, including sequence-based localization approaches and the need for better local features. Our benchmark is available at visuallocalization.net. Figure 2 presents some example results.

7.1.3. Changing Views on Curves and Surfaces

Participants: Kathlen Kohn, Bernd Sturmfels, Matthew Trager, Boris Bukh, Xavier Goaoc, Alfredo Hubard, Matthew Trager.

Visual events in computer vision are studied from the perspective of algebraic geometry. Given a sufficiently general curve or surface in 3-space, we consider the image or contour curve that arises by projecting from a viewpoint. Qualitative changes in that curve occur when the viewpoint crosses the visual event surface as illustrated in 3 . We examine the components of this ruled surface, and observe that these coincide with the iterated singular loci of the coisotropic hypersurfaces associated with the original curve or surface. We derive formulas, due to Salmon and Petitjean, for the degrees of these surfaces, and show how to compute exact representations for all visual event surfaces using algebraic methods. This work has been published in [8].

subsectionConsistent Sets of Lines with no Colorful Incidence

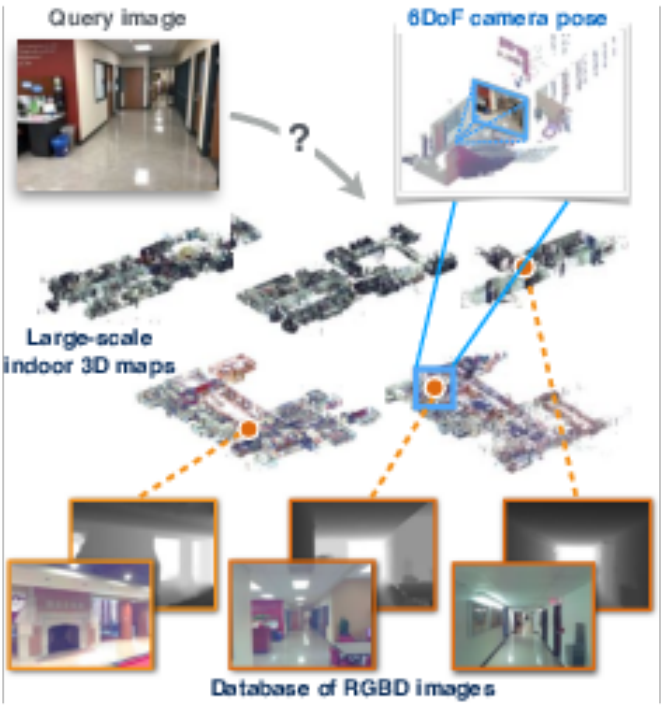

Figure 1. Large-scale indoor visual localization. Given a database of geometrically-registered RGBD images, we predict the 6DoF camera pose of a query RGB image by retrieving candidate images, estimating candicate camera poses, and selecting the best matching camera pose. To address inherent difficulties in indoor visual localization, we introduce the ?InLoc? approach that performs a sequence of progressively stricter verification steps.

Figure 2. Visual localization in changing urban conditions. We present three new datasets, Aachen Day-Night, RobotCar Seasons (shown) and CMU Seasons for evaluating 6DOF localization against a prior 3D map (top) using registered query images taken from a wide variety of conditions (bottom), including day-night variation, weather, and seasonal changes over long periods of time.

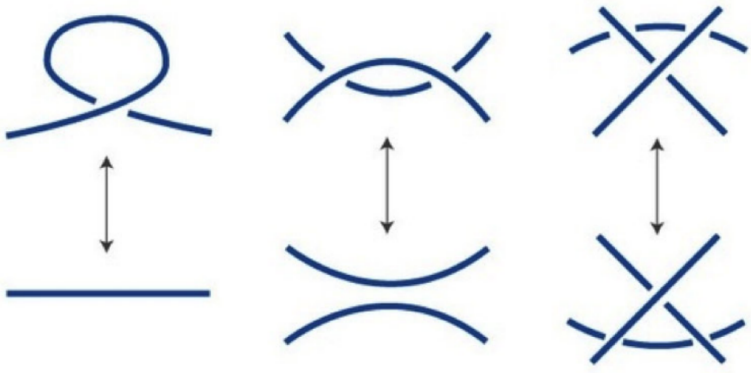

Figure 3. Changing views of a curve correspond to Reidemeister moves. The viewpoint z crosses the tangential surface (left), edge surface (middle), or trisecant surface (right).

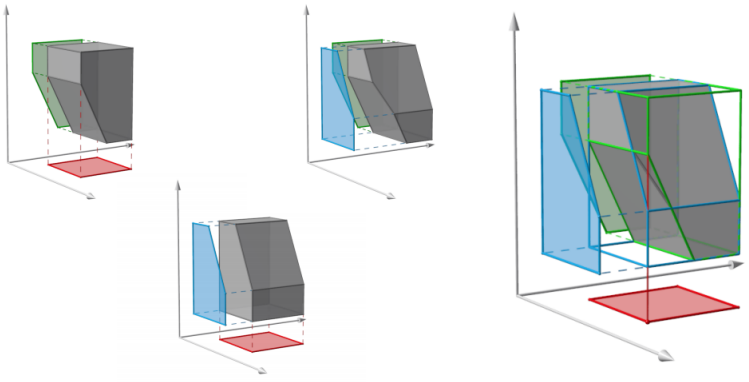

Figure 4. Three silhouettes that are 2-consistent but not globally consistent for three orthogonal projections. Each of the first three figures shows a three-dimensional set that projects onto two of the three silhouettes. The fourth figure illustrates that no set can project simultaneously onto all three silhouettes: the highlighted red image point cannot be lifted in 3D, since no point that projects onto it belongs to the pre-images of both the blue and green silhouettes.

We consider incidences among colored sets of lines in $\mathbb{R}^d$ and examine whether the existence of certain concurrences between lines of k colors force the existence of at least one concurrence between lines of $k+1$ colors. This question is relevant for problems in 3D reconstruction in computer vision such as the one illustrated in Figure 4 . This work has been published in [12].

7.1.4. On the Solvability of Viewing Graphs

Participants: Matthew Trager, Brian Osserman, Jean Ponce.

A set of fundamental matrices relating pairs of cameras in some configuration can be represented as edges of a ” viewing graph ". Whether or not these fundamental matrices are generically sufficient to recover the global camera configuration depends on the structure of this graph. We study characterizations of ” solvable " viewing graphs, and present several new results that can be applied to determine which pairs of views may be used to recover all camera parameters. We also discuss strategies for verifying the solvability of a graph computationally. This work has been published in [21].

7.1.5. In Defense of Relative Multi-View Geometry

Participants: Matthew Trager, Jean Ponce.

The idea of studying multi-view geometry and structure-from-motion problems *relative* to the scene and camera configurations, without appeal to external coordinate systems, dates back to the early days of modern geometric computer vision. Yet, it has a bad rap, the scene reconstructions obtained often being deemed as inaccurate despite careful implementations. The aim of this article is to correct this perception with a series of new results. In particular, we show that using a small subset of scene and image points to parameterize their relative configurations offers a natural coordinate-free formulation of Carlsson-Weinshall duality for arbitrary numbers of images. An example is shown in Figure 5 . For three views, this approach also yields novel purely- and quasi-linear formulations of structure from motion using *reduced trilinearities*, without the complex polynomial constraints associated with trifocal tensors, revealing in passing the strong link between “3D” ($\mathbb{P}^3 \to \mathbb{P}^2$) and “2D” ($\mathbb{P}^2 \to \mathbb{P}^1$) models of trinocular vision. Finally, we demonstrate through preliminary experiments that the proposed relative reconstruction methods gives good results on real data. This works is available as a preprint [32].

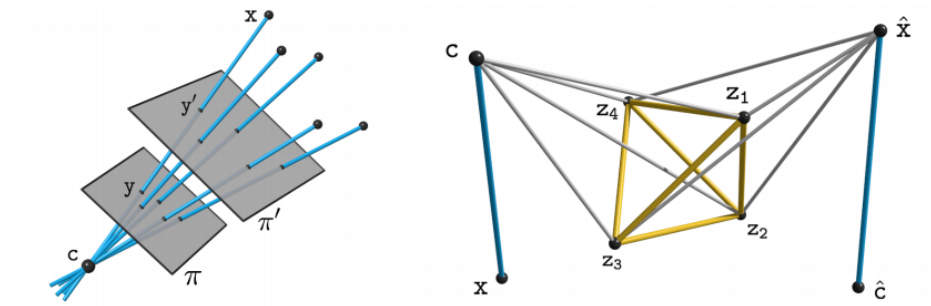

Figure 5. Configurations. ***Left****: Image point and viewing ray configurations are isomorphic and independent of the retinal plane.* ***Right****: Geometric Carlsson-Weinshall duality between scene point and pinhole configurations.*

7.1.6. Multigraded Cayley-Chow Forms

Participants: Brian Osserman, Matthew Trager.

We introduce a theory of multigraded Cayley-Chow forms associated to subvarieties of products of projective spaces. Figure 6 illustrares some examples of projective speces. Two new phenomena arise: first, the construction turns out to require certain inequalities on the dimensions of projections; and second, in positive characteristic the multigraded Cayley-Chow forms can have higher multiplicities. The theory also provides a natural framework for understanding multifocal tensors in computer vision. This works is available as a preprint [30].

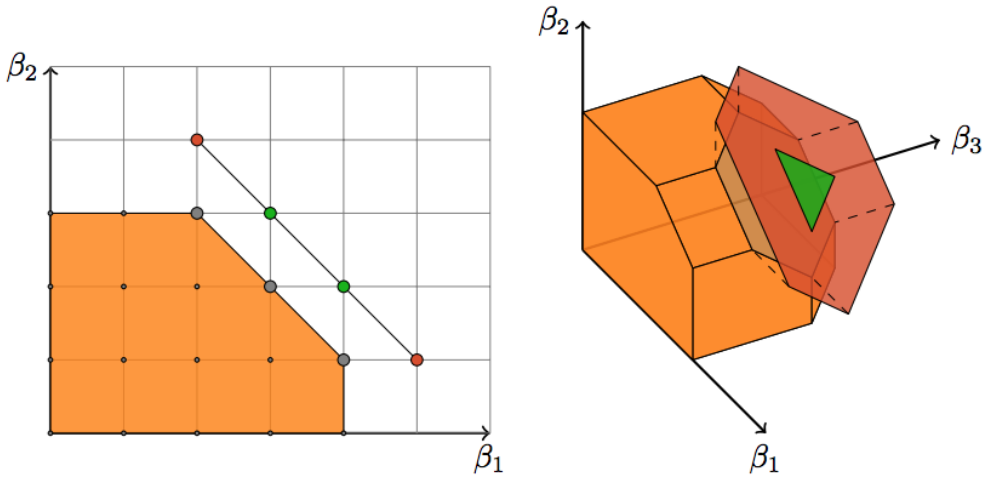

Figure 6. Two polymatroids. The sets of bases (corresponding to our multidegree supports) are in gray; while the sets of circuits and of non-circuit 1-deficient vectors are in green and red, respectively.

7.2. Category-level object and scene recognition

7.2.1. Detecting rare visual relations using analogies

Participants: Julia Peyre, Cordelia Schmid, Ivan Laptev, Josef Sivic.

We seek to detect visual relations in images of the form of triplets t = (subject, predicate, object), such as "person riding dog", where training examples of the individual entities are available but their combinations are rare or unseen at training such as shown in Figure 7 . This is an important set-up due to the combinatorial nature of visual relations : collecting sufficient training data for all possible triplets would be very hard. The contributions of this work are three-fold. First, we learn a representation of visual relations that combines (i) individual embeddings for subject, object and predicate together with (ii) a visual phrase embedding that represents the relation triplet. Second, we learn how to transfer visual phrase embeddings from existing training triplets to unseen test triplets using analogies between relations that involve similar objects. Third, we demonstrate the benefits of our approach on two challenging datasets involving rare and unseen relations : on HICO-DET, our model achieves significant improvement over a strong baseline, and we confirm this improvement on retrieval of unseen triplets on the UnRel rare relation dataset. This work, currently under review, can be found at [31].

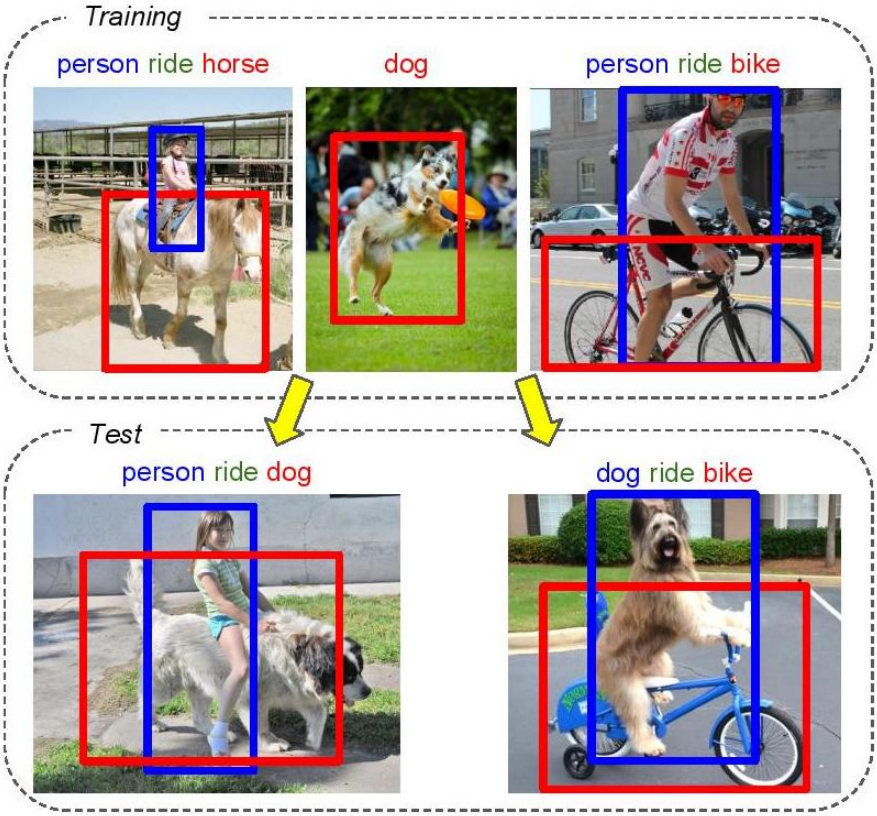

Figure 7. Illustration of transfer by analogy from seen training triplets (e.g. "person ride horse") to unseen or rare ones (e.g. "person ride dog"))

7.2.2. Convolutional neural network architecture for geometric matching

Participants: Ignacio Rocco, Relja Arandjelović, Josef Sivic.

In [9], we address the problem of determining correspondences between two images in agreement with a geometric model such as an affine, homography or thin-plate spline transformation, and estimating its parameters. The contributions of this work are threefold. First, we propose a convolutional neural network architecture for geometric matching. The architecture is based on three main components that mimic the standard steps of feature extraction, matching and simultaneous inlier detection and model parameter estimation, while being trainable end-to-end. Second, we demonstrate that the network parameters can be trained from synthetically generated imagery without the need for manual annotation and that our matching layer significantly increases generalization capabilities to never seen before images. Finally, we show that the same model can perform both instance-level and category-level matching giving state-of-the-art results on the challenging PF, TSS and Caltech-101 datasets.

7.2.3. End-to-end weakly-supervised semantic alignment

Participants: Ignacio Rocco, Relja Arandjelović, Josef Sivic.

In [17], we tackle the task of semantic alignment where the goal is to compute dense semantic correspondence aligning two images depicting objects of the same category. This is a challenging task due to large intra-class variation, changes in viewpoint and background clutter. We present the following three principal contributions. First, we develop a convolutional neural network architecture for semantic alignment that is trainable in an end-to-end manner from weak image-level supervision in the form of matching image pairs. The outcome is that parameters are learnt from rich appearance variation present in different but semantically related images without the need for tedious manual annotation of correspondences at training time. Second, the main component of this architecture is a differentiable soft inlier scoring module, inspired by the RANSAC inlier scoring procedure, that computes the quality of the alignment based on only geometrically consistent correspondences thereby reducing the effect of background clutter. Third, we demonstrate that the proposed approach achieves state-of-the-art performance on multiple standard benchmarks for semantic alignment. Figure 8 presents some example results.

7.2.4. Neighbourhood Consensus Networks

Participants: Ignacio Rocco, Mircea Cimpoi, Relja Arandjelović, Akihiko Torii, Tomas Pajdla, Josef Sivic.

In [18], we address the problem of finding reliable dense correspondences between a pair of images. This is a challenging task due to strong appearance differences between the corresponding scene elements and ambiguities generated by repetitive patterns. The contributions of this work are threefold. First, inspired by the classic idea of disambiguating feature matches using semi-local constraints, we develop an end-to-end trainable convolutional neural network architecture that identifies sets of spatially consistent matches by analyzing neighbourhood consensus patterns in the 4D space of all possible correspondences between a pair of images without the need for a global geometric model. Second, we demonstrate that the model can be trained effectively from weak supervision in the form of matching and non-matching image pairs without the need for costly manual annotation of point to point correspondences. Third, we show the proposed neighbourhood consensus network can be applied to a range of matching tasks including both category- and instance-level matching, obtaining the state-of-the-art results on the PF Pascal dataset and the InLoc indoor visual localization benchmark. Figure 9 shows the network architecture of the proposed Neighbourhood Consensus Network, that features 3 layers of 4D convolutions.

7.2.5. Compressing the Input for CNNs with the First-Order Scattering Transform

Participants: Edouard Oyallon, Eugene Belilovsky, Sergey Zagoruyko, Michal Valko.

In [16], we study the first-order scattering transform as a candidate for reducing the signal processed by a convolutional neural network (CNN). We study this transformation and show theoretical and empirical evidence that in the case of natural images and sufficiently small translation invariance, this transform preserves most of the signal information needed for classification while substantially reducing the spatial resolution and total signal size. We show that cascading a CNN with this representation performs on par with ImageNet classification models commonly used in downstream tasks such as the ResNet-50. We subsequently apply our trained hybrid ImageNet model as a base model on a detection system, which has typically larger

Figure 8. Each row corresponds to one example and shows the (right) automatic semantic alignment of the (left) source and (middle) target images.

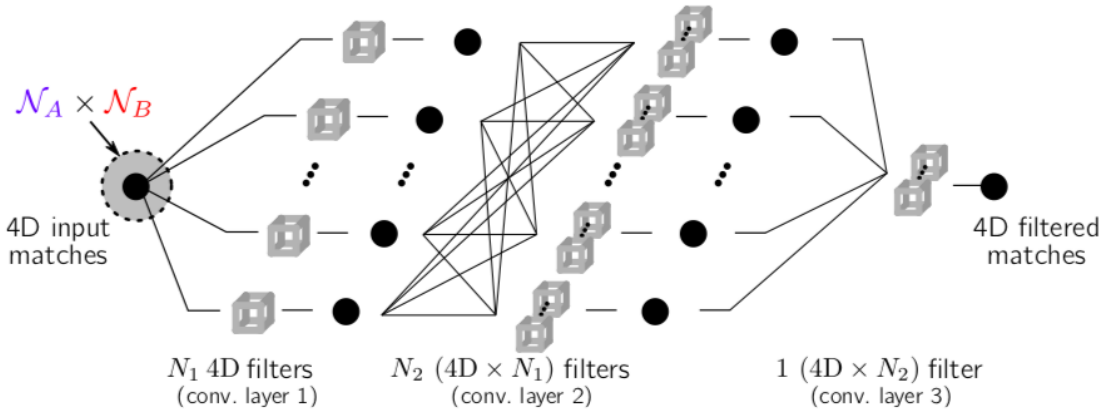

Figure 9. A neighbourhood consensus CNN operates on the 4D space of feature matches. The first 4D convolutional layer filters span $\mathcal{N}_A \times \mathcal{N}_B$, the Cartesian product of local neighbourhoods $\mathcal{N}_A$ and $\mathcal{N}_B$ in images A and B respectively. The proposed 4D neighbourhood consensus CNN can learn to identify the matching patterns of reliable and unreliable matches, and filter the matches accordingly

image inputs. On Pascal VOC and COCO detection tasks we deliver substantial improvements in the inference speed and training memory consumption compared to models trained directly on the input image.

7.2.6. Exploring Weight Symmetry in Deep Neural Networks

Participants: Xu Shell Hu, Sergey Zagoruyko, Nikos Komodakis.

In [27], we propose to impose symmetry in neural network parameters to improve parameter usage and make use of dedicated convolution and matrix multiplication routines. Due to significant reduction in the number of parameters as a result of the symmetry constraints, one would expect a dramatic drop in accuracy. Surprisingly, we show that this is not the case, and, depending on network size, symmetry can have little or no negative effect on network accuracy, especially in deep overparameterized networks. We propose several ways to impose local symmetry in recurrent and convolutional neural networks, and show that our symmetry parameterizations satisfy universal approximation property for single hidden layer networks. We extensively evaluate these parameterizations on CIFAR, ImageNet and language modeling datasets, showing significant benefits from the use of symmetry. For instance, our ResNet-101 with channel-wise symmetry has almost 25% less parameters and only 0.2% accuracy loss on ImageNet.

7.3. Image restoration, manipulation and enhancement

7.3.1. Neural Embedding of an Iterative Deconvolution Algorithm for Motion Blur Estimation and Removal

Participants: Thomas Eboli, Jian Sun, Jean Ponce.

We introduce a new two-steps learning-based approach to motion blur estimation and removal decomposed into two trainable modules. A local linear motion model is estimated at each pixel using a first convolutional neural network (CNN) in a regression setting. It is then used to drive an algorithm that casts non-blind, non-uniform image deblurring as a least-squares problem regularized by natural image priors in the form of sparsity constraints. This problem is solved by combining the alternative direction method of multipliers with an iterative residual compensation algorithm, with a finite number of iterations embedded into a second CNN whose trainable parameters are deconvolution filters. The second network outputs the sharp image, and the

two CNNs can be trained together in an end-to-end manner. Our experiments demonstrate that the proposed method is significantly faster than existing ones, and provides competitive results with the state of the art on synthetic and real data. This work is available as a pre-print[25] and an example is illustrated in Figure 10 .

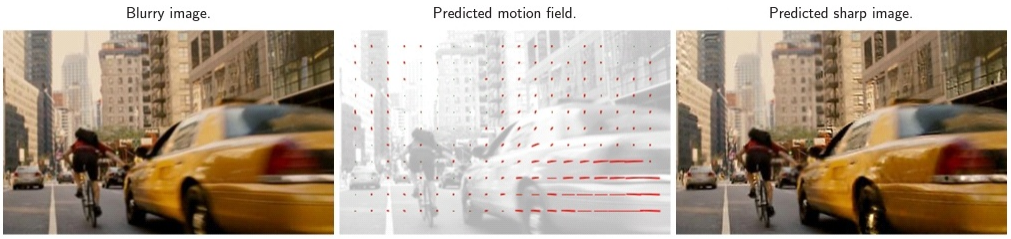

Figure 10. From a blurry image, we first use CNN-based regressor to predict a motion field with local linear motions before using it in a trainable iterative residual compensation algorithm to restore the image.

7.3.2. Deformable Kernel Networks for Joint Image Filtering

Participants: Beomjun Kim, Jean Ponce, Bumsub Ham.

Joint image filters are used to transfer structural details from a guidance picture used as a prior to a target image, in tasks such as enhancing spatial resolution and suppressing noise. Previous methods based on convolutional neural networks (CNNs) combine nonlinear activations of spatially-invariant kernels to estimate structural details and regress the filtering result. In this paper, we instead learn explicitly sparse and spatially-variant kernels. We propose a CNN architecture and its efficient implementation, called the deformable kernel network (DKN), that outputs sets of neighbors and the corresponding weights adaptively for each pixel. The filtering result is then computed as a weighted average. We also propose a fast version of DKN that runs about four times faster for an image of size 640 by 480. We demonstrate the effectiveness and flexibility of our models on the tasks of depth map upsampling, saliency map upsampling, cross-modality image restoration, texture removal, and semantic segmentation. In particular, we show that the weighted averaging process with sparsely sampled 3 by 3 kernels outperforms the state of the art by a significant margin. This works has been submitted to the IEEE Trans. on Pattern Analysis and Machine Intelligence and is available as a pre-print [28].

7.4. Human activity capture and classification

7.4.1. Learning a Text-Video Embedding from Incomplete and Heterogeneous Data

Participants: Antoine Miech, Ivan Laptev, Josef Sivic.

Joint understanding of video and language is an active research area with many applications. Prior work in this domain typically relies on learning text-video embeddings. One difficulty with this approach, however, is the lack of large-scale annotated video-caption datasets for training. To address this issue, in [29] we aim at learning text-video embeddings from heterogeneous data sources. To this end, we propose a Mixture-of-Embedding-Experts (MEE) model with ability to handle missing input modalities during training. As a result, our framework can learn improved text-video embeddings simultaneously from image and video datasets. We also show the generalization of MEE to other input modalities such as face descriptors. We evaluate our method on the task of video retrieval and report results for the MPII Movie Description and MSR-VTT datasets. The proposed MEE model demonstrates significant improvements and outperforms previously reported methods on both text-to-video and video-to-text retrieval tasks. Figure 11 illustrates application of our method in text-to-video retrieval.

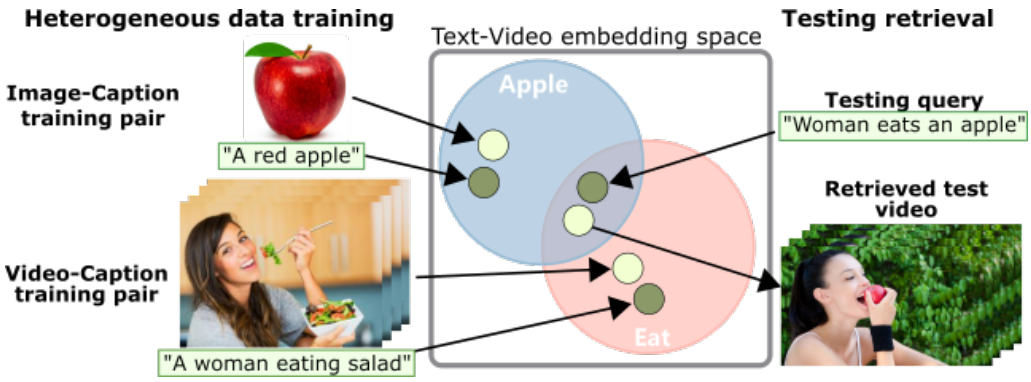

Figure 11. We learn a text-video embedding from heterogenous (here Image-Text and Video-Text) data sources. At test time, we can query concepts learnt from both Image-Caption and Video-Caption training pair (e.g. the eating notion being learnt from video and the apple notion from image).

7.4.2. A flexible model for training action localization with varying levels of supervision

Participants: Guilhem Chéron, Jean-Baptiste Alayrac, Ivan Laptev, Cordelia Schmid.

Spatio-temporal action detection in videos is typically addressed in a fully-supervised setup with manual annotation of training videos required at every frame. Since such annotation is extremely tedious and prohibits scalability, there is a clear need to minimize the amount of manual supervision. In this work we propose a unifying framework that can handle and combine varying types of less-demanding weak supervision. Our model is based on discriminative clustering and integrates different types of supervision as constraints on the optimization as illustrated in Figure 12 . We investigate applications of such a model to training setups with alternative supervisory signals ranging from video-level class labels to the full per-frame annotation of action bounding boxes. Experiments on the challenging UCF101-24 and DALY datasets demonstrate competitive performance of our method at a fraction of supervision used by previous methods. The flexibility of our model enables joint learning from data with different levels of annotation. Experimental results demonstrate a significant gain by adding a few fully supervised examples to otherwise weakly labeled videos. This work has been published in [14].

7.4.3. BodyNet: Volumetric Inference of 3D Human Body Shapes

Participants: Gül Varol, Duygu Ceylan, Bryan Russell, Jimei Yang, Ersin Yumer, Ivan Laptev, Cordelia Schmid.

Human shape estimation is an important task for video editing, animation and fashion industry. Predicting 3D human body shape from natural images, however, is highly challenging due to factors such as variation in human bodies, clothing and viewpoint. Prior methods addressing this problem typically attempt to fit parametric body models with certain priors on pose and shape. In this work we argue for an alternative representation and propose BodyNet, a neural network for direct inference of volumetric body shape from a single image. BodyNet is an end-to-end trainable network that benefits from (i) a volumetric 3D loss, (ii) a multi-view re-projection loss, and (iii) intermediate supervision of 2D pose, 2D body part segmentation, and 3D pose. Each of them results in performance improvement as demonstrated by our experiments. To evaluate the method, we fit the SMPL model to our network output and show state-of-the-art results on the SURREAL and Unite the People datasets, outperforming recent approaches. Besides achieving state-of-the-art performance, our method also enables volumetric body-part segmentation. Figure 13 illustrates the volumetric outputs given two sample input images. This work has been published at ECCV 2018 [22].

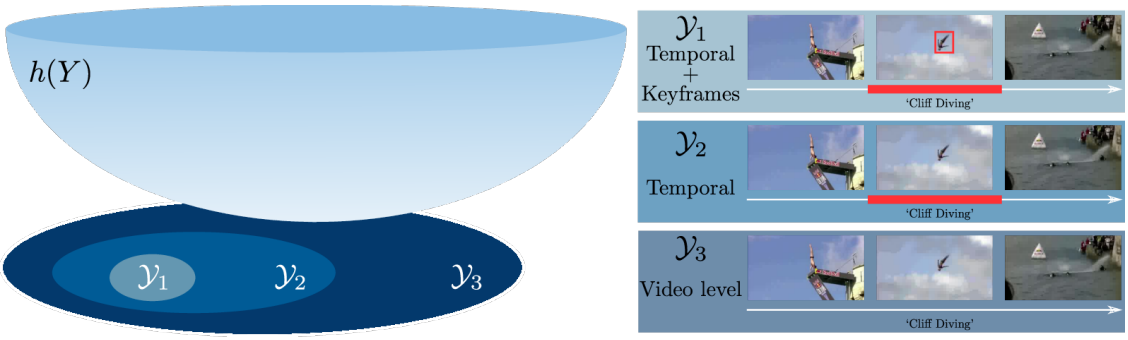

Figure 12. Our method estimates a matrix Y assigning human tracklets to action labels in training videos by optimizing an objective function $h(Y)$ under constraints $\mathcal{Y}_s$. Different types of supervision define particular constraints $\mathcal{Y}_s$ and do not affect the form of the objective function. The increasing level of supervision imposes stricter constraints, e.g. $\mathcal{Y}_1 \supset \mathcal{Y}_2 \supset \mathcal{Y}_3 \supset \mathcal{Y}_4$ as illustrated for the Cliff Diving example above.

Figure 13. Our BodyNet predicts a volumetric 3D human body shape and 3D body parts from a single image. We show the input image, the predicted human voxels, and the predicted part voxels.

7.4.4. Localizing Moments in Video with Temporal Language

Participants: Lisa Anne Hendricks, Oliver Wang, Eli Schechtman, Josef Sivic, Trevor Darrell, Bryan Russell.

Localizing moments in a longer video via natural language queries is a new, challenging task at the intersection of language and video understanding. Though moment localization with natural language is similar to other language and vision tasks like natural language object retrieval in images, moment localization offers an interesting opportunity to model temporal dependencies and reasoning in text. In [15], we propose a new model that explicitly reasons about different temporal segments in a video, and shows that temporal context is important for localizing phrases which include temporal language. To benchmark whether our model, and other recent video localization models, can effectively reason about temporal language, we collect the novel TEMPO-ral reasoning in video and language (TEMPO) dataset. Our dataset consists of two parts: a dataset with real videos and template sentences (TEMPO - Template Language) which allows for controlled studies on temporal language, and a human language dataset which consists of temporal sentences annotated by humans (TEMPO - Human Language).

7.4.5. The Pinocchio C++ library ? A fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives

Participants: Justin Carpentier, Guilhem Saurel, Gabriele Buondonno, Joseph Mirabel, Florent Lamiraux, Olivier Stasse, Nicolas Mansard.

In this work, we introduce Pinocchio, an open-source software framework that implements rigid body dynamics algorithms and their analytical derivatives. Pinocchio does not only include standard algorithms employed in robotics (e.g., forward and inverse dynamics) but provides additional features essential for the control, the planning and the simulation of robots. In this paper, we describe these features and detail the programming patterns and design which make Pinocchio efficient. We evaluate the performances against RBDL, another framework with broad dissemination inside the robotics community. We also demonstrate how the source code generation embedded in Pinocchio outperforms other approaches of state of the art.

7.4.6. Modeling Spatio-Temporal Human Track Structure for Action Localization

Participants: Guilhem Chéron, Anton Osokin, Ivan Laptev, Cordelia Schmid.

This paper [24] addresses spatio-temporal localization of human actions in video. In order to localize actions in time, we propose a recurrent localization network (RecLNet) designed to model the temporal structure of actions on the level of person tracks. Our model is trained to simultaneously recognize and localize action classes in time and is based on two layer gated recurrent units (GRU) applied separately to two streams, i.e. appearance and optical flow streams. When used together with state-of-the-art person detection and tracking, our model is shown to improve substantially spatio-temporal action localization in videos. The gain is shown to be mainly due to improved temporal localization as illustrated in Figure 14 . We evaluate our method on two recent datasets for spatio-temporal action localization, UCF101-24 and DALY, demonstrating a significant improvement of the state of the art.

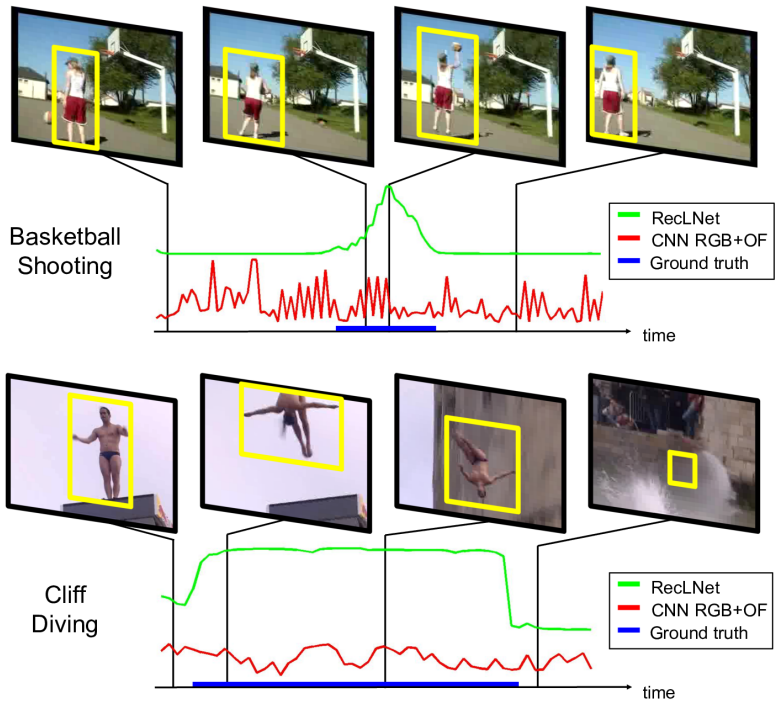

Figure 14. Spatio-temporal action localization using a CNN baseline (red) and our RecLNet (green) both applied on the level of person tracks. Our approach provides accurate temporal boundaries when the action happens.