RESEARCH CENTER
Sophia Antipolis - Méditerranée

FIELD

Activity Report 2019

Section Scientific Foundations

AROMATH Project-Team

3. Research Program

3.1. High order geometric modeling

The accurate description of shapes is a long standing problem in mathematics, with an important impact in many domains, inducing strong interactions between geometry and computation. Developing precise geometric modeling techniques is a critical issue in CAD-CAM. Constructing accurate models, that can be exploited in geometric applications, from digital data produced by cameras, laser scanners, observations or simulations is also a major issue in geometry processing. A main challenge is to construct models that can capture the geometry of complex shapes, using few parameters while being precise.

Our first objective is to develop methods, which are able to describe accurately and in an efficient way, objects or phenomena of geometric nature, using algebraic representations.

The approach followed in CAGD, to describe complex geometry is based on parametric representations called NURBS (Non Uniform Rational B-Spline). The models are constructed by trimming and gluing together high order patches of algebraic surfaces. These models are built from the so-called B-Spline functions that encode a piecewise algebraic function with a prescribed regularity at knots. Although these models have many advantages and have become the standard for designing nowadays CAD models, they also have important drawbacks. Among them, the difficulty to locally refine a NURBS surface and also the topological rigidity of NURBS patches that imposes to use many such patches with trims for designing complex models, with the consequence of the appearing of cracks at the seams. To overcome these difficulties, an active area of research is to look for new blending functions for the representation of CAD models. Some examples are the so-called T-Splines, LR-Spline blending functions, or hierarchical splines, that have been recently devised in order to perform efficiently local refinement. An important problem is to analyze spline spaces associated to general subdivisions, which is of particular interest in higher order Finite Element Methods. Another challenge in geometric modeling is the efficient representation and/or reconstruction of complex objects, and the description of computational domains in numerical simulation. To construct models that can represent efficiently the geometry of complex shapes, we are interested in developing modeling methods, based on alternative constructions such as skeleton-based representations. The change of representation, in particular between parametric and implicit representations, is of particular interest in geometric computations and in its applications in CAGD.

We also plan to investigate adaptive hierarchical techniques, which can locally improve the approximation of a shape or a function. They shall be exploited to transform digital data produced by cameras, laser scanners, observations or simulations into accurate and structured algebraic models.

The precise and efficient representation of shapes also leads to the problem of extracting and exploiting characteristic properties of shapes such as symmetry, which is very frequent in geometry. Reflecting the symmetry of the intended shape in the representation appears as a natural requirement for visual quality, but also as a possible source of sparsity of the representation. Recognizing, encoding and exploiting symmetry requires new paradigms of representation and further algebraic developments. Algebraic foundations for the exploitation of symmetry in the context of non linear differential and polynomial equations are addressed. The intent is to bring this expertise with symmetry to the geometric models and computations developed by AROMATH.

3.2. Robust algebraic-geometric computation

In many problems, digital data are approximated and cannot just be used as if they were exact. In the context of geometric modeling, polynomial equations appear naturally, as a way to describe constraints between the unknown variables of a problem. *An important challenge is to take into account the input error in order to*

develop robust methods for solving these algebraic constraints. Robustness means that a small perturbation of the input should produce a controlled variation of the output, that is forward stability, when the input-output map is regular. In non-regular cases, robustness also means that the output is an exact solution, or the most coherent solution, of a problem with input data in a given neighborhood, that is backward stability.

Our second long term objective is to develop methods to robustly and efficiently solve algebraic problems that occur in geometric modeling.

Robustness is a major issue in geometric modeling and algebraic computation. Classical methods in computer algebra, based on the paradigm of exact computation, cannot be applied directly in this context. They are not designed for stability against input perturbations. New investigations are needed to develop methods, which integrate this additional dimension of the problem. Several approaches are investigated to tackle these difficulties.

One relies on linearization of algebraic problems based on "elimination of variables" or projection into a space of smaller dimension. Resultant theory provides strong foundation for these methods, connecting the geometric properties of the solutions with explicit linear algebra on polynomial vector spaces, for families of polynomial systems (e.g., homogeneous, multi-homogeneous, sparse). Important progresses have been made in the last two decades to extend this theory to new families of problems with specific geometric properties. Additional advances have been achieved more recently to exploit the syzygies between the input equations. This approach provides matrix based representations, which are particularly powerful for approximate geometric computation on parametrized curves and surfaces. They are tuned to certain classes of problems and an important issue is to detect and analyze degeneracies and to adapt them to these cases.

A more adaptive approach involves linear algebra computation in a hierarchy of polynomial vector spaces. It produces a description of quotient algebra structures, from which the solutions of polynomial systems can be recovered. This family of methods includes Gröbner Basis, which provides general tools for solving polynomial equations. Border Basis is an alternative approach, offering numerically stable methods for solving polynomial equations with approximate coefficients. An important issue is to understand and control the numerical behavior of these methods as well as their complexity and to exploit the structure of the input system.

In order to compute "only" the (real) solutions of a polynomial system in a given domain, duality techniques can also be employed. They consist in analyzing and adding constraints on the space of linear forms which vanish on the polynomial equations. Combined with semi-definite programming techniques, they provide efficient methods to compute the real solutions of algebraic equations or to solve polynomial optimization problems. The main issues are the completness of the approach, their scalability with the degree and dimension and the certification of bounds.

Singular solutions of polynomial systems can be analyzed by computing differentials, which vanish at these points. This leads to efficient deflation techniques, which transform a singular solution of a given problem into a regular solution of the transformed problem. These local methods need to be combined with more global root localisation methods.

Subdivision methods are another type of methods which are interesting for robust geometric computation. They are based on exclusion tests which certify that no solution exists in a domain and inclusion tests, which certify the uniqueness of a solution in a domain. They have shown their strength in addressing many algebraic problems, such as isolating real roots of polynomial equations or computing the topology of algebraic curves and surfaces. The main issues in these approaches is to deal with singularities and degenerate solutions.

<p style="text-align: center; color: red;">**DATASHAPE Project-Team**</p>

3. Research Program

3.1. Algorithmic aspects of topological and geometric data analysis

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

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

3.2. Statistical aspects of topological and geometric data analysis

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

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

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

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

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

3.3. Topological approach for multimodal data processing

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

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

3.4. Experimental research and software development

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

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

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

---

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

<p style="text-align:center; color:red">Kairos Project-Team</p>

3. Research Program

3.1. Cyber-Physical co-modeling

Cyber-Physical System modeling requires joint representation of digital/cyber controllers and natural physics environments. Heterogeneous modeling must then be articulated to support accurate (co-)simulation, (co-)analysis, and (co-)verification. The picture above sketches the overall design framework. It comprises functional requirements, to be met provided surrounding platform guarantees, in a contract approach. All relevant aspects are modeled with proper Domain Specific Languages (DSL), so that constraints can be gathered globally, then analyzed to build a mapping proposal with both a structural aspect (functions allocated to platform resources), but also a behavioral ones, scheduling activities. Mapping may be computed automatically or not, provably correct or not, obtained by static analytic methods or abstract execution. Physical phenomena (in a very broad acceptance of the term) are usually modeled using continuous-time models and differential equations. Then the "proper" discretization opportunities for numerical simulation form a large spectrum of mathematical engineering practices. This is not at all the domain of expertise of Kairos members, but it should not be a limitation as long as one can assume a number of properties from the discretized version. On the other hand, we do have a strong expertise on modeling of both embedded processing architectures and embedded software (i.e., the kind of usually concurrent, sometimes distributed software that reacts to and control the physical environment). This is important as, unlike in the "physical" areas where modeling is common-place, modeling of software and programs is far from mainstream in the Software Engineering community. These domains are also an area of computer science where modeling, and even formal modeling, of the real objects that are originally of discrete/cyber nature, takes some importance with formal Models of Computation and Communications. It seems therefore quite natural to combine physical and cyber modeling in a more global design approach (even multi-physic domains and systems of systems possibly, but always with software-intensive aspects involved). Our objective is certainly not to become experts in physical modeling and/or simulation process, but to retain from it only the essential and important aspects to include them into System-Level Engineering design, based on Model-Driven approaches allowing formal analysis.

This sets an original research agenda: Model-Based System Engineering environments exist, at various stages of maturity and specificity, in the academic and industrial worlds. Formal Methods and Verification/Certification techniques also exist, but generally in a point-wise fashion. Our approach aims at raising the level of formality describing relevant features of existing individual models, so that formal methods can have a greater general impact on usual, "industrial-level", modeling practices. Meanwhile, the relevance of formal methods is enhanced as it now covers various aspects in a uniform setting (timeliness, energy budget, dependability, safety/security...).

New research directions on formal CPS design should focus on the introduction of uncertainty (stochastic models) in our particular framework, on relations between (logical) real-time and security, on relations between common programming languages paradigms and logical time, on extending logical frameworks with logical time, on the concern with resource discovery also in presence of mobility inherent to connected objects and Internet of Things.

3.2. Cyber-Physical co-simulation

The FMI standard (Functional Mock-Up Interface) has been proposed for "purely physical" (i.e., based on persistent signals) co-simulation, and then adopted in over 100 industrial tools including frameworks such as Matlab/Simulink and Ansys, to mention two famous model editors. With the recent use of co-simulation to cyber-physical systems, dealing with the discrete and transient nature of cyber systems became mandatory. Together with other people from the community, we shown that FMI and other frameworks for co-simulation

badly support co-simulation of cyber-physical systems; leading to bad accuracy and performances. More precisely, the way to interact with the different parts of the co-simulation require a specific knowledge about its internal semantics and the kind of data exposed (e.g., continuous, piecewise-constant). Towards a better co-simulation of cyber-physical systems, we are looking for conservative abstractions of the parts and formalisms that aim to describe the functional and temporal constraints that are required to bind several simulation models together.

3.3. Formal analysis and verification

Because the nature of our constraints is specific, we want to adjust verification methods to the goals and expressiveness of our modeling approach. Quantitative (interval) timing conditions on physical models combined with (discrete) cyber modes suggest the use of SMT (Satisfiability Modulo Theories) automatic solvers, but the natural expressiveness requested (as for instance in our CCSL constructs) shows this is not always feasible. Either interactive proofs, or suboptimal solutions (essentially resulting of abstract run-time simulations) should be considered. Complementarily to these approaches, we are experimenting with new variants of symbolic behavioural semantics, allowing to construct finite representations of the behaviour of CPS systems with explicit handling of data, time, or other non-functional aspects.

3.4. Relation between Model and Code

While models considered in Kairos can also be considered as executable specifications (through abstract simulation schemes), they can also lead to code synthesis and deployment. Conversely, code execution of smaller, elementary software components can lead to performance estimation enriching the models before global mapping optimization. CPS introduce new challenging problems for code performance stability. Indeed, two additional factors for performance variability appear, which were not present in classical embedded systems: 1) variable and continuous data input from the physical world and 2) variable underlying hardware platform. For the first factor, CPS software must be analysed in conjunction with its data input coming from the physics, so the variability of the performance may come from the various data. For the second factor, the underlying hardware of the CPS may change during the time (new computing actors appear or disappear, some actors can be reconfigured during execution). The new challenge is to understand how these factors influence performance variability exactly, and how to provide solutions to reduce it or to model it. The modeling of performance variability becomes a new input.

3.5. Code generation and optimization

A significant part CPS design happens at model level, through activities such as model construction, analysis, or verification. However, in most cases the objective of the design process is implementation. We mostly consider the implementation problem in the context of embedded, real-time, or edge computing applications, which are subject to stringent performance, embedding, and safety *non-functional requirements*.

The implementation of such systems usually involves a mix of synthesis—(real-time) scheduling, code generation, compilation—and performance (*e.g.* timing) analysis. One key difficulty here is that synthesis and performance analysis depend on each other. As enumerating the various solutions is not possible for complexity reasons, heuristic implementation methods are needed in all cases. One popular solution here is to build the system first using unsafe performance estimations for its components, and then check system *schedulability* through a global analysis. Another solution is to use safe, over-approximated performance estimations and perform their mapping in a way that ensures by construction the schedulability of the system.

In both cases, the specification of the design space—functional specification, execution platform model, non-functional requirements, implementation model—is a key problem. Another problem is the definition of scalable and efficient mapping methods based on both "exact" approaches (ILP/SMT/CP solving) and compilation-like heuristics.

3.6. Extending logical frameworks with logical time

The Curry-Howard isomorphism (*proposition-as-types and proofs-as-typed-λ-terms*) represent the logical and computational basis to interactive theorem provers: our challenge is to investigate and design time constraints within a Dependent Type Theory (e.g. if event A happened-before event B, then the timestamp/type of A is less (i.e. a subtype) than the timestamp/type of B). We are currently extending the Edinburgh Logical Framework (LF) of Harper-Honsell-Plotkin with relevant constructs expressing logical time and synchronization between processes. Also, union and intersection types with their subtyping constraints theories could capture some constraints expressions *à la* CCSL needed to formalize logical clocks (in particular CCSL expressions like subclock, clock union, intersection and concatenation) and provide opportunities for an *ad hoc* polymorphic timed Type Theory. Logical time constraints seen as property types can beneficially be handled by logical frameworks. The new challenge here is to demonstrate the relevance of Type Theory to work on logical and multiform timing constraint resolution.

3.7. Object-oriented programming and logical time

We formalize in the past object-oriented programming features and safe static type systems featuring delegation-based or trait inheritance: well typed program will never produce into the `message-not-found` infamous run-time error. We view the logical time as a means to enhance the description of timing constraints and properties on top of existing language semantics. When considering general purpose object-oriented languages, like Java, Type Theory is a natural way to provide such properties. Currently, few languages have special types to manage instants, time structures and instant relations like subclocking, precedence, causality, equality, coincidence, exclusion, independence, etc. CCSL provides ad-hoc constructors to specify clock constraints and logical time: enriching object-oriented type theories with CCSL expressions could constitute an interesting research perspective towards a wider usage of CCSL. The new challenge is to consider logical time constraints as behavioral type properties, and the design of programming language constructs and *ad-hoc* type systems. Advances of typed-calculi featuring those static time features will be applied to our extension [42] of the lambda-calculus of objects of Fisher-Honsell-Mitchell.

3.8. Extensions for spatio-temporal modeling and mobile systems

While Time is clearly a primary ingredient in the proper design of CPS systems, in some cases Space, and related notions of local proximity or conversely long distance, play also a key role for correct modeling, often in part because of the constraints this puts on interactions and time for communications. Once space is taken into account, one has to recognize also that many systems will request to consider mobility, originated as change of location through time. Mobile CPS (or mCPS) systems occur casually, e.g., in the case of Intelligent Transportation Systems, or in roaming connected objects of the IoT. Spatio-temporal and mobility modeling may each lead to dynamicity in the representation of constraints, with the creation/deletion/discovering of new components in the system. This opportunity for new expressiveness will certainly cause new needs in handling constraint systems and topological graph locations. The new challenge is to provide an algebraic support with a constraint description language that could be as simple and expressive as possible, and of use in the semantic annotations for mobile CPS design. We also aims to provide fully distributed routing protocols to manage Semantic Resource Discovery in IoT.

PRIVATICS Project-Team (section vide)

<div align="center" style="color:red">**STAMP Project-Team**</div>

3. Research Program

3.1. Theoretical background

The proof assistants that we consider provide both a programming language, where users can describe algorithms performing tasks in their domain of interest, and a logical language to reason about the programs, thus making it possible to ensure that the algorithms do solve the problems for which they were designed. trustability is gained because algorithms and logical statements provide multiple views of the same topic, thus making it possible to detect errors coming from mismatch between expected and established properties. The verification process is itself a logical process, where the computer can bring rigor in aligning expectations and guarantees.

The foundations of proof assistants rest on the very foundations of mathematics. As a consequence, all aspects of reasoning must be made completely explicit in the process of formally verifying an algorithm. All aspects of the formal verification of an algorithm are expressed in a discourse whose consistency is verified by the computer, so that unclear or intuitive arguments need to be replaced by precise logical inferences.

One of the foundational features on which we rely extensively is *Type Theory*. In this approach a very simple programming language is equiped with a powerful discipline to check the consistency of usage: types represent sets of data with similar behavior, functions represent algorithms mapping types to other types, and the consistency can be verified by a simple computer program, a *type-checker*. Although they can be verified by a simple program, types can express arbitrary complex objects or properties, so that the verification work lives in an interesting realm, where verifying proofs is decidable, but finding the proofs is undecidable.

This process for producing new algorithms and theorems is a novelty in the development of mathematical knowledge or algorithms, and new working methods must be devised for it to become a productive approach to high quality software development. Questions that arise are numerous. How do we avoid requiring human assistance to work on mundane aspects of proofs? How do we take advantage of all the progress made in automatic theorem proving? How do we organize the maintenance of ambitious corpora of formally verified knowledge in the long term?

To acquire hands-on expertise, we concentrate our activity on three aspects. The first one is foundational: we develop and maintain a library of mathematical facts that covers many aspects of algebra. In the past, we applied this library to proofs in group theory, but it is increasingly used for many different areas of mathematics and by other teams around the world, from combinatorics to elliptic cryptography, for instance. The second aspect is applicative: we develop a specific tool for proofs in cryptography, where we need to reason on the probability that opponents manage to access information we wish to protect. For this activity, we develop a specific proof system, relying on a wider set of automatic tools, with the objective of finding the tools that are well adapted to this domain and to attract users that are initially specialists in cryptography but not in formal verification. The third domain is robotics, as we believe that the current trend towards more an more autonomous robots and vehicles will raise questions of safety and trustability where formal verification can bring significant added value.

ACUMES Project-Team

3. Research Program

3.1. Research directions

The project develops along the following two axes:

- modeling complex systems through novel (unconventional) PDE systems, accounting for multi-scale phenomena and uncertainty;
- optimization and optimal control algorithms for systems governed by the above PDE systems.

These themes are motivated by the specific problems treated in the applications, and represent important and up-to-date issues in engineering sciences. For example, improving the design of transportation means and civil buildings, and the control of traffic flows, would result not only in better performances of the object of the optimization strategy (vehicles, buildings or road networks level of service), but also in enhanced safety and lower energy consumption, contributing to reduce costs and pollutant emissions.

3.1.1. PDE models accounting for multi-scale phenomena and uncertainties

Dynamical models consisting of evolutionary PDEs, mainly of hyperbolic type, appear classically in the applications studied by the previous Project-Team Opale (compressible flows, traffic, cell-dynamics, medicine, etc). Yet, the classical purely macroscopic approach is not able to account for some particular phenomena related to specific interactions occurring at smaller scales. These phenomena can be of greater importance when dealing with particular applications, where the "first order" approximation given by the purely macroscopic approach reveals to be inadequate. We refer for example to self-organizing phenomena observed in pedestrian flows [126], or to the dynamics of turbulent flows for which large scale / small scale vortical structures interfere [155].

Nevertheless, macroscopic models offer well known advantages, namely a sound analytical framework, fast numerical schemes, the presence of a low number of parameters to be calibrated, and efficient optimization procedures. Therefore, we are convinced of the interest of keeping this point of view as dominant, while completing the models with information on the dynamics at the small scale / microscopic level. This can be achieved through several techniques, like hybrid models, homogenization, mean field games. In this project, we will focus on the aspects detailed below.

The development of adapted and efficient numerical schemes is a mandatory completion, and sometimes ingredient, of all the approaches listed below. The numerical schemes developed by the team are based on finite volumes or finite elements techniques, and constitute an important tool in the study of the considered models, providing a necessary step towards the design and implementation of the corresponding optimization algorithms, see Section 3.1.2 .

3.1.1.1. Micro-macro couplings

Modeling of complex problems with a dominant macroscopic point of view often requires couplings with small scale descriptions. Accounting for systems heterogeneity or different degrees of accuracy usually leads to coupled PDE-ODE systems.

In the case of heterogeneous problems the coupling is "intrinsic", i.e. the two models evolve together and mutually affect each-other. For example, accounting for the impact of a large and slow vehicle (like a bus or a truck) on traffic flow leads to a strongly coupled system consisting of a (system of) conservation law(s) coupled with an ODE describing the bus trajectory, which acts as a moving bottleneck.The coupling is realized through a local unilateral moving constraint on the flow at the bus location, see [94] for an existence result and [77], [93] for numerical schemes.

If the coupling is intended to offer higher degree of accuracy at some locations, a macroscopic and a microscopic model are connected through an artificial boundary, and exchange information across it through suitable boundary conditions. See [84], [115] for some applications in traffic flow modelling, and [106], [111], [113] for applications to cell dynamics.

The corresponding numerical schemes are usually based on classical finite volume or finite element methods for the PDE, and Euler or Runge-Kutta schemes for the ODE, coupled in order to take into account the interaction fronts. In particular, the dynamics of the coupling boundaries require an accurate handling capturing the possible presence of non-classical shocks and preventing diffusion, which could produce wrong solutions, see for example [77], [93].

We plan to pursue our activity in this framework, also extending the above mentioned approaches to problems in two or higher space dimensions, to cover applications to crowd dynamics or fluid-structure interaction.

3.1.1.2. Micro-macro limits

Rigorous derivation of macroscopic models from microscopic ones offers a sound basis for the proposed modeling approach, and can provide alternative numerical schemes, see for example [85], [96] for the derivation of Lighthill-Whitham-Richards [138], [154] traffic flow model from Follow-the-Leader and [107] for results on crowd motion models (see also [128]). To tackle this aspect, we will rely mainly on two (interconnected) concepts: measure-valued solutions and mean-field limits.

The notion of **measure-valued solutions** for conservation laws was first introduced by DiPerna [97], and extensively used since then to prove convergence of approximate solutions and deduce existence results, see for example [108] and references therein. Measure-valued functions have been recently advocated as the appropriate notion of solution to tackle problems for which analytical results (such as existence and uniqueness of weak solutions in distributional sense) and numerical convergence are missing [64], [110]. We refer, for example, to the notion of solution for non-hyperbolic systems [116], for which no general theoretical result is available at present, and to the convergence of finite volume schemes for systems of hyperbolic conservation laws in several space dimensions, see [110].

In this framework, we plan to investigate and make use of measure-based PDE models for vehicular and pedestrian traffic flows. Indeed, a modeling approach based on (multi-scale) time-evolving measures (expressing the agents probability distribution in space) has been recently introduced (see the monograph [89]), and proved to be successful for studying emerging self-organised flow patterns [88]. The theoretical measure framework proves to be also relevant in addressing micro-macro limiting procedures of mean field type [117], where one lets the number of agents going to infinity, while keeping the total mass constant. In this case, one must prove that the *empirical measure*, corresponding to the sum of Dirac measures concentrated at the agents positions, converges to a measure-valued solution of the corresponding macroscopic evolution equation. We recall that a key ingredient in this approach is the use of the *Wasserstein distances* [163], [164]. Indeed, as observed in [147], the usual L^1 spaces are not natural in this context, since they don't guarantee uniqueness of solutions.

This procedure can potentially be extended to more complex configurations, like for example road networks or different classes of interacting agents, or to other application domains, like cell-dynamics.

Another powerful tool we shall consider to deal with micro-macro limits is the so-called **Mean Field Games (MFG)** technique (see the seminal paper [136]). This approach has been recently applied to some of the systems studied by the team, such as traffic flow and cell dynamics. In the context of crowd dynamics, including the case of several populations with different targets, the mean field game approach has been adopted in [72], [73], [98], [135], under the assumption that the individual behavior evolves according to a stochastic process, which gives rise to parabolic equations greatly simplifying the analysis of the system. Besides, a deterministic context is studied in [150], which considers a non-local velocity field. For cell dynamics, in order to take into account the fast processes that occur in the migration-related machinery, a framework such the one developed in   [92] to handle games "where agents evolve their strategies according to the best-reply scheme on a much faster time scale than their social configuration variables" may turn out to be suitable. An alternative framework to MFG is also considered. This framework is based on the formulation of -Nash- games

constrained by the **Fokker-Planck** (FP, [62]) partial differential equations that govern the time evolution of the probability density functions -PDF- of stochastic systems and on objectives that may require to follow a given PDF trajectory or to minimize an expectation functional.

3.1.1.3. Non-local flows

Non-local interactions can be described through macroscopic models based on integro-differential equations. Systems of the type

$$\partial_t u + \text{div}_{\mathbf{x}} F(t, \mathbf{x}, u, W) = 0, \qquad t > 0, x \in R^d, d \geq 1, \tag{1}$$

where $u = u(t, \mathbf{x}) \in R^N$, $N \geq 1$ is the vector of conserved quantities and the variable $W = W(t, x, u)$ depends on an integral evaluation of u, arise in a variety of physical applications. Space-integral terms are considered for example in models for granular flows [59], sedimentation [66], supply chains [120], conveyor belts [121], biological applications like structured populations dynamics [146], or more general problems like gradient constrained equations [60]. Also, non-local in time terms arise in conservation laws with memory, starting from [91]. In particular, equations with non-local flux have been recently introduced in traffic flow modeling to account for the reaction of drivers or pedestrians to the surrounding density of other individuals, see [68], [75], [81], [118], [158]. While pedestrians are likely to react to the presence of people all around them, drivers will mainly adapt their velocity to the downstream traffic, assigning a greater importance to closer vehicles. In particular, and in contrast to classical (without integral terms) macroscopic equations, these models are able to display finite acceleration of vehicles through Lipschitz bounds on the mean velocity [68], [118] and lane formation in crossing pedestrian flows.

General analytical results on non-local conservation laws, proving existence and eventually uniqueness of solutions of the Cauchy problem for 1 , can be found in [61] for scalar equations in one space dimension ($N = d = 1$), in [82] for scalar equations in several space dimensions ($N = 1$, $d \geq 1$) and in [55], [83], [87] for multi-dimensional systems of conservation laws. Besides, specific finite volume numerical methods have been developed recently in [55], [118] and [134].

Relying on these encouraging results, we aim to push a step further the analytical and numerical study of non-local models of type 1 , in particular concerning well-posedness of initial - regularity of solutions, boundary value problems and high-order numerical schemes.

3.1.1.4. Uncertainty in parameters and initial-boundary data

Different sources of uncertainty can be identified in PDE models, related to the fact that the problem of interest is not perfectly known. At first, initial and boundary condition values can be uncertain. For instance, in traffic flows, the time-dependent value of inlet and outlet fluxes, as well as the initial distribution of vehicles density, are not perfectly determined [74]. In aerodynamics, inflow conditions like velocity modulus and direction, are subject to fluctuations [124], [145]. For some engineering problems, the geometry of the boundary can also be uncertain, due to structural deformation, mechanical wear or disregard of some details [100]. Another source of uncertainty is related to the value of some parameters in the PDE models. This is typically the case of parameters in turbulence models in fluid mechanics, which have been calibrated according to some reference flows but are not universal [156], [162], or in traffic flow models, which may depend on the type of road, weather conditions, or even the country of interest (due to differences in driving rules and conductors behaviour). This leads to equations with flux functions depending on random parameters [157], [160], for which the mean and the variance of the solutions can be computed using different techniques. Indeed, uncertainty quantification for systems governed by PDEs has become a very active research topic in the last years. Most approaches are embedded in a probabilistic framework and aim at quantifying statistical moments of the PDE solutions, under the assumption that the characteristics of uncertain parameters are known. Note that classical Monte-Carlo approaches exhibit low convergence rate and consequently accurate simulations require huge computational times. In this respect, some enhanced algorithms have been proposed, for example in the balance law framework [143]. Different approaches propose to modify the PDE solvers to account for this probabilistic context, for instance by defining the non-deterministic part of the solution on an orthogonal

basis (Polynomial Chaos decomposition) and using a Galerkin projection [124], [133], [139], [166] or an entropy closure method [95], or by discretizing the probability space and extending the numerical schemes to the stochastic components [54]. Alternatively, some other approaches maintain a fully deterministic PDE resolution, but approximate the solution in the vicinity of the reference parameter values by Taylor series expansions based on first- or second-order sensitivities [151], [162], [165].

Our objective regarding this topic is twofold. In a pure modeling perspective, we aim at including uncertainty quantification in models calibration and validation for predictive use. In this case, the choice of the techniques will depend on the specific problem considered [65]. Besides, we plan to extend previous works on sensitivity analysis [100], [140] to more complex and more demanding problems. In particular, high-order Taylor expansions of the solution (greater than two) will be considered in the framework of the Sensitivity Equation Method [69] (SEM) for unsteady aerodynamic applications, to improve the accuracy of mean and variance estimations. A second targeted topic in this context is the study of the uncertainty related to turbulence closure parameters, in the sequel of [162]. We aim at exploring the capability of the SEM approach to detect a change of flow topology, in case of detached flows. Our ambition is to contribute to the emergence of a new generation of simulation tools, which will provide solution densities rather than values, to tackle real-life uncertain problems. This task will also include a reflection about numerical schemes used to solve PDE systems, in the perspective of constructing a unified numerical framework able to account for exact geometries (isogeometric methods), uncertainty propagation and sensitivity analysis w.r.t. control parameters.

3.1.2. *Optimization and control algorithms for systems governed by PDEs*

The non-classical models described above are developed in the perspective of design improvement for real-life applications. Therefore, control and optimization algorithms are also developed in conjunction with these models. The focus here is on the methodological development and analysis of optimization algorithms for PDE systems in general, keeping in mind the application domains in the way the problems are mathematically formulated.

3.1.2.1. *Sensitivity vs. adjoint equation*

Adjoint methods (achieved at continuous or discrete level) are now commonly used in industry for steady PDE problems. Our recent developments [153] have shown that the (discrete) adjoint method can be efficiently applied to cost gradient computations for time-evolving traffic flow on networks, thanks to the special structure of the associated linear systems and the underlying one dimensionality of the problem. However, this strategy is questionable for more complex (e.g. 2D/3D) unsteady problems, because it requires sophisticated and time-consuming check-pointing and/or re-computing strategies [63], [119] for the backward time integration of the adjoint variables. The sensitivity equation method (SEM) offers a promising alternative [99], [129], if the number of design parameters is moderate. Moreover, this approach can be employed for other goals, like fast evaluation of neighboring solutions or uncertainty propagation [100].

Regarding this topic, we intend to apply the continuous sensitivity equation method to challenging problems. In particular, in aerodynamics, multi-scale turbulence models like Large-Eddy Simulation (LES) [155] , Detached-Eddy Simulation (DES) [159] or Organized-Eddy Simulation (OES) [70], are more and more employed to analyse the unsteady dynamics of the flows around bluff-bodies, because they have the ability to compute the interactions of vortices at different scales, contrary to classical Reynolds-Averaged Navier-Stokes models. However, their use in design optimization is tedious, due to the long time integration required. In collaboration with turbulence specialists (M. Braza, CNRS - IMFT), we aim at developing numerical methods for effective sensitivity analysis in this context, and apply them to realistic problems, like the optimization of active flow control devices. Note that the use of SEM allows computing cost functional gradients at any time, which permits to construct new gradient-based optimization strategies like instantaneous-feedback method [131] or multiobjective optimization algorithm (see section below).

3.1.2.2. *Multi-objective descent algorithms for multi-disciplinary, multi-point, unsteady optimization or robust-design*

In differentiable optimization, multi-disciplinary, multi-point, unsteady optimization or robust-design can all be formulated as multi-objective optimization problems. In this area, we have proposed the *Multiple-Gradient*

Descent Algorithm (MGDA) to handle all criteria concurrently [102] [103]. Originally, we have stated a principle according which, given a family of local gradients, a descent direction common to all considered objective-functions simultaneously is identified, assuming the Pareto-stationarity condition is not satisfied. When the family is linearly-independent, we dispose of a direct algorithm. Inversely, when the family is linearly-dependent, a quadratic-programming problem should be solved. Hence, the technical difficulty is mostly conditioned by the number m of objective functions relative to the search space dimension n. In this respect, the basic algorithm has recently been revised [104] to handle the case where $m > n$, and even $m \gg n$, and is currently being tested on a test-case of robust design subject to a periodic time-dependent Navier-Stokes flow.

The multi-point situation is very similar and, being of great importance for engineering applications, will be treated at large.

Moreover, we intend to develop and test a new methodology for robust design that will include uncertainty effects. More precisely, we propose to employ MGDA to achieve an effective improvement of all criteria simultaneously, which can be of statistical nature or discrete functional values evaluated in confidence intervals of parameters. Some recent results obtained at ONERA [148] by a stochastic variant of our methodology confirm the viability of the approach. A PhD thesis has also been launched at ONERA/DADS.

Lastly, we note that in situations where gradients are difficult to evaluate, the method can be assisted by a meta-model [168].

3.1.2.3. Bayesian Optimization algorithms for efficient computation of general equilibria

Bayesian Optimization (BO) relies on Gaussian processes, which are used as emulators (or surrogates) of the black-box model outputs based on a small set of model evaluations. Posterior distributions provided by the Gaussian process are used to design acquisition functions that guide sequential search strategies that balance between exploration and exploitation. Such approaches have been transposed to frameworks other than optimization, such as uncertainty quantification. Our aim is to investigate how the BO apparatus can be applied to the search of general game equilibria, and in particular the classical Nash equilibrium (NE). To this end, we propose two complementary acquisition functions, one based on a greedy search approach and one based on the Stepwise Uncertainty Reduction paradigm [112]. Our proposal is designed to tackle derivative-free, expensive models, hence requiring very few model evaluations to converge to the solution.

3.1.2.4. Decentralized strategies for inverse problems

Most if not all the mathematical formulations of inverse problems (a.k.a. reconstruction, identification, data recovery, non destructive engineering,...) are known to be ill posed in the Hadamard sense. Indeed, in general, inverse problems try to fulfill (minimize) two or more very antagonistic criteria. One classical example is the Tikhonov regularization, trying to find artificially smoothed solutions close to naturally non-smooth data.

We consider here the theoretical general framework of parameter identification coupled to (missing) data recovery. Our aim is to design, study and implement algorithms derived within a game theoretic framework, which are able to find, with computational efficiency, equilibria between the "identification related players" and the "data recovery players". These two parts are known to pose many challenges, from a theoretical point of view, like the identifiability issue, and from a numerical one, like convergence, stability and robustness problems. These questions are tricky [56] and still completely open for systems like e.g. coupled heat and thermoelastic joint data and material detection.

<div align="center">

ECUADOR Project-Team

</div>

3. Research Program

3.1. Algorithmic Differentiation

Participants: Laurent Hascoët, Valérie Pascual.

> **algorithmic differentiation**   (AD, aka Automatic Differentiation) Transformation of a program, that returns a new program that computes derivatives of the initial program, i.e. some combination of the partial derivatives of the program's outputs with respect to its inputs.
>
> **adjoint**   Mathematical manipulation of the Partial Differential Equations that define a problem, obtaining new differential equations that define the gradient of the original problem's solution.
>
> **checkpointing**   General trade-off technique, used in adjoint AD, that trades duplicate execution of a part of the program to save some memory space that was used to save intermediate results.

Algorithmic Differentiation (AD) differentiates *programs*. The input of AD is a source program P that, given some $X \in \mathbb{R}^n$, returns some $Y = F(X) \in \mathbb{R}^m$, for a differentiable F. AD generates a new source program P' that, given X, computes some derivatives of F [4].

Any execution of P amounts to a sequence of instructions, which is identified with a composition of vector functions. Thus, if

$$
\begin{aligned}
P \quad &\text{runs} \quad \{I_1; I_2; \cdots I_p; \}, \\
F \quad &\text{then is} \quad f_p \circ f_{p-1} \circ \cdots \circ f_1,
\end{aligned}
\tag{2}
$$

where each f_k is the elementary function implemented by instruction I_k. AD applies the chain rule to obtain derivatives of F. Calling X_k the values of all variables after instruction I_k, i.e. $X_0 = X$ and $X_k = f_k(X_{k-1})$, the Jacobian of F is

$$
F'(X) = f'_p(X_{p-1}) \cdot f'_{p-1}(X_{p-2}) \cdot \cdots \cdot f'_1(X_0)
\tag{3}
$$

which can be mechanically written as a sequence of instructions I'_k. This can be generalized to higher level derivatives, Taylor series, etc. Combining the I'_k with the control of P yields P', and therefore this differentiation is piecewise.

The above computation of $F'(X)$, albeit simple and mechanical, can be prohibitively expensive on large codes. In practice, many applications only need cheaper projections of $F'(X)$ such as:

- **Sensitivities**, defined for a given direction $\dot{X}$ in the input space as:

$$
F'(X).\dot{X} = f'_p(X_{p-1}) \cdot f'_{p-1}(X_{p-2}) \cdot \cdots \cdot f'_1(X_0) \cdot \dot{X} \quad .
\tag{4}
$$

> This expression is easily computed from right to left, interleaved with the original program instructions. This is the *tangent mode* of AD.

- **Adjoints**, defined after transposition (F'^*), for a given weighting $\overline{Y}$ of the outputs as:

$$
F'^*(X).\overline{Y} = f'^*_1(X_0).f'^*_2(X_1). \cdots .f'^*_{p-1}(X_{p-2}).f'^*_p(X_{p-1}).\overline{Y} \quad .
\tag{5}
$$

This expression is most efficiently computed from right to left, because matrix×vector products are cheaper than matrix×matrix products. This is the *adjoint mode* of AD, most effective for optimization, data assimilation [31], adjoint problems [25], or inverse problems.

Adjoint AD builds a very efficient program [27], which computes the gradient in a time independent from the number of parameters n. In contrast, computing the same gradient with the *tangent mode* would require running the tangent differentiated program n times.

However, the X_k are required in the *inverse* of their computation order. If the original program *overwrites* a part of X_k, the differentiated program must restore X_k before it is used by $f'^*_{k+1}(X_k)$. Therefore, the central research problem of adjoint AD is to make the X_k available in reverse order at the cheapest cost, using strategies that combine storage, repeated forward computation from available previous values, or even inverted computation from available later values.

Another research issue is to make the AD model cope with the constant evolution of modern language constructs. From the old days of Fortran77, novelties include pointers and dynamic allocation, modularity, structured data types, objects, vectorial notation and parallel programming. We keep developing our models and tools to handle these new constructs.

3.2. Static Analysis and Transformation of programs

Participants: Laurent Hascoët, Valérie Pascual.

abstract syntax tree Tree representation of a computer program, that keeps only the semantically significant information and abstracts away syntactic sugar such as indentation, parentheses, or separators.

control flow graph Representation of a procedure body as a directed graph, whose nodes, known as basic blocks, each contain a sequence of instructions and whose arrows represent all possible control jumps that can occur at run-time.

abstract interpretation Model that describes program static analysis as a special sort of execution, in which all branches of control switches are taken concurrently, and where computed values are replaced by abstract values from a given *semantic domain*. Each particular analysis gives birth to a specific semantic domain.

data flow analysis Program analysis that studies how a given property of variables evolves with execution of the program. Data Flow analysis is static, therefore studying all possible run-time behaviors and making conservative approximations. A typical data-flow analysis is to detect, at any location in the source program, whether a variable is initialized or not.

The most obvious example of a program transformation tool is certainly a compiler. Other examples are program translators, that go from one language or formalism to another, or optimizers, that transform a program to make it run better. AD is just one such transformation. These tools share the technological basis that lets them implement the sophisticated analyses [14] required. In particular there are common mathematical models to specify these analyses and analyze their properties.

An important principle is *abstraction*: the core of a compiler should not bother about syntactic details of the compiled program. The optimization and code generation phases must be independent from the particular input programming language. This is generally achieved using language-specific *front-ends*, language-independent *middle-ends*, and target-specific *back-ends*. In the middle-end, analysis can concentrate on the semantics of a reduced set of constructs. This analysis operates on an abstract representation of programs made of one *call graph*, whose nodes are themselves *flow graphs* whose nodes (*basic blocks*) contain abstract *syntax trees* for the individual atomic instructions. To each level are attached symbol tables, nested to capture scoping.

Static program analysis can be defined on this internal representation, which is largely language independent. The simplest analyses on trees can be specified with inference rules [18], [28], [15]. But many *data-flow analyses* are more complex, and better defined on graphs than on trees. Since both call graphs and flow graphs may be cyclic, these global analyses will be solved iteratively. *Abstract Interpretation* [19] is a theoretical framework to study complexity and termination of these analyses.

Data flow analyses must be carefully designed to avoid or control combinatorial explosion. At the call graph level, they can run bottom-up or top-down, and they yield more accurate results when they take into account the different call sites of each procedure, which is called *context sensitivity*. At the flow graph level, they can run forwards or backwards, and yield more accurate results when they take into account only the possible execution flows resulting from possible control, which is called *flow sensitivity*.

Even then, data flow analyses are limited, because they are static and thus have very little knowledge of actual run-time values. Far before reaching the very theoretical limit of *undecidability*, one reaches practical limitations to how much information one can infer from programs that use arrays [34], [20] or pointers. Therefore, conservative *over-approximations* must be made, leading to derivative code less efficient than ideal.

3.3. Algorithmic Differentiation and Scientific Computing

Participants: Alain Dervieux, Laurent Hascoët, Bruno Koobus, Eléonore Gauci, Emmanuelle Itam, Olivier Allain, Stephen Wornom.

linearization   In Scientific Computing, the mathematical model often consists of Partial Differential Equations, that are discretized and then solved by a computer program. Linearization of these equations, or alternatively linearization of the computer program, predict the behavior of the model when small perturbations are applied. This is useful when the perturbations are effectively small, as in acoustics, or when one wants the sensitivity of the system with respect to one parameter, as in optimization.

adjoint state   Consider a system of Partial Differential Equations that define some characteristics of a system with respect to some parameters. Consider one particular scalar characteristic. Its sensitivity (or gradient) with respect to the parameters can be defined by means of *adjoint* equations, deduced from the original equations through linearization and transposition. The solution of the adjoint equations is known as the adjoint state.

Scientific Computing provides reliable simulations of complex systems. For example it is possible to *simulate* the steady or unsteady 3D air flow around a plane that captures the physical phenomena of shocks and turbulence. Next comes *optimization*, one degree higher in complexity because it repeatedly simulates and applies gradient-based optimization steps until an optimum is reached. The next sophistication is *robustness*, that detects undesirable solutions which, although maybe optimal, are very sensitive to uncertainty on design parameters or on manufacturing tolerances. This makes second derivatives come into play. Similarly *Uncertainty Quantification* can use second derivatives to evaluate how uncertainty on the simulation inputs imply uncertainty on its outputs.

To obtain this gradient and possibly higher derivatives, we advocate adjoint AD (*cf*3.1) of the program that discretizes and solves the direct system. This gives the exact gradient of the discrete function computed by the program, which is quicker and more sound than differentiating the original mathematical equations [25]. Theoretical results [24] guarantee convergence of these derivatives when the direct program converges. This approach is highly mechanizable. However, it requires careful study and special developments of the AD model [29], [32] to master possibly heavy memory usage. Among these additional developments, we promote in particular specialized AD models for Fixed-Point iterations [26], [17], efficient adjoints for linear algebra operators such as solvers, or exploitation of parallel properties of the adjoint code.

FACTAS Project-Team

3. Research Program

3.1. Introduction

Within the extensive field of inverse problems, much of the research by Factas deals with reconstructing solutions of classical elliptic PDEs from their boundary behavior. Perhaps the simplest example lies with harmonic identification of a stable linear dynamical system: the transfer-function f can be evaluated at a point $i\omega$ of the imaginary axis from the response to a periodic input at frequency ω. Since f is holomorphic in the right half-plane, it satisfies there the Cauchy-Riemann equation $\overline{\partial} f = 0$, and recovering f amounts to solve a Dirichlet problem which can be done in principle using, *e.g.* the Cauchy formula.

Practice is not nearly as simple, for f is only measured pointwise in the pass-band of the system which makes the problem ill-posed [70]. Moreover, the transfer function is usually sought in specific form, displaying the necessary physical parameters for control and design. For instance if f is rational of degree n, then $\overline{\partial} f = \sum_1^n a_j \delta_{z_j}$ where the z_j are its poles and δ_{z_j} is a Dirac unit mass at z_j. Thus, to find the domain of holomorphy (*i.e.* to locate the z_j) amounts to solve a (degenerate) free-boundary inverse problem, this time on the left half-plane. To address such questions, the team has developed a two-step approach as follows.

> **Step 1:** To determine a complete model, that is, one which is defined at every frequency, in a sufficiently versatile function class (*e.g.* Hardy spaces). This ill-posed issue requires regularization, for instance constraints on the behavior at non-measured frequencies.

> **Step 2:** To compute a reduced order model. This typically consists of rational approximation of the complete model obtained in step 1, or phase-shift thereof to account for delays. We emphasize that deriving a complete model in step 1 is crucial to achieve stability of the reduced model in step 2.

Step 1 relates to extremal problems and analytic operator theory, see Section 3.3.1 . Step 2 involves optimization, and some Schur analysis to parametrize transfer matrices of given Mc-Millan degree when dealing with systems having several inputs and outputs, see Section 3.3.2.2 . It also makes contact with the topology of rational functions, in particular to count critical points and to derive bounds, see Section 3.3.2 . Step 2 raises further issues in approximation theory regarding the rate of convergence and the extent to which singularities of the approximant (*i.e.* its poles) tend to singularities of the approximated function; this is where logarithmic potential theory becomes instrumental, see Section 3.3.3 .

Applying a realization procedure to the result of step 2 yields an identification procedure from incomplete frequency data which was first demonstrated in [76] to tune resonant microwave filters. Harmonic identification of nonlinear systems around a stable equilibrium can also be envisaged by combining the previous steps with exact linearization techniques from [34].

A similar path can be taken to approach design problems in the frequency domain, replacing the measured behavior by some desired behavior. However, describing achievable responses in terms of the design parameters is often cumbersome, and most constructive techniques rely on specific criteria adapted to the physics of the problem. This is especially true of filters, the design of which traditionally appeals to polynomial extremal problems [72], [57]. To this area, Apics contributed the use of Zolotarev-like problems for multi-band synthesis, although we presently favor interpolation techniques in which parameters arise in a more transparent manner, as well as convex relaxation of hyperbolic approximation problems, see Sections 3.2.2 and 6.2.2 .

The previous example of harmonic identification quickly suggests a generalization of itself. Indeed, on identifying $\mathbb{C}$ with $\mathbb{R}^2$, holomorphic functions become conjugate-gradients of harmonic functions, so that harmonic identification is, after all, a special case of a classical issue: to recover a harmonic function on a domain from partial knowledge of the Dirichlet-Neumann data; when the portion of boundary where data are not available is itself unknown, we meet a free boundary problem. This framework for 2-D non-destructive control was first advocated in [62] and subsequently received considerable attention. It makes clear how

to state similar problems in higher dimensions and for more general operators than the Laplacian, provided solutions are essentially determined by the trace of their gradient on part of the boundary which is the case for elliptic equations [0] [32], [80]. Such questions are particular instances of the so-called inverse potential problem, where a measure μ has to be recovered from the knowledge of the gradient of its potential (*i.e.*, the field) on part of a hypersurface (a curve in 2-D) encompassing the support of μ. For Laplace's operator, potentials are logarithmic in 2-D and Newtonian in higher dimensions. For elliptic operators with non constant coefficients, the potential depends on the form of fundamental solutions and is less manageable because it is no longer of convolution type. Nevertheless it is a useful concept bringing perspective on how problems could be raised and solved, using tools from harmonic analysis.

Inverse potential problems are severely indeterminate because infinitely many measures within an open set of $\mathbb{R}^n$ produce the same field outside this set; this phenomenon is called *balayage* [69]. In the two steps approach previously described, we implicitly removed this indeterminacy by requiring in step 1 that the measure be supported on the boundary (because we seek a function holomorphic throughout the right half-space), and by requiring in step 2 that the measure be discrete in the left half-plane (in fact: a finite sum of point masses $\sum_1^N a_j \delta_{z_j}$). The discreteness assumption also prevails in 3-D inverse source problems, see Section 4.3 . Conditions that ensure uniqueness of the solution to the inverse potential problem are part of the so-called regularizing assumptions which are needed in each case to derive efficient algorithms.

To recap, the gist of our approach is to approximate boundary data by (boundary traces of) fields arising from potentials of measures with specific support. This differs from standard approaches to inverse problems, where descent algorithms are applied to integration schemes of the direct problem; in such methods, it is the equation which gets approximated (in fact: discretized).

Along these lines, Factas advocates the use of steps 1 and 2 above, along with some singularity analysis, to approach issues of nondestructive control in 2-D and 3-D [2], [41], [45]. The team is currently engaged in the generalization to inverse source problems for the Laplace equation in 3-D, to be described further in Section 3.2.1 . There, holomorphic functions are replaced by harmonic gradients; applications are to inverse source problems in neurosciences (in particular in EEG/MEG) and inverse magnetization problems in geosciences, see Section 4.3 .

The approximation-theoretic tools developed by Apics and now by Factas to handle issues mentioned so far are outlined in Section 3.3 . In Section 3.2 to come, we describe in more detail which problems are considered and which applications are targeted.

Note that the Inria project-team Apics reached the end of its life cycle by the end of 2017. The proposal for our new team Factas was processed by the CEP (Comité des Équipes-Projets) of the Research Center in 2018, and approved by the head of the Institute in 2019.

3.2. Range of inverse problems

3.2.1. *Elliptic partial differential equations (PDE)*

Participants: Paul Asensio, Laurent Baratchart, Sylvain Chevillard, Juliette Leblond, Masimba Nemaire, Konstantinos Mavreas.

By standard properties of conjugate differentials, reconstructing Dirichlet-Neumann boundary conditions for a function harmonic in a plane domain, when these conditions are already known on a subset E of the boundary, is equivalent to recover a holomorphic function in the domain from its boundary values on E. This is the problem raised on the half-plane in step 1 of Section 3.1 . It makes good sense in holomorphic Hardy spaces where functions are entirely determined by their values on boundary subsets of positive linear measure, which is the framework for Problem (P) that we set up in Section 3.3.1 . Such issues naturally

---

[0]There is a subtle difference here between dimension 2 and higher. Indeed, a function holomorphic on a plane domain is defined by its non-tangential limit on a boundary subset of positive linear measure, but there are non-constant harmonic functions in the 3-D ball, C^1 up to the boundary sphere, yet having vanishing gradient on a subset of positive measure of the sphere. Such a "bad" subset, however, cannot have interior points on the sphere.

arise in nondestructive testing of 2-D (or 3-D cylindrical) materials from partial electrical measurements on the boundary. For instance, the ratio between the tangential and the normal currents (the so-called Robin coefficient) tells one about corrosion of the material. Thus, solving Problem (P) where ψ is chosen to be the response of some uncorroded piece with identical shape yields non destructive testing of a potentially corroded piece of material, part of which is inaccessible to measurements. This was an initial application of holomorphic extremal problems to non-destructive control [55], [58].

Another application by the team deals with non-constant conductivity over a doubly connected domain, the set E being now the outer boundary. Measuring Dirichlet-Neumann data on E, one wants to recover level lines of the solution to a conductivity equation, which is a so-called free boundary inverse problem. For this, given a closed curve inside the domain, we first quantify how constant the solution on this curve. To this effect, we state and solve an analog of Problem (P), where the constraint bears on the real part of the function on the curve (it should be close to a constant there), in a Hardy space of a conjugate Beltrami equation, of which the considered conductivity equation is the compatibility condition (just like the Laplace equation is the compatibility condition of the Cauchy-Riemann system). Subsequently, a descent algorithm on the curve leads one to improve the initial guess. For example, when the domain is regarded as separating the edge of a tokamak's vessel from the plasma (rotational symmetry makes this a 2-D situation), this method can be used to estimate the shape of a plasma subject to magnetic confinement.

This was actually carried out in collaboration with CEA (French nuclear agency) and the University of Nice (JAD Lab.), to data from *Tore Supra* in [61]. The procedure is fast because no numerical integration of the underlying PDE is needed, as an explicit basis of solutions to the conjugate Beltrami equation in terms of Bessel functions was found in this case. Generalizing this approach in a more systematic manner to free boundary problems of Bernoulli type, using descent algorithms based on shape-gradient for such approximation-theoretic criteria, is an interesting prospect to the team.

The piece of work we just mentioned requires defining and studying Hardy spaces of conjugate Beltrami equations, which is an interesting topic. For Sobolev-smooth coefficients of exponent greater than 2, they were investigated in [6], [35]. The case of the critical exponent 2 is treated in [31], which apparently provides the first example of well-posed Dirichlet problem in the non-strictly elliptic case: the conductivity may be unbounded or zero on sets of zero capacity and, accordingly, solutions need not be locally bounded. More importantly perhaps, the exponent 2 is also the key to a corresponding theory on very general (still rectifiable) domains in the plane, as coefficients of pseudo-holomorphic functions obtained by conformal transformation onto a disk are merely of L^2-class in general, even if the initial problem deals with coefficients of L^r-class for some $r > 2$. Such generalizations are now under study within the team.

Generalized Hardy classes as above are used in [32] where we address the uniqueness issue in the classical Robin inverse problem on a Lipschitz domain of $\Omega \subset \mathbb{R}^n$, $n \geq 2$, with uniformly bounded Robin coefficient, L^2 Neumann data and conductivity of Sobolev class $W^{1,r}(\Omega)$, $r > n$. We show that uniqueness of the Robin coefficient on a subset of the boundary, given Cauchy data on the complementary part, does hold in dimension $n = 2$, thanks to a unique continuation result, but needs not hold in higher dimension. In higher dimension, this raises an open issue on harmonic gradients, namely whether the positivity of the Robin coefficient is compatible with identical vanishing of the boundary gradient on a subset of positive measure.

The 3-D version of step 1 in Section 3.1 is another subject investigated by Factas: to recover a harmonic function (up to an additive constant) in a ball or a half-space from partial knowledge of its gradient. This prototypical inverse problem (*i.e.* inverse to the Cauchy problem for the Laplace equation) often recurs in electromagnetism. At present, Factas is involved with solving instances of this inverse problem arising in two fields, namely medical imaging *e.g.* for electroencephalography (EEG) or magneto-encephalography (MEG), and paleomagnetism (recovery of rocks magnetization) [2], [37], see Section 6.1 . In this connection, we collaborate with two groups of partners: Athena Inria project-team and INS (Institut de Neurosciences des Systèmes, http://ins.univ-amu.fr/), hospital la Timone, Aix-Marseille Univ., on the one hand, Geosciences Lab. at MIT and Cerege CNRS Lab. on the other hand. The question is considerably more difficult than its 2-D counterpart, due mainly to the lack of multiplicative structure for harmonic gradients. Still, substantial progress has been made over the last years using methods of harmonic analysis and operator theory.

The team is further concerned with 3-D generalizations and applications to non-destructive control of step 2 in Section 3.1 . A typical problem is here to localize inhomogeneities or defaults such as cracks, sources or occlusions in a planar or 3-dimensional object, knowing thermal, electrical, or magnetic measurements on the boundary. These defaults can be expressed as a lack of harmonicity of the solution to the associated Dirichlet-Neumann problem, thereby posing an inverse potential problem in order to recover them. In 2-D, finding an optimal discretization of the potential in Sobolev norm amounts to solve a best rational approximation problem, and the question arises as to how the location of the singularities of the approximant (*i.e.* its poles) reflects the location of the singularities of the potential (*i.e.* the defaults we seek). This is a fairly deep issue in approximation theory, to which Apics contributed convergence results for certain classes of fields expressed as Cauchy integrals over extremal contours for the logarithmic potential [8], [38], [52]. Initial schemes to locate cracks or sources *via* rational approximation on planar domains were obtained this way [41], [45], [55]. It is remarkable that finite inverse source problems in 3-D balls, or more general algebraic surfaces, can be approached using these 2-D techniques upon slicing the domain into planar sections [9], [42]. More precisely, each section cuts out a planar domain, the boundary of which carries data which can be proved to match an algebraic function. The singularities of this algebraic function are not located at the 3-D sources, but are related to them: the section contains a source if and only if some function of the singularities in that section meets a relative extremum. Using bisection it is thus possible to determine an extremal place along all sections parallel to a given plane direction, up to some threshold which has to be chosen small enough that one does not miss a source. This way, we reduce the original source problem in 3-D to a sequence of inverse poles and branchpoints problems in 2-D. This bottom line generates a steady research activity within Factas, and again applications are sought to medical imaging and geosciences, see Sections 4.3 , 4.2 and 6.1 .

Conjectures may be raised on the behavior of optimal potential discretization in 3-D, but answering them is an ambitious program still in its infancy.

3.2.2. *Systems, transfer and scattering*

Participants: Laurent Baratchart, Sylvain Chevillard, Adam Cooman, Martine Olivi, Fabien Seyfert.

Through contacts with CNES (French space agency), members of the team became involved in identification and tuning of microwave electromagnetic filters used in space telecommunications, see Section 4.4 . The initial problem was to recover, from band-limited frequency measurements, physical parameters of the device under examination. The latter consists of interconnected dual-mode resonant cavities with negligible loss, hence its scattering matrix is modeled by a 2×2 unitary-valued matrix function on the frequency line, say the imaginary axis to fix ideas. In the bandwidth around the resonant frequency, a modal approximation of the Helmholtz equation in the cavities shows that this matrix is approximately rational, of Mc-Millan degree twice the number of cavities.

This is where system theory comes into play, through the so-called *realization* process mapping a rational transfer function in the frequency domain to a state-space representation of the underlying system of linear differential equations in the time domain. Specifically, realizing the scattering matrix allows one to construct a virtual electrical network, equivalent to the filter, the parameters of which mediate in between the frequency response and the geometric characteristics of the cavities (*i.e.* the tuning parameters).

Hardy spaces provide a framework to transform this ill-posed issue into a series of regularized analytic and meromorphic approximation problems. More precisely, the procedure sketched in Section 3.1 goes as follows:

1. infer from the pointwise boundary data in the bandwidth a stable transfer function (*i.e.* one which is holomorphic in the right half-plane), that may be infinite dimensional (numerically: of high degree). This is done by solving a problem analogous to (P) in Section 3.3.1 , while taking into account prior knowledge on the decay of the response outside the bandwidth, see [13] for details.

2. A stable rational approximation of appropriate degree to the model obtained in the previous step is performed. For this, a descent method on the compact manifold of inner matrices of given size and degree is used, based on an original parametrization of stable transfer functions developed within the team [27], [13].

3. Realizations of this rational approximant are computed. To be useful, they must satisfy certain constraints imposed by the geometry of the device. These constraints typically come from the coupling topology of the equivalent electrical network used to model the filter. This network is composed of resonators, coupled according to some specific graph. This realization step can be recast, under appropriate compatibility conditions [56], as solving a zero-dimensional multivariate polynomial system. To tackle this problem in practice, we use Gröbner basis techniques and continuation methods which team up in the Dedale-HF software (see Section 3.4.2).

We recently started a collaboration with the Chinese Hong Kong University on the topic of frequency depending couplings appearing in the equivalent circuits we compute [19] continuing our work [1] on wide-band design applications.

Factas also investigates issues pertaining to design rather than identification. Given the topology of the filter, a basic problem in this connection is to find the optimal response subject to specifications that bear on rejection, transmission and group delay of the scattering parameters. Generalizing the classical approach based on Chebyshev polynomials for single band filters, we recast the problem of multi-band response synthesis as a generalization of the classical Zolotarev min-max problem for rational functions [26] [12]. Thanks to quasi-convexity, the latter can be solved efficiently using iterative methods relying on linear programming. These were implemented in the software easy-FF (see easy-FF). Currently, the team is engaged in the synthesis of more complex microwave devices like multiplexers and routers, which connect several filters through wave guides. Schur analysis plays an important role here, because scattering matrices of passive systems are of Schur type (*i.e.* contractive in the stability region). The theory originates with the work of I. Schur [75], who devised a recursive test to check for contractivity of a holomorphic function in the disk. The so-called Schur parameters of a function may be viewed as Taylor coefficients for the hyperbolic metric of the disk, and the fact that Schur functions are contractions for that metric lies at the root of Schur's test. Generalizations thereof turn out to be efficient to parametrize solutions to contractive interpolation problems [28]. Dwelling on this, Factas contributed differential parametrizations (atlases of charts) of lossless matrix functions [27], [71], [66] which are fundamental to our rational approximation software RARL2 (see Section 3.4.5). Schur analysis is also instrumental to approach de-embedding issues, and provides one with considerable insight into the so-called matching problem. The latter consists in maximizing the power a multiport can pass to a given load, and for reasons of efficiency it is all-pervasive in microwave and electric network design, *e.g.* of antennas, multiplexers, wifi cards and more. It can be viewed as a rational approximation problem in the hyperbolic metric, and the team presently deals with this hot topic using contractive interpolation with constraints on boundary peak points, within the framework of the (defense funded) ANR Cocoram, see Sections 6.2 .

In recent years, our attention was driven by CNES and UPV (Bilbao) to questions about stability of high-frequency amplifiers. Contrary to previously discussed devices, these are *active* components. The response of an amplifier can be linearized around a set of primary current and voltages, and then admittances of the corresponding electrical network can be computed at various frequencies, using the so-called harmonic balance method. The initial goal is to check for stability of the linearized model, so as to ascertain existence of a well-defined working state. The network is composed of lumped electrical elements namely inductors, capacitors, negative *and* positive resistors, transmission lines, and controlled current sources. Our research so far has focused on describing the algebraic structure of admittance functions, so as to set up a function-theoretic framework where the two-steps approach outlined in Section 3.1 can be put to work. The main discovery is that the unstable part of each partial transfer function is rational and can be computed by analytic projection, see Section 6.3 . We now start investigating the linearized harmonic transfer-function around a periodic cycle, to check for stability under non necessarily small inputs. This topic generates the doctoral work of S. Fueyo.

3.3. Approximation

Participants: Laurent Baratchart, Sylvain Chevillard, Juliette Leblond, Martine Olivi, Fabien Seyfert.

3.3.1. *Best analytic approximation*

In dimension 2, the prototypical problem to be solved in step 1 of Section 3.1 may be described as: given a domain $D \subset \mathbb{R}^2$, to recover a holomorphic function from its values on a subset K of the boundary of D. For the discussion it is convenient to normalize D, which can be done by conformal mapping. So, in the simply connected case, we fix D to be the unit disk with boundary unit circle T. We denote by H^p the Hardy space of exponent p, which is the closure of polynomials in $L^p(T)$-norm if $1 \le p < \infty$ and the space of bounded holomorphic functions in D if $p = \infty$. Functions in H^p have well-defined boundary values in $L^p(T)$, which makes it possible to speak of (traces of) analytic functions on the boundary.

To find an analytic function g in D matching some measured values f approximately on a sub-arc K of T, we formulate a constrained best approximation problem as follows.

> (P) Let $1 \le p \le \infty$, K a sub-arc of T, $f \in L^p(K)$, $\psi \in L^p(T \smallsetminus K)$ and $M > 0$; find a function $g \in H^p$ such that $\|g - \psi\|_{L^p(T \smallsetminus K)} \le M$ and $g - f$ is of minimal norm in $L^p(K)$ under this constraint.

Here ψ is a reference behavior capturing *a priori* assumptions on the behavior of the model off K, while M is some admissible deviation thereof. The value of p reflects the type of stability which is sought and how much one wants to smooth out the data. The choice of L^p classes is suited to handle pointwise measurements.

To fix terminology, we refer to (P) as a *bounded extremal problem*. As shown in [40], [43], [49], the solution to this convex infinite-dimensional optimization problem can be obtained when $p \ne 1$ upon iterating with respect to a Lagrange parameter the solution to spectral equations for appropriate Hankel and Toeplitz operators. These spectral equations involve the solution to the special case $K = T$ of (P), which is a standard extremal problem [64]:

> (P_0) Let $1 \le p \le \infty$ and $\varphi \in L^p(T)$; find a function $g \in H^p$ such that $g - \varphi$ is of minimal norm in $L^p(T)$.

In the case $p = 1$, partial results are known but computational issues remain open.

Various modifications of (P) can be tailored to meet specific needs. For instance when dealing with lossless transfer functions (see Section 4.4), one may want to express the constraint on $T \smallsetminus K$ in a pointwise manner: $|g - \psi| \le M$ a.e. on $T \smallsetminus K$, see [44]. In this form, the problem comes close to (but still is different from) H^∞ frequency optimization used in control [67], [74]. One can also impose bounds on the real or imaginary part of $g - \psi$ on $T \smallsetminus K$, which is useful when considering Dirichlet-Neumann problems.

The analog of Problem (P) on an annulus, K being now the outer boundary, can be seen as a means to regularize a classical inverse problem occurring in nondestructive control, namely to recover a harmonic function on the inner boundary from Dirichlet-Neumann data on the outer boundary (see Sections 3.2.1 , 4.3 , 6.1.3). It may serve as a tool to approach Bernoulli type problems, where we are given data on the outer boundary and we *seek the inner boundary*, knowing it is a level curve of the solution. In this case, the Lagrange parameter indicates how to deform the inner contour in order to improve data fitting. Similar topics are discussed in Section 3.2.1 for more general equations than the Laplacian, namely isotropic conductivity equations of the form $\operatorname{div}(\sigma \nabla u) = 0$ where σ is no longer constant (*i.e.*, varies in the space). Then, the Hardy spaces in Problem (P) are those of a so-called conjugate Beltrami equation: $\overline{\partial} f = \nu \overline{\partial f}$ [68], which are studied for $1 < p < \infty$ in [6], [31], [35] and [59]. Expansions of solutions needed to constructively handle such issues in the specific case of linear fractional conductivities (occurring for instance in plasma shaping) have been expounded in [61].

Though originally considered in dimension 2, Problem (P) carries over naturally to higher dimensions where analytic functions get replaced by gradients of harmonic functions. Namely, given some open set $\Omega \subset \mathbb{R}^n$ and some $\mathbb{R}^n$-valued vector field V on an open subset O of the boundary of Ω, we seek a harmonic function in Ω whose gradient is close to V on O.

When Ω is a ball or a half-space, a substitute for holomorphic Hardy spaces is provided by the Stein-Weiss Hardy spaces of harmonic gradients [78]. Conformal maps are no longer available when $n > 2$, so that Ω can no longer be normalized. More general geometries than spheres and half-spaces have not been much studied so far.

On the ball, the analog of Problem (P) is

> (P_1) Let $1 \leq p \leq \infty$ and $B \subset \mathbb{R}^n$ the unit ball. Fix O an open subset of the unit sphere $S \subset \mathbb{R}^n$. Let further $V \in L^p(O)$ and $W \in L^p(S \setminus O)$ be $\mathbb{R}^n$-valued vector fields. Given $M > 0$, find a harmonic gradient $G \in H^p(B)$ such that $\|G - W\|_{L^p(S \setminus O)} \leq M$ and $G - V$ is of minimal norm in $L^p(O)$ under this constraint.

When $p = 2$, Problem (P_1) was solved in [2] as well as its analog on a shell, when the tangent component of V is a gradient (when O is Lipschitz the general case follows easily from this). The solution extends the work in [40] to the 3-D case, using a generalization of Toeplitz operators. The case of the shell was motivated by applications to the processing of EEG data. An important ingredient is a refinement of the Hodge decomposition, that we call the *Hardy-Hodge* decomposition, allowing us to express a $\mathbb{R}^n$-valued vector field in $L^p(S)$, $1 < p < \infty$, as the sum of a vector field in $H^p(B)$, a vector field in $H^p(\mathbb{R}^n \setminus \overline{B})$, and a tangential divergence free vector field on S; the space of such divergence-free fields is denoted by $D(S)$. If $p = 1$ or $p = \infty$, L^p must be replaced by the real Hardy space or the space of functions with bounded mean oscillation. More generally this decomposition, which is valid on any sufficiently smooth surface (see Section 6.1), seems to play a fundamental role in inverse potential problems. In fact, it was first introduced formally on the plane to describe silent magnetizations supported in $\mathbb{R}^2$ (*i.e.* those generating no field in the upper half space) [37].

Just like solving problem (P) appeals to the solution of problem (P_0), our ability to solve problem (P_1) will depend on the possibility to tackle the special case where $O = S$:

> (P_2) Let $1 \leq p \leq \infty$ and $V \in L^p(S)$ be a $\mathbb{R}^n$-valued vector field. Find a harmonic gradient $G \in H^p(B)$ such that $\|G - V\|_{L^p(S)}$ is minimum.

Problem (P_2) is simple when $p = 2$ by virtue of the Hardy-Hodge decomposition together with orthogonality of $H^2(B)$ and $H^2(\mathbb{R}^n \setminus \overline{B})$, which is the reason why we were able to solve (P_1) in this case. Other values of p cannot be treated as easily and are still under investigation, especially the case $p = \infty$ which is of particular interest and presents itself as a 3-D analog to the Nehari problem [73].

Companion to problem (P_2) is problem (P_3) below.

> (P_3) Let $1 \leq p \leq \infty$ and $V \in L^p(S)$ be a $\mathbb{R}^n$-valued vector field. Find $G \in H^p(B)$ and $D \in D(S)$ such that $\|G + D - V\|_{L^p(S)}$ is minimum.

Note that (P_2) and (P_3) are identical in 2-D, since no non-constant tangential divergence-free vector field exists on T. It is no longer so in higher dimension, where both (P_2) and (P_3) arise in connection with inverse potential problems in divergence form, like source recovery in electro/magneto encephalography and paleomagnetism, see Sections 3.2.1 and 4.3 .

3.3.2. *Best meromorphic and rational approximation*

The techniques set forth in this section are used to solve step 2 in Section 3.2 and they are instrumental to approach inverse boundary value problems for the Poisson equation $\Delta u = \mu$, where μ is some (unknown) measure.

3.3.2.1. *Scalar meromorphic and rational approximation*

We put R_N for the set of rational functions with at most N poles in D. By definition, meromorphic functions in $L^p(T)$ are (traces of) functions in $H^p + R_N$.

A natural generalization of problem (P_0) is:

> (P_N) Let $1 \leq p \leq \infty$, $N \geq 0$ an integer, and $f \in L^p(T)$; find a function $g_N \in H^p + R_N$ such that $g_N - f$ is of minimal norm in $L^p(T)$.

Only for $p = \infty$ and f continuous is it known how to solve (P_N) in semi-closed form. The unique solution is given by AAK theory (named after Adamjan, Arov and Krein), which connects the spectral decomposition of Hankel operators with best approximation [73].

The case where $p = 2$ is of special importance for it reduces to rational approximation. Indeed, if we write the Hardy decomposition $f = f^+ + f^-$ where $f^+ \in H^2$ and $f^- \in H^2(\mathbb{C} \smallsetminus \overline{D})$, then $g_N = f^+ + r_N$ where r_N is a best approximant to f^- from R_N in $L^2(T)$. Moreover, r_N has no pole outside D, hence it is a *stable* rational approximant to f^-. However, in contrast to the case where $p = \infty$, this best approximant may *not* be unique.

The Miaou project (predecessor of Apics) already designed a dedicated steepest-descent algorithm for the case $p = 2$ whose convergence to a *local minimum* is guaranteed; the algorithm ha evolved over years and still now, it seems to be the only procedure meeting this property. This gradient algorithm proceeds recursively with respect to N on a compactification of the parameter space [33]. Although it has proved to be effective in all applications carried out so far (see Sections 4.3 , 4.4), it is still unknown whether the absolute minimum can always be obtained by choosing initial conditions corresponding to *critical points* of lower degree (as is done by the RARL2 software, Section 3.4.5).

In order to establish global convergence results, Apics has undertaken a deeper study of the number and nature of critical points (local minima, saddle points, ...), in which tools from differential topology and operator theory team up with classical interpolation theory [46], [48]. Based on this work, uniqueness or asymptotic uniqueness of the approximant was proved for certain classes of functions like transfer functions of relaxation systems (*i.e.* Markov functions) [50] and more generally Cauchy integrals over hyperbolic geodesic arcs [53]. These are the only results of this kind. Research by Apics on this topic remained dormant for a while by reasons of opportunity, but revisiting the work [29] in higher dimension is a worthy and timely endeavor today. Meanwhile, an analog to AAK theory was carried out for $2 \leq p < \infty$ in [49]. Although not as effective computationally, it was recently used to derive lower bounds [5]. When $1 \leq p < 2$, problem (P_N) is still quite open.

A common feature to the above-mentioned problems is that critical point equations yield non-Hermitian orthogonality relations for the denominator of the approximant. This stresses connections with interpolation, which is a standard way to build approximants, and in many respects best or near-best rational approximation may be regarded as a clever manner to pick interpolation points. This was exploited in [54], [51], and is used in an essential manner to assess the behavior of poles of best approximants to functions with branched singularities, which is of particular interest for inverse source problems (*cf.* Sections 3.4.3 and 6.1).

In higher dimensions, the analog of Problem (P_N) is best approximation of a vector field by gradients of discrete potentials generated by N point masses. This basic issue is by no means fully understood, and it is an exciting field of research. It is connected with certain generalizations of Toeplitz or Hankel operators, and with constructive approaches to so-called weak factorizations for real Hardy functions [60].

Besides, certain constrained rational approximation problems, of special interest in identification and design of passive systems, arise when putting additional requirements on the approximant, for instance that it should be smaller than 1 in modulus (*i.e.* a Schur function). In particular, Schur interpolation lately received renewed attention from the team, in connection with matching problems. There, interpolation data are subject to a well-known compatibility condition (positive definiteness of the so-called Pick matrix), and the main difficulty is to put interpolation points on the boundary of D while controlling both the degree and the extremal points (peak points for the modulus) of the interpolant. Results obtained by Apics in this direction generalize a variant of contractive interpolation with degree constraint as studied in [65]. We mention that contractive interpolation with nodes approaching the boundary has been a subsidiary research topic by the team in the past, which plays an interesting role in the spectral representation of certain non-stationary stochastic processes [36], [39].

3.3.2.2. Matrix-valued rational approximation

Matrix-valued approximation is necessary to handle systems with several inputs and outputs but it generates additional difficulties as compared to scalar-valued approximation, both theoretically and algorithmically. In the matrix case, the McMillan degree (*i.e.* the degree of a minimal realization in the System-Theoretic sense) generalizes the usual notion of degree for rational functions. For instance when poles are simple, the McMillan degree is the sum of the ranks of the residues.

The basic problem that we consider now goes as follows: *let* $\mathcal{F} \in (H^2)^{m \times l}$ *and* n *an integer; find a rational matrix of size* $m \times l$ *without poles in the unit disk and of McMillan degree at most* n *which is nearest possible to* $\mathcal{F}$ *in* $(H^2)^{m \times l}$. Here the L^2 norm of a matrix is the square root of the sum of the squares of the norms of its entries.

The scalar approximation algorithm derived in [33] and mentioned in Section 3.3.2.1 generalizes to the matrix-valued situation [63]. The first difficulty here is to parametrize inner matrices (*i.e.* matrix-valued functions analytic in the unit disk and unitary on the unit circle) of given McMillan degree degree n. Indeed, inner matrices play the role of denominators in fractional representations of transfer matrices (using the so-called Douglas-Shapiro-Shields factorization). The set of inner matrices of given degree is a smooth manifold that allows one to use differential tools as in the scalar case. In practice, one has to produce an atlas of charts (local parametrizations) and to handle changes of charts in the course of the algorithm. Such parametrization can be obtained using interpolation theory and Schur-type algorithms, the parameters of which are vectors or matrices ([27], [66], [71]). Some of these parametrizations are also interesting to compute realizations and achieve filter synthesis ([66], [71]). The rational approximation software "RARL2" developed by the team is described in Section 3.4.5 .

Difficulties relative to multiple local minima of course arise in the matrix-valued case as well, and deriving criteria that guarantee uniqueness is even more difficult than in the scalar case. The case of rational functions of degree n or small perturbations thereof (the consistency problem) was solved in [47]. Matrix-valued Markov functions are the only known example beyond this one [30].

Let us stress that RARL2 seems the only algorithm handling rational approximation in the matrix case that demonstrably converges to a local minimum while meeting stability constraints on the approximant. It is still a working pin of many developments by Factas on frequency optimization and design.

3.3.3. *Behavior of poles of meromorphic approximants*
Participant: Laurent Baratchart.

We refer here to the behavior of poles of best meromorphic approximants, in the L^p-sense on a closed curve, to functions f defined as Cauchy integrals of complex measures whose support lies inside the curve. Normalizing the contour to be the unit circle T, we are back to Problem (P_N) in Section 3.3.2.1 ; invariance of the latter under conformal mapping was established in [45]. Research so far has focused on functions whose singular set inside the contour is polar, meaning that the function can be continued analytically (possibly in a multiple-valued manner) except over a set of logarithmic capacity zero.

Generally speaking in approximation theory, assessing the behavior of poles of rational approximants is essential to obtain error rates as the degree goes large, and to tackle constructive issues like uniqueness. However, as explained in Section 3.2.1 , the original twist by Apics, now Factas, is to consider this issue also as a means to extract information on singularities of the solution to a Dirichlet-Neumann problem. The general theme is thus: *how do the singularities of the approximant reflect those of the approximated function?* This approach to inverse problem for the 2-D Laplacian turns out to be attractive when singularities are zero- or one-dimensional (see Section 4.3). It can be used as a computationally cheap initial condition for more precise but much heavier numerical optimizations which often do not even converge unless properly initialized. As regards crack detection or source recovery, this approach boils down to analyzing the behavior of best meromorphic approximants of given pole cardinality to a function with branch points, which is the prototype of a polar singular set. For piecewise analytic cracks, or in the case of sources, we were able to prove ([8], [45], [38]), that the poles of the approximants accumulate, when the degree goes large, to some extremal cut of minimum weighted logarithmic capacity connecting the singular points of the crack, or the sources [41]. Moreover, the asymptotic density of the poles turns out to be the Green equilibrium distribution on this cut in D, therefore it charges the singular points if one is able to approximate in sufficiently high degree (this is where the method could fail, because high-order approximation requires rather precise data).

The case of two-dimensional singularities is still an outstanding open problem.

It is remarkable that inverse source problems inside a sphere or an ellipsoid in 3-D can be approached with such 2-D techniques, as applied to planar sections, see Section 6.1 . The technique is implemented in the software FindSources3D, see Section 3.4.3 .

3.4. Software tools of the team

In addition to the above-mentioned research activities, Factas develops and maintains a number of long-term software tools that either implement and illustrate effectiveness of the algorithms theoretically developed by the team or serve as tools to help further research by team members. We present briefly the most important of them.

3.4.1. *Pisa*

KEYWORDS: Electrical circuit - Stability

FUNCTIONAL DESCRIPTION: To minimise prototyping costs, the design of analog circuits is performed using computer-aided design tools which simulate the circuit's response as accurately as possible.

Some commonly used simulation tools do not impose stability, which can result in costly errors when the prototype turns out to be unstable. A thorough stability analysis is therefore a very important step in circuit design. This is where pisa is used.

pisa is a Matlab toolbox that allows designers of analog electronic circuits to determine the stability of their circuits in the simulator. It analyses the impedance presented by a circuit to determine the circuit's stability. When an instability is detected, pisa can estimate location of the unstable poles to help designers fix their stability issue.

RELEASE FUNCTIONAL DESCRIPTION: First version
- Authors: Adam Cooman, David Martinez Martinez, Fabien Seyfert and Martine Olivi
- Contact: Fabien Seyfert
- Publications: Model-Free Closed-Loop Stability Analysis: A Linear Functional Approach - On Transfer Functions Realizable with Active Electronic Components
- URL: https://project.inria.fr/pisa

3.4.2. *DEDALE-HF*

SCIENTIFIC DESCRIPTION

Dedale-HF consists in two parts: a database of coupling topologies as well as a dedicated predictor-corrector code. Roughly speaking each reference file of the database contains, for a given coupling topology, the complete solution to the coupling matrix synthesis problem (C.M. problem for short) associated to particular filtering characteristics. The latter is then used as a starting point for a predictor-corrector integration method that computes the solution to the C.M. corresponding to the user-specified filter characteristics. The reference files are computed off-line using Gröbner basis techniques or numerical techniques based on the exploration of a monodromy group. The use of such continuation techniques, combined with an efficient implementation of the integrator, drastically reduces the computational time.

Dedale-HF has been licensed to, and is currently used by TAS-Espana
FUNCTIONAL DESCRIPTION

Dedale-HF is a software dedicated to solve exhaustively the coupling matrix synthesis problem in reasonable time for the filtering community. Given a coupling topology, the coupling matrix synthesis problem consists in finding all possible electromagnetic coupling values between resonators that yield a realization of given filter characteristics. Solving the latter is crucial during the design step of a filter in order to derive its physical dimensions, as well as during the tuning process where coupling values need to be extracted from frequency measurements.
- Participant: Fabien Seyfert
- Contact: Fabien Seyfert
- URL: http://www-sop.inria.fr/apics/Dedale/

3.4.3. FindSources3D

KEYWORDS: Health - Neuroimaging - Visualization - Compilers - Medical - Image - Processing

FindSources3D is a software program dedicated to the resolution of inverse source problems in electroencephalography (EEG). From pointwise measurements of the electrical potential taken by electrodes on the scalp, FindSources3D estimates pointwise dipolar current sources within the brain in a spherical model.

After a first data transmission "cortical mapping" step, it makes use of best rational approximation on 2-D planar cross-sections and of the software RARL2 in order to locate singularities. From those planar singularities, the 3-D sources are estimated in a last step, see [9].

The present version of FindSources3D (called FindSources3D-bolis) provides a modular, ergonomic, accessible and interactive platform, with a convenient graphical interface for EEG medical imaging. Modularity is now granted (using the tools dtk, Qt, with compiled Matlab libraries). It offers a detailed and nice visualization of data and tuning parameters, processing steps, and of the computed results (using VTK).

A new version is being developed that will incorporate a first Singular Value Decomposition (SVD) step in order to be able to handle time dependent data and to find the corresponding principal static components.

- Participants: Juliette Leblond, Maureen Clerc (team Athena, Inria Sophia), Jean-Paul Marmorat, Théodore Papadopoulo (team Athena).
- Contact: Juliette Leblond
- URL: http://www-sop.inria.fr/apics/FindSources3D/en/index.html

3.4.4. PRESTO-HF

SCIENTIFIC DESCRIPTION

For the matrix-valued rational approximation step, Presto-HF relies on RARL2. Constrained realizations are computed using the Dedale-HF software. As a toolbox, Presto-HF has a modular structure, which allows one for example to include some building blocks in an already existing software.

The delay compensation algorithm is based on the following assumption: far off the pass-band, one can reasonably expect a good approximation of the rational components of S11 and S22 by the first few terms of their Taylor expansion at infinity, a small degree polynomial in $1/s$. Using this idea, a sequence of quadratic convex optimization problems are solved, in order to obtain appropriate compensations. In order to check the previous assumption, one has to measure the filter on a larger band, typically three times the pass band.

This toolbox has been licensed to (and is currently used by) Thales Alenia Space in Toulouse and Madrid, Thales airborne systems and Flextronics (two licenses). Xlim (University of Limoges) is a heavy user of Presto-HF among the academic filtering community and some free license agreements have been granted to the microwave department of the University of Erlangen (Germany) and the Royal Military College (Kingston, Canada).

FUNCTIONAL DESCRIPTION

Presto-HF is a toolbox dedicated to low-pass parameter identification for microwave filters. In order to allow the industrial transfer of our methods, a Matlab-based toolbox has been developed, dedicated to the problem of identification of low-pass microwave filter parameters. It allows one to run the following algorithmic steps, either individually or in a single stroke:

• Determination of delay components caused by the access devices (automatic reference plane adjustment),

• Automatic determination of an analytic completion, bounded in modulus for each channel,

• Rational approximation of fixed McMillan degree,

• Determination of a constrained realization.

- Participants: Fabien Seyfert, Jean-Paul Marmorat and Martine Olivi
- Contact: Fabien Seyfert
- URL: https://project.inria.fr/presto-hf/

3.4.5. RARL2

Réalisation interne et Approximation Rationnelle L2
SCIENTIFIC DESCRIPTION

The method is a steepest-descent algorithm. A parametrization of MIMO systems is used, which ensures that the stability constraint on the approximant is met. The implementation, in Matlab, is based on state-space representations.

RARL2 performs the rational approximation step in the software tools PRESTO-HF and FindSources3D. It is distributed under a particular license, allowing unlimited usage for academic research purposes. It was released to the universities of Delft and Maastricht (the Netherlands), Cork (Ireland), Brussels (Belgium), Macao (China) and BITS-Pilani Hyderabad Campus (India).
FUNCTIONAL DESCRIPTION

RARL2 is a software for rational approximation. It computes a stable rational L2-approximation of specified order to a given L2-stable (L2 on the unit circle, analytic in the complement of the unit disk) matrix-valued function. This can be the transfer function of a multivariable discrete-time stable system. RARL2 takes as input either:

• its internal realization,

• its first N Fourier coefficients,

• discretized (uniformly distributed) values on the circle. In this case, a least-square criterion is used instead of the L2 norm.

It thus performs model reduction in the first or the second case, and leans on frequency data identification in the third. For band-limited frequency data, it could be necessary to infer the behavior of the system outside the bandwidth before performing rational approximation.

An appropriate Möbius transformation allows to use the software for continuous-time systems as well.

- Participants: Jean-Paul Marmorat and Martine Olivi
- Contact: Martine Olivi
- URL: http://www-sop.inria.fr/apics/RARL2/rarl2.html

3.4.6. Sollya

KEYWORDS: Numerical algorithm - Supremum norm - Curve plotting - Remez algorithm - Code generator - Proof synthesis
FUNCTIONAL DESCRIPTION

Sollya is an interactive tool where the developers of mathematical floating-point libraries (libm) can experiment before actually developing code. The environment is safe with respect to floating-point errors, i.e. the user precisely knows when rounding errors or approximation errors happen, and rigorous bounds are always provided for these errors.

Among other features, it offers a fast Remez algorithm for computing polynomial approximations of real functions and also an algorithm for finding good polynomial approximants with floating-point coefficients to any real function. As well, it provides algorithms for the certification of numerical codes, such as Taylor Models, interval arithmetic or certified supremum norms.

It is available as a free software under the CeCILL-C license.

- Participants: Sylvain Chevillard, Christoph Lauter, Mioara Joldes and Nicolas Jourdan
- Partners: CNRS - ENS Lyon - UCBL Lyon 1
- Contact: Sylvain Chevillard
- URL: http://sollya.gforge.inria.fr/

<h1 style="text-align:center; color:red">MCTAO Project-Team</h1>

3. Research Program

3.1. Control Problems

McTAO's major field of expertise is control theory in the large sense. Let us give an overview of this field.

Modelling.   Our effort is directed toward efficient methods for the control of real (physical) *systems*, based on a *model* of the system to be controlled. Choosing accurate models yet simple enough to allow control design is in itself a key issue. The typical continuous-time model is of the form $dx/dt = f(x, u)$ where x is the *state*, ideally finite dimensional, and u the *control*; the control is left free to be a function of time, or a function of the state, or obtained as the solution of another dynamical system that takes x as an input. Modelling amounts to deciding the nature and dimension of x, as well as the dynamics (roughly speaking the function f). Connected to modeling is identification of parameters when a finite number of parameters are left free in "f".

Controllability, path planning.   Controllability is a property of a control system (in fact of a model) that two states in the state space can be connected by a trajectory generated by some control, here taken as an explicit function of time. Deciding on local or global controllability is still a difficult open question in general. In most cases, controllability can be decided by linear approximation, or non-controllability by "physical" first integrals that the control does not affect. For some critically actuated systems, it is still difficult to decide local or global controllability, and the general problem is anyway still open. Path planning is the problem of constructing the control that actually steers one state to another.

Optimal control.   In optimal control, one wants to find, among the controls that satisfy some contraints at initial and final time (for instance given initial and final state as in path planning), the ones that minimize some criterion. This is important in many control engineering problems, because minimizing a cost is often very relevant. Mathematically speaking, optimal control is the modern branch of the calculus of variations, rather well established and mature [71], [47], [34], but with a lot of hard open questions. In the end, in order to actually compute these controls, ad-hoc numerical schemes have to be derived for effective computations of the optimal solutions. See more about our research program in optimal control in section 3.2 .

Feedback control.   In the above two paragraphs, the control is an explicit function of time. To address in particular the stability issues (sensitivity to errors in the model or the initial conditions for example), the control has to be taken as a function of the (measured) state, or part of it. This is known as closed-loop control; it must be combined with optimal control in many real problems. On the problem of stabilization, there is longstanding research record from members of the team, in particular on the construction of "Control Lyapunov Functions", see [62], [73].

Classification of control systems   One may perform various classes of transformations acting on systems, or rather on models... The simpler ones come from point-to-point transformations (changes of variables) on the state and control, and more intricate ones consist in embedding an extraneous dynamical system into the model, these are dynamic feedback transformations, they change the dimension of the state. In most problems, choosing the proper coordinates, or the right quantities that describe a phenomenon, sheds light on a path to the solution; these proper choices may sometimes be found from an understanding of the modelled phenomenons, or it can come from the study of the geometry of the equations and the transformation acting on them. This justifies the investigations of these transformations on models for themselves. These topics are central in control theory; they are present in the team, see for instance the classification aspect in [52] or —although this research has not been active very recently— the study [70] of dynamic feedback and the so-called "flatness" property [65].

3.2. Optimal Control and its Geometry

Let us detail our research program concerning optimal control. Relying on Hamiltonian dynamics is now prevalent, instead of the Lagrangian formalism in classical calculus of variations. The two points of view run parallel when computing geodesics and shortest path in Riemannian Geometry for instance, in that there is a clear one-to-one correspondance between the solutions of the geodesic equation in the tangent bundle and the solution of the Pontryagin Maximum Principle in the cotangent bundle. In most optimal control problems, on the contrary, due to the differential constraints (velocities of feasible trajectories do not cover all directions in the state space), the Lagrangian formalism becomes more involved, while the Pontryagin Maximum Principle keeps the same form, its solutions still live in the cotangent bundle, their projections are the extremals, and a minimizing curve must be the projection of such a solution.

Cut and conjugate loci.   The cut locus —made of the points where the extremals lose optimality— is obviously crucial in optimal control, but usually out of reach (even in low dimensions), and anyway does not have an analytic characterization because it is a non-local object. Fortunately, conjugate points —where the extremals lose *local* optimality— can be effectively computed with high accuracy for many control systems. Elaborating on the seminal work of the Russian and French schools (see [76], [35], [36] and [53] among others), efficient algorithms were designed to treat the smooth case. This was the starting point of a series of papers of members of the team culminating in the outcome of the *cotcot* software [46], followed by the *Hampath* [55] code. Over the years, these codes have allowed for the computation of conjugate loci in a wealth of situations including applications to space mechanics, quantum control, and more recently swimming at low Reynolds number. With in mind the two-dimensional analytic Riemannian framework, a heuristic approach to the global issue of determining cut points is to search for singularities of the conjugate loci; this line is however very delicate to follow on problems stemming from applications in three or more dimensions (see *e.g.* [56] and [43]). In all these situations, the fundamental object underlying the analysis is the curvature tensor. In Hamiltonian terms, one considers the dynamics of subspaces (spanned by Jacobi fields) in the Lagrangian Grassmannian [33]. This point of view withstands generalizations far beyond the smooth case: In L^1-minimization, for instance, discontinuous curves in the Grassmannian have to be considered (instantaneous rotations of Lagrangian subspaces still obeying symplectic rules [60]). The cut locus is a central object in Riemannian geometry, control and optimal transport. This is the motivation for a series of conferences on "The cut locus: A bridge over differential geometry, optimal control, and transport", co-organized by team members and Japanese colleagues, the next one should take place in Nice in 2020.

Riemann and Finsler geometry.   Studying the distance and minimising geodesics in Riemannian Geometry or Finsler Geometry is a particular case of optimal control, simpler because there are no differential constraints; it is studied in the team for the following two reasons. On the one hand, after some tranformations, like averaging (see section 3.2) or reduction, some more difficult optimal control problems lead to a Riemann or Finsler geometry problem. On the other hand, optimal control, mostly the Hamiltonian setting, brings a fresh viewpoint on problems in Riemann and Finsler geometry. On Riemannian ellipsoids of revolution, the optimal control approach allowed to decide on the convexity of the injectivity domain, which, associated with non-negativity of the Ma-Trudinger-Wang curvature tensor, ensures continuity of the optimal transport on the ambient Riemannian manifold [64], [63]. The analysis in the oblate geometry [44] was completed in [59] in the prolate one, including a preliminary analysis of non-focal domains associated with conjugate loci. Averaging in systems coming from space mechanics control (see sections 3.2 and 4.1) with L^2-minimization yields a Riemannian metric, thoroughly computed in [42] together with its geodesic flow; in reduced dimension, its conjugate and cut loci were computed in [45] with Japanese Riemannian geometers. Averaging the same systems for minimum time yields a Finsler Metric, as noted in [41]. In [51], the geodesic convexity properties of these two types of metrics were compared. When perturbations (other than the control) are considered, they introduce a "drift", *i.e.* the Finsler metric is no longer symmetric.

Sub-Riemannian Geometry.   Optimal control problems that pertain to sub-Riemannian Geometry bear all the difficulties of optimal control, like the role of singular/abnormal trajectories, while having some useful structure. They lead to many open problems, like smoothness of minimisers, see the recent monograph [69] for an introduction. Let us detail one open question related to these singular trajectories: the Sard conjecture in sub-Riemannian geometry. Given a totally non-holonomic distribution on a smooth manifold, the Sard Conjecture is concerned with the size of the set of points that can be reached by singular horizontal paths starting from a given point. In the setting of rank-two distributions in dimension three, the Sard conjecture is that this set should be a subset of the so-called Martinet surface, indeed small both in measure and in dimension. In [39], it has been proved that the conjecture holds in the case where the Martinet surface is smooth. Moreover, the case of singular real-analytic Martinet surfaces was also addressed. In this case, it was shown that the Sard Conjecture holds true under an assumption of non-transversality of the distribution on the singular set of the Martinet surface. It is, of course, very intersting to get rid of the remaining technical assumption, or to go to higher dimension. Note that any Sard-type result has strong consequences on the regularity of sub-Riemannian distance functions and in turn on optimal transport problems in the sub-Riemannian setting.

Small controls and conservative systems, averaging.   Using averaging techniques to study small perturbations of integrable Hamiltonian systems is as old an idea as celestial mechanics. It is very subtle in the case of multiple periods but more elementary in the single period case, here it boils down to taking the average of the perturbation along each periodic orbit [37], [75]. This line of research stemmed out of applications to space engineering (see section 4.1): the control of the super-integrable Keplerian motion of a spacecraft orbiting around the Earth is an example of a slow-fast controlled system. Since weak propulsion is used, the control itself acts as a perturbation, among other perturbations of similar magnitudes: higher order terms of the Earth potential (including J_2 effect, first), potential of more distant celestial bodies (such as the Sun and the Moon), atmospheric drag, or even radiation pressure. Properly qualifying the convergence properties (when the small parameter goes to zero) is important and is made difficult by the presence of control. In [41], convergence is seen as convergence to a differential inclusion; this applies to minimum time; a contribution of this work is to put forward the metric character of the averaged system by yielding a Finsler metric (see section 3.2). Proving convergence of the extremals (solutions of the Pontryagin Maximum Principle) is more intricate. In [58], standard averaging ([37], [75]) is performed on the minimum time extremal flow after carefully identifying slow variables of the system thanks to a symplectic reduction. This alternative approach allows to retrieve the previous metric approximation, and to partly address the question of convergence. Under suitable assumptions on a given geodesic of the averaged system (disconjugacy conditions, namely), one proves existence of a family of quasi-extremals for the original system that converge towards the geodesic when the small perturbation parameter goes to zero. This needs to be improved, but convergence of all extremals to extremals of an "averaged Pontryagin Maximum Principle" certainly fails. In particular, one cannot hope for C^1-regularity on the value function when the small parameter goes to zero as swallowtail-like singularities due to the structure of local minima in the problem are expected. (A preliminary analysis has been made in [57].)

Optimality of periodic solutions/periodic controls.   When seeking to minimize a cost with the constraint that the controls and/or part of the states are periodic (and with other initial and final conditions), the notion of conjugate points is more difficult than with straightforward fixed initial point. In [48], for the problem of optimizing the efficiency of the displacement of some micro-swimmers (see section 4.3) with periodic deformations, we used the sufficient optimality conditions established by R. Vinter's group [80], [66] for systems with non unique minimizers due to the existence of a group of symmetry (always present with a periodic minimizer-candidate control). This takes place in a long term collaboration with P. Bettiol (Univ. Bretagne Ouest) on second order sufficient optimality conditions for periodic solutions, or in the presence of higher dimensional symmetry groups, following [80], [66]. Another question relevant to locomotion is the following. Observing animals (or humans), or numerically solving the optimal control problem associated with driftless micro-swimmers for various initial and final conditions, we remark that the optimal strategies of deformation seem to be periodic, at least asymptotically for large distances. This observation is the starting point for characterizing dynamics for which some optimal solutions are periodic, and asymptotically attract

other solutions as the final time grows large; this is reminiscent of the "turnpike theorem" (classical, recently applied to nonlinear situations in [79]).

Software.   These applications (but also the development of theory where numerical experiments can be very enlightening) require many algorithmic and numerical developments that are an important side of the team activity. The software *Hampath* (see section 6.1) is maintained by former members of the team in close collaboration with McTAO. We also use direct discretization approaches (such as the Bocop solver developed by COMMANDS) in parallel. Apart from this, we develop on-demand algorithms and pieces of software, for instance we have to interact with a production software developed by Thales Alenia Space. A strong asset of the team is the interplay of its expertise in geometric control theory with applications and algorithms (see sections 4.1 to 4.3) on one hand, and with optimal transport, and more recently Hamiltonian dynamics, on the other. In 2019, the ADT ct (Control Toolbox) has started with a first sprint in "AMDT mode" with Sophia SED during spring 2019. In addition to McTAO, researchers from the CAGE team (Inria Paris) and the APO team (CNRS Toulouse) are involved. The idea is to put together the efforts on BOCOP and HamPath to go towards a reference toolbox in optimal control. After the first sprint cycle (24 months being planned on the whole action), some starting points have been addressed including: continuous integration for BOCOP and HamPath, refresh on collaborative development tools, first steps of software refactoring, first test of a high-end interface (throught scripting, notebooks, or an *ad hoc* GUI). The next sprint is planned during spring 2020.

3.3. Optimal Transport

Given two measures, and calling transport maps the maps that transport the first measure into the second one, the Monge-Kantorovich problem of Optimal Transport is the search of the minimum of some cost on the set of transport maps. The cost of a map usually comes from some point to point cost and the transport measure. This topic attracted renewed attention in the last decade, and has ongoing applications of many types. Matching optimal transport with geometric control theory is one originality of our team. Let us sketch an important class of open problems. In collaboration with R. McCann [68], we worked towards identifying the costs that admit unique optimizers in the Monge-Kantorovich problem of optimal transport between arbitrary probability densities. For smooth costs and densities on compact manifolds, the only known examples for which the optimal solution is always unique require at least one of the two underlying spaces to be homeomorphic to a sphere. We have introduced a multivalued dynamics induced by the transportation cost between the target and source space, for which the presence or absence of a sufficiently large set of periodic trajectories plays a role in determining whether or not optimal transport is necessarily unique. This insight allows us to construct smooth costs on a pair of compact manifolds with arbitrary topology, so that the optimal transport between any pair of probability densities is unique. We investigated further this problem of uniquely minimizing costs and obtained in collaboration with Abbas Moameni [10] a result of density of uniquely minimizing costs in the C^0-topology. The results in higher topology should be the subject of some further research.

<p style="text-align:center">**NACHOS Project-Team**</p>

3. Research Program

3.1. Scientific foundations

The research activities undertaken by the team aim at developing innovative numerical methodologies putting the emphasis on several features:

- **Accuracy**. The foreseen numerical methods should rely on discretization techniques that best fit to the geometrical characteristics of the problems at hand. Methods based on unstructured, locally refined, even non-conforming, simplicial meshes are particularly attractive in this regard. In addition, the proposed numerical methods should also be capable to accurately describe the underlying physical phenomena that may involve highly variable space and time scales. Both objectives are generally addressed by studying so-called hp-adaptive solution strategies which combine h-adaptivity using local refinement/coarsening of the mesh and p-adaptivity using adaptive local variation of the interpolation order for approximating the solution variables. However, for physical problems involving strongly heterogeneous or high contrast propagation media, such a solution strategy may not be sufficient. Then, for dealing accurately with these situations, one has to design numerical methods that specifically address the multiscale nature of the underlying physical phenomena.

- **Numerical efficiency**. The simulation of unsteady problems most often relies on explicit time integration schemes. Such schemes are constrained by a stability criterion, linking some space and time discretization parameters, that can be very restrictive when the underlying mesh is highly non-uniform (especially for locally refined meshes). For realistic 3D problems, this can represent a severe limitation with regards to the overall computing time. One possible overcoming solution consists in resorting to an implicit time scheme in regions of the computational domain where the underlying mesh size is very small, while an explicit time scheme is applied elsewhere in the computational domain. The resulting hybrid explicit-implicit time integration strategy raises several challenging questions concerning both the mathematical analysis (stability and accuracy, especially for what concern numerical dispersion), and the computer implementation on modern high performance systems (data structures, parallel computing aspects). A second, often considered approach is to devise a local time stepping strategy. Beside, when considering time-harmonic (frequency-domain) wave propagation problems, numerical efficiency is mainly linked to the solution of the system of algebraic equations resulting from the discretization in space of the underlying PDE model. Various strategies exist ranging from the more robust and efficient sparse direct solvers to the more flexible and cheaper (in terms of memory resources) iterative methods. Current trends tend to show that the ideal candidate will be a judicious mix of both approaches by relying on domain decomposition principles.

- **Computational efficiency**. Realistic 3D wave propagation problems involve the processing of very large volumes of data. The latter results from two combined parameters: the size of the mesh i.e the number of mesh elements, and the number of degrees of freedom per mesh element which is itself linked to the degree of interpolation and to the number of physical variables (for systems of partial differential equations). Hence, numerical methods must be adapted to the characteristics of modern parallel computing platforms taking into account their hierarchical nature (e.g multiple processors and multiple core systems with complex cache and memory hierarchies). In addition, appropriate parallelization strategies need to be designed that combine SIMD and MIMD programming paradigms.

From the methodological point of view, the research activities of the team are concerned with four main topics: (1) high order finite element type methods on unstructured or hybrid structured/unstructured meshes for the discretization of the considered systems of PDEs, (2) efficient time integration strategies for dealing with grid induced stiffness when using non-uniform (locally refined) meshes, (3) numerical treatment of complex propagation media models (e.g. physical dispersion models), (4) algorithmic adaptation to modern high performance computing platforms.

3.2. High order discretization methods

3.2.1. *The Discontinuous Galerkin method*

The Discontinuous Galerkin method (DG) was introduced in 1973 by Reed and Hill to solve the neutron transport equation. From this time to the 90's a review on the DG methods would likely fit into one page. In the meantime, the Finite Volume approach (FV) has been widely adopted by computational fluid dynamics scientists and has now nearly supplanted classical finite difference and finite element methods in solving problems of non-linear convection and conservation law systems. The success of the FV method is due to its ability to capture discontinuous solutions which may occur when solving non-linear equations or more simply, when convecting discontinuous initial data in the linear case. Let us first remark that DG methods share with FV methods this property since a first order FV scheme may be viewed as a 0th order DG scheme. However a DG method may also be considered as a Finite Element (FE) one where the continuity constraint at an element interface is released. While keeping almost all the advantages of the FE method (large spectrum of applications, complex geometries, etc.), the DG method has other nice properties which explain the renewed interest it gains in various domains in scientific computing as witnessed by books or special issues of journals dedicated to this method [47]- [48]- [49]- [54]:

- It is naturally adapted to a high order approximation of the unknown field. Moreover, one may increase the degree of the approximation in the whole mesh as easily as for spectral methods but, with a DG method, this can also be done very locally. In most cases, the approximation relies on a polynomial interpolation method but the DG method also offers the flexibility of applying local approximation strategies that best fit to the intrinsic features of the modeled physical phenomena.

- When the space discretization is coupled to an explicit time integration scheme, the DG method leads to a block diagonal mass matrix whatever the form of the local approximation (e.g. the type of polynomial interpolation). This is a striking difference with classical, continuous FE formulations. Moreover, the mass matrix may be diagonal if the basis functions are orthogonal.

- It easily handles complex meshes. The grid may be a classical conforming FE mesh, a non-conforming one or even a hybrid mesh made of various elements (tetrahedra, prisms, hexahedra, etc.). The DG method has been proven to work well with highly locally refined meshes. This property makes the DG method more suitable (and flexible) to the design of some hp-adaptive solution strategy.

- It is also flexible with regards to the choice of the time stepping scheme. One may combine the DG spatial discretization with any global or local explicit time integration scheme, or even implicit, provided the resulting scheme is stable.

- It is naturally adapted to parallel computing. As long as an explicit time integration scheme is used, the DG method is easily parallelized. Moreover, the compact nature of DG discretization schemes is in favor of high computation to communication ratio especially when the interpolation order is increased.

As with standard FE methods, a DG method relies on a variational formulation of the continuous problem at hand. However, due to the discontinuity of the global approximation, this variational formulation has to be defined locally, at the element level. Then, a degree of freedom in the design of a DG method stems from the approximation of the boundary integral term resulting from the application of an integration by parts to the element-wise variational form. In the spirit of FV methods, the approximation of this boundary integral term calls for a numerical flux function which can be based on either a centered scheme or an upwind scheme, or a blending between these two schemes.

3.2.2. High order DG methods for wave propagation models

DG methods are at the heart of the activities of the team regarding the development of high order discretization schemes for the PDE systems modeling electromagnetic and elatsodynamic wave propagation.

- **Nodal DG methods for time-domain problems**. For the numerical solution of the time-domain Maxwell equations, we have first proposed a non-dissipative high order DGTD (Discontinuous Galerkin Time-Domain) method working on unstructured conforming simplicial meshes [9]. This DG method combines a central numerical flux function for the approximation of the integral term at the interface of two neighboring elements with a second order leap-frog time integration scheme. Moreover, the local approximation of the electromagnetic field relies on a nodal (Lagrange type) polynomial interpolation method. Recent achievements by the team deal with the extension of these methods towards non-conforming unstructured [6]-[7] and hybrid structured/unstructured meshes [4], their coupling with hybrid explicit/implicit time integration schemes in order to improve their efficiency in the context of locally refined meshes [3]-[14]-[13]. A high order DG method has also been proposed for the numerical resolution of the elastodynamic equations modeling the propagation of seismic waves [2].

- **Hybridizable DG (HDG) method for time-domain and time-harmonic problems**. For the numerical treatment of the time-harmonic Maxwell equations, nodal DG methods can also be considered [5]. However, such DG formulations are highly expensive, especially for the discretization of 3D problems, because they lead to a large sparse and undefinite linear system of equations coupling all the degrees of freedom of the unknown physical fields. Different attempts have been made in the recent past to improve this situation and one promising strategy has been recently proposed by Cockburn *et al.*[52] in the form of so-called hybridizable DG formulations. The distinctive feature of these methods is that the only globally coupled degrees of freedom are those of an approximation of the solution defined only on the boundaries of the elements. This work is concerned with the study of such Hybridizable Discontinuous Galerkin (HDG) methods for the solution of the system of Maxwell equations in the time-domain when the time integration relies on an implicit scheme, or in the frequency-domain. The team has been a precursor in the development of HDG methods for the frequency-domain Maxwell equations[12].

- **Multiscale DG methods for time-domain problems**. More recently, in collaboration with LNCC in Petropolis (Frédéric Valentin) the framework of the HOMAR assoacite team, we are investigating a family of methods specifically designed for an accurate and efficient numerical treatment of multiscale wave propagation problems. These methods, referred to as Multiscale Hybrid Mixed (MHM) methods, are currently studied in the team for both time-domain electromagnetic and elastodynamic PDE models. They consist in reformulating the mixed variational form of each system into a global (arbitrarily coarse) problem related to a weak formulation of the boundary condition (carried by a Lagrange multiplier that represents e.g. the normal stress tensor in elastodynamic sytems), and a series of small, element-wise, fully decoupled problems resembling to the initial one and related to some well chosen partition of the solution variables on each element. By construction, that methodology is fully parallelizable and recursivity may be used in each local problem as well, making MHM methods belonging to multi-level highly parallelizable methods. Each local problem may be solved using DG or classical Galerkin FE approximations combined with some appropriate time integration scheme (θ-scheme or leap-frog scheme).

3.3. Efficient time integration strategies

The use of unstructured meshes (based on triangles in two space dimensions and tetrahedra in three space dimensions) is an important feature of the DGTD methods developed in the team which can thus easily deal with complex geometries and heterogeneous propagation media. Moreover, DG discretization methods are naturally adapted to local, conforming as well as non-conforming, refinement of the underlying mesh. Most of the existing DGTD methods rely on explicit time integration schemes and lead to block diagonal mass matrices which is often recognized as one of the main advantages with regards to continuous finite element methods.

However, explicit DGTD methods are also constrained by a stability condition that can be very restrictive on highly refined meshes and when the local approximation relies on high order polynomial interpolation. There are basically three strategies that can be considered to cure this computational efficiency problem. The first approach is to use an unconditionally stable implicit time integration scheme to overcome the restrictive constraint on the time step for locally refined meshes. In a second approach, a local time stepping strategy is combined with an explicit time integration scheme. In the third approach, the time step size restriction is overcome by using a hybrid explicit-implicit procedure. In this case, one blends a time implicit and a time explicit schemes where only the solution variables defined on the smallest elements are treated implicitly. The first and third options are considered in the team in the framework of DG [3]-[14]-[13] and HDG discretization methods.

3.4. Numerical treatment of complex material models

Towards the general aim of being able to consider concrete physical situations, we are interested in taking into account in the numerical methodologies that we study, a better description of the propagation of waves in realistic media. In the case of electromagnetics, a typical physical phenomenon that one has to consider is *dispersion*. It is present in almost all media and expresses the way the material reacts to an electromagnetic field. In the presence of an electric field a medium does not react instantaneously and thus presents an electric polarization of the molecules or electrons that itself influences the electric displacement. In the case of a linear homogeneous isotropic media, there is a linear relation between the applied electric field and the polarization. However, above some range of frequencies (depending on the considered material), the dispersion phenomenon cannot be neglected and the relation between the polarization and the applied electric field becomes complex. This is rendered via a frequency-dependent complex permittivity. Several models of complex permittivity exist. Concerning biological media, the Debye model is commonly adopted in the presence of water, biological tissues and polymers, so that it already covers a wide range of applications [11]. In the context of nanoplasmonics, one is interested in modeling the dispersion effects on metals on the nanometer scale and at optical frequencies. In this case, the Drude or the Drude-Lorentz models are generally chosen [17]. In the context of seismic wave propagation, we are interested by the intrinsic attenuation of the medium [15]. In realistic configurations, for instance in sedimentary basins where the waves are trapped, we can observe site effects due to local geological and geotechnical conditions which result in a strong increase in amplification and duration of the ground motion at some particular locations. During the wave propagation in such media, a part of the seismic energy is dissipated because of anelastic losses relied to the internal friction of the medium. For these reasons, numerical simulations based on the basic assumption of linear elasticity are no more valid since this assumption results in a severe overestimation of amplitude and duration of the ground motion, even when we are not in presence of a site effect, since intrinsic attenuation is not taken into account.

3.5. High performance numerical computing

Beside basic research activities related to the design of numerical methods and resolution algorithms for the wave propagation models at hand, the team is also committed to demonstrate the benefits of the proposed numerical methodologies in the simulation of challenging three-dimensional problems pertaining to computational electromagnetics and computational geoseismics. For such applications, parallel computing is a mandatory path. Nowadays, modern parallel computers most often take the form of clusters of heterogeneous multiprocessor systems, combining multiple core CPUs with accelerator cards (e.g Graphical Processing Units - GPUs), with complex hierarchical distributed-shared memory systems. Developing numerical algorithms that efficiently exploit such high performance computing architectures raises several challenges, especially in the context of a massive parallelism. In this context, current efforts of the team are towards the exploitation of multiple levels of parallelism (computing systems combining CPUs and GPUs) through the study of hierarchical SPMD (Single Program Multiple Data) strategies for the parallelization of unstructured mesh based solvers.

TOSCA Team

3. Research Program

3.1. Research Program

Most often physicists, economists, biologists and engineers need a stochastic model because they cannot describe the physical, economical, biological, etc., experiment under consideration with deterministic systems, either because of its complexity and/or its dimension or because precise measurements are impossible. Therefore, they abandon trying to get the exact description of the state of the system at future times given its initial conditions, and try instead to get a statistical description of the evolution of the system. For example, they desire to compute occurrence probabilities for critical events such as the overstepping of a given thresholds by financial losses or neuronal electrical potentials, or to compute the mean value of the time of occurrence of interesting events such as the fragmentation to a very small size of a large proportion of a given population of particles. By nature such problems lead to complex modelling issues: one has to choose appropriate stochastic models, which require a thorough knowledge of their qualitative properties, and then one has to calibrate them, which requires specific statistical methods to face the lack of data or the inaccuracy of these data. In addition, having chosen a family of models and computed the desired statistics, one has to evaluate the sensitivity of the results to the unavoidable model specifications. The TOSCA team, in collaboration with specialists of the relevant fields, develops theoretical studies of stochastic models, calibration procedures, and sensitivity analysis methods.

In view of the complexity of the experiments, and thus of the stochastic models, one cannot expect to use closed form solutions of simple equations in order to compute the desired statistics. Often one even has no other representation than the probabilistic definition (e.g., this is the case when one is interested in the quantiles of the probability law of the possible losses of financial portfolios). Consequently the practitioners need Monte Carlo methods combined with simulations of stochastic models. As the models cannot be simulated exactly, they also need approximation methods which can be efficiently used on computers. The TOSCA team develops mathematical studies and numerical experiments in order to determine the global accuracy and the global efficiency of such algorithms.

The simulation of stochastic processes is not motivated by stochastic models only. The stochastic differential calculus allows one to represent solutions of certain deterministic partial differential equations in terms of probability distributions of functionals of appropriate stochastic processes. For example, elliptic and parabolic linear equations are related to classical stochastic differential equations (SDEs), whereas nonlinear equations such as the Burgers and the Navier–Stokes equations are related to McKean stochastic differential equations describing the asymptotic behavior of stochastic particle systems. In view of such probabilistic representations one can get numerical approximations by using discretization methods of the stochastic differential systems under consideration. These methods may be more efficient than deterministic methods when the space dimension of the PDE is large or when the viscosity is small. The TOSCA team develops new probabilistic representations in order to propose probabilistic numerical methods for equations such as conservation law equations, kinetic equations, and nonlinear Fokker–Planck equations.

<div align="center">

ABS Project-Team

</div>

3. Research Program

3.1. Introduction

The research conducted by ABS focuses on three main directions in Computational Structural Biology (CSB), together with the associated methodological developments:
– Modeling interfaces and contacts,
– Modeling macro-molecular assemblies,
– Modeling the flexibility of macro-molecules,
– Algorithmic foundations.

3.2. Modeling interfaces and contacts

Keywords: Docking, interfaces, protein complexes, structural alphabets, scoring functions, Voronoi diagrams, arrangements of balls.

The Protein Data Bank, http://www.rcsb.org/pdb, contains the structural data which have been resolved experimentally. Most of the entries of the PDB feature isolated proteins [0], the remaining ones being protein - protein or protein - drug complexes. These structures feature what Nature does – up to the bias imposed by the experimental conditions inherent to structure elucidation, and are of special interest to investigate non-covalent contacts in biological complexes. More precisely, given two proteins defining a complex, interface atoms are defined as the atoms of one protein *interacting* with atoms of the second one. Understanding the structure of interfaces is central to understand biological complexes and thus the function of biological molecules [44]. Yet, in spite of almost three decades of investigations, the basic principles guiding the formation of interfaces and accounting for its stability are unknown [47]. Current investigations follow two routes. From the experimental perspective [30], directed mutagenesis enables one to quantify the energetic importance of residues, important residues being termed *hot* residues. Such studies recently evidenced the *modular* architecture of interfaces [41]. From the modeling perspective, the main issue consists of guessing the hot residues from sequence and/or structural informations [36].

The description of interfaces is also of special interest to improve *scoring functions*. By scoring function, two things are meant: either a function which assigns to a complex a quantity homogeneous to a free energy change [0], or a function stating that a complex is more stable than another one, in which case the value returned is a score and not an energy. Borrowing to statistical mechanics [25], the usual way to design scoring functions is to mimic the so-called potentials of mean force. To put it briefly, one reverts Boltzmann's law, that is, denoting $p_i(r)$ the probability of two atoms –defining type i– to be located at distance r, the (free) energy assigned to the pair is computed as $E_i(r) = -kT \log p_i(r)$. Estimating from the PDB one function $p_i(r)$ for each type of pair of atoms, the energy of a complex is computed as the sum of the energies of the pairs located within a distance threshold [45], [32]. To compare the energy thus obtained to a reference state, one may compute $E = \sum_i p_i \log p_i/q_i$, with p_i the observed frequencies, and q_i the frequencies stemming from an a priori model [37]. In doing so, the energy defined is nothing but the Kullback-Leibler divergence between the distributions $\{p_i\}$ and $\{q_i\}$.

Describing interfaces poses problems in two settings: static and dynamic.

---

[0]For structures resolved by crystallography, the PDB contains the asymmetric unit of the crystal. Determining the biological unit from the asymmetric unit is a problem in itself.

[0]The Gibbs free energy of a system is defined by $G = H - TS$, with $H = U + PV$. G is minimum at an equilibrium, and differences in G drive chemical reactions.

In the static setting, one seeks the minimalist geometric model providing a relevant bio-physical signal. A first step in doing so consists of identifying interface atoms, so as to relate the geometry and the bio-chemistry at the interface level [12]. To elaborate at the atomic level, one seeks a structural alphabet encoding the spatial structure of proteins. At the side-chain and backbone level, an example of such alphabet is that of [26]. At the atomic level and in spite of recent observations on the local structure of the neighborhood of a given atom [46], no such alphabet is known. Specific important local conformations are known, though. One of them is the so-called dehydron structure, which is an under-desolvated hydrogen bond – a property that can be directly inferred from the spatial configuration of the C_α carbons surrounding a hydrogen bond [29].

In the dynamic setting, one wishes to understand whether selected (hot) residues exhibit specific dynamic properties, so as to serve as anchors in a binding process [40]. More generally, any significant observation raised in the static setting deserves investigations in the dynamic setting, so as to assess its stability. Such questions are also related to the problem of correlated motions, which we discuss next.

3.3. Modeling macro-molecular assemblies

Keywords: Macro-molecular assembly, reconstruction by data integration, proteomics, modeling with uncertainties, curved Voronoi diagrams, topological persistence.

3.3.1. Reconstruction by Data Integration

Large protein assemblies such as the Nuclear Pore Complex (NPC), chaperonin cavities, the proteasome or ATP synthases, to name a few, are key to numerous biological functions. To improve our understanding of these functions, one would ideally like to build and animate atomic models of these molecular machines. However, this task is especially tough, due to their size and their plasticity, but also due to the flexibility of the proteins involved. In a sense, the modeling challenges arising in this context are different from those faced for binary docking, and also from those encountered for intermediate size complexes which are often amenable to a processing mixing (cryo-EM) image analysis and classical docking. To face these new challenges, an emerging paradigm is that of reconstruction by data integration [24]. In a nutshell, the strategy is reminiscent from NMR and consists of mixing experimental data from a variety of sources, so as to find out the model(s) best complying with the data. This strategy has been in particular used to propose plausible models of the Nuclear Pore Complex [23], the largest assembly known to date in the eukaryotic cell, and consisting of 456 protein *instances* of 30 *types*.

3.3.2. Modeling with Uncertainties and Model Assessment

Reconstruction by data integration requires three ingredients. First, a parametrized model must be adopted, typically a collection of balls to model a protein with pseudo-atoms. Second, as in NMR, a functional measuring the agreement between a model and the data must be chosen. In [22], this functional is based upon *restraints*, namely penalties associated to the experimental data. Third, an optimization scheme must be selected. The design of restraints is notoriously challenging, due to the ambiguous nature and/or the noise level of the data. For example, Tandem Affinity Purification (TAP) gives access to a *pullout* i.e. a list of protein types which are known to interact with one tagged protein type, but no information on the number of complexes or on the stoichiometry of proteins types within a complex is provided. In cryo-EM, the envelope enclosing an assembly is often imprecisely defined, in particular in regions of low density. For immuno-EM labelling experiments, positional uncertainties arise from the microscope resolution.

These uncertainties coupled with the complexity of the functional being optimized, which in general is non convex, have two consequences. First, it is impossible to single out a unique reconstruction, and a set of plausible reconstructions must be considered. As an example, 1000 plausible models of the NPC were reported in [22]. Interestingly, averaging the positions of all balls of a particular protein type across these models resulted in 30 so-called *probability density maps*, each such map encoding the probability of presence of a particular protein type at a particular location in the NPC. Second, the assessment of all models (individual and averaged) is non trivial. In particular, the lack of straightforward statistical analysis of the individual models and the absence of assessment for the averaged models are detrimental to the mechanistic exploitation of the reconstruction results. At this stage, such models therefore remain qualitative.

3.4. Modeling the flexibility of macro-molecules

Keywords: Folding, docking, energy landscapes, induced fit, molecular dynamics, conformers, conformer ensembles, point clouds, reconstruction, shape learning, Morse theory.

Proteins in vivo vibrate at various frequencies: high frequencies correspond to small amplitude deformations of chemical bonds, while low frequencies characterize more global deformations. This flexibility contributes to the entropy thus the *free energy* of the system *protein - solvent*. From the experimental standpoint, NMR studies generate ensembles of conformations, called *conformers*, and so do molecular dynamics (MD) simulations. Of particular interest while investigating flexibility is the notion of correlated motion. Intuitively, when a protein is folded, all atomic movements must be correlated, a constraint which gets alleviated when the protein unfolds since the steric constraints get relaxed [0]. Understanding correlations is of special interest to predict the folding pathway that leads a protein towards its native state. A similar discussion holds for the case of partners within a complex, for example in the third step of the *diffusion - conformer selection - induced fit* complex formation model.

Parameterizing these correlated motions, describing the corresponding energy landscapes, as well as handling collections of conformations pose challenging algorithmic problems.

At the side-chain level, the question of improving rotamer libraries is still of interest [28]. This question is essentially a clustering problem in the parameter space describing the side-chains conformations.

At the atomic level, flexibility is essentially investigated resorting to methods based on a classical potential energy (molecular dynamics), and (inverse) kinematics. A molecular dynamics simulation provides a point cloud sampling the conformational landscape of the molecular system investigated, as each step in the simulation corresponds to one point in the parameter space describing the system (the conformational space) [43]. The standard methodology to analyze such a point cloud consists of resorting to normal modes. Recently, though, more elaborate methods resorting to more local analysis [39], to Morse theory [34] and to analysis of meta-stable states of time series [35] have been proposed.

3.5. Algorithmic foundations

Keywords: Computational geometry, computational topology, optimization, data analysis.

Making a stride towards a better understanding of the biophysical questions discussed in the previous sections requires various methodological developments, which we briefly discuss now.

3.5.1. *Modeling Interfaces and Contacts*

In modeling interfaces and contacts, one may favor geometric or topological information.

On the geometric side, the problem of modeling contacts at the atomic level is tantamount to encoding multi-body relations between an atom and its neighbors. On the one hand, one may use an encoding of neighborhoods based on geometric constructions such as Voronoi diagrams (affine or curved) or arrangements of balls. On the other hand, one may resort to clustering strategies in higher dimensional spaces, as the p neighbors of a given atom are represented by $3p - 6$ degrees of freedom – the neighborhood being invariant upon rigid motions. The information gathered while modeling contacts can further be integrated into interface models.

On the topological side, one may favor constructions which remain stable if each atom in a structure *retains* the same neighbors, even though the 3D positions of these neighbors change to some extent. This process is observed in flexible docking cases, and call for the development of methods to encode and compare shapes undergoing tame geometric deformations.

3.5.2. *Modeling Macro-molecular Assemblies*

In dealing with large assemblies, a number of methodological developments are called for.

---

[0]Assuming local forces are prominent, which in turn subsumes electrostatic interactions are not prominent.

On the experimental side, of particular interest is the disambiguation of proteomics signals. For example, TAP and mass spectrometry data call for the development of combinatorial algorithms aiming at unraveling pairwise contacts between proteins within an assembly. Likewise, density maps coming from electron microscopy, which are often of intermediate resolution (5-10Å) call the development of noise resilient segmentation and interpretation algorithms. The results produced by such algorithms can further be used to guide the docking of high resolutions crystal structures into maps.

As for modeling, two classes of developments are particularly stimulating. The first one is concerned with the design of algorithms performing reconstruction by data integration, a process reminiscent from non convex optimization. The second one encompasses assessment methods, in order to single out the reconstructions which best comply with the experimental data. For that endeavor, the development of geometric and topological models accommodating uncertainties is particularly important.

3.5.3. *Modeling the Flexibility of Macro-molecules*

Given a sampling on an energy landscape, a number of fundamental issues actually arise: how does the point cloud describe the topography of the energy landscape (a question reminiscent from Morse theory)? Can one infer the effective number of degrees of freedom of the system over the simulation, and is this number varying? Answers to these questions would be of major interest to refine our understanding of folding and docking, with applications to the prediction of structural properties. It should be noted in passing that such questions are probably related to modeling phase transitions in statistical physics where geometric and topological methods are being used [38].

From an algorithmic standpoint, such questions are reminiscent of *shape learning*. Given a collection of samples on an (unknown) *model*, *learning* consists of guessing the model from the samples – the result of this process may be called the *reconstruction*. In doing so, two types of guarantees are sought: topologically speaking, the reconstruction and the model should (ideally!) be isotopic; geometrically speaking, their Hausdorff distance should be small. Motivated by applications in Computer Aided Geometric Design, surface reconstruction triggered a major activity in the Computational Geometry community over the past ten years. Aside from applications, reconstruction raises a number of deep issues: the study of distance functions to the model and to the samples, and their comparison; the study of Morse-like constructions stemming from distance functions to points; the analysis of topological invariants of the model and the samples, and their comparison.

ATHENA Project-Team

3. Research Program

3.1. Computational diffusion MRI

Diffusion MRI (dMRI) provides a non-invasive way of estimating in-vivo CNS fiber structures using the average random thermal movement (diffusion) of water molecules as a probe. It's a relatively recent field of research with a history of roughly three decades. It was introduced in the mid 80's by Le Bihan et al [63], Merboldt et al [68] and Taylor et al [78]. As of today, it is the unique non-invasive technique capable of describing the neural connectivity in vivo by quantifying the anisotropic diffusion of water molecules in biological tissues.

3.1.1. Diffusion Tensor Imaging & High Angular Resolution Diffusion Imaging

In dMRI, the acquisition and reconstruction of the diffusion signal allows for the reconstruction of the water molecules displacement probability, known as the Ensemble Average Propagator (EAP) [77], [47]. Historically, the first model in dMRI is the 2nd order diffusion tensor (DTI) [45], [44] which assumes the EAP to be Gaussian centered at the origin. DTI (Diffusion Tensor Imaging) has now proved to be extremely useful to study the normal and pathological human brain [64], [55]. It has led to many applications in clinical diagnosis of neurological diseases and disorder, neurosciences applications in assessing connectivity of different brain regions, and more recently, therapeutic applications, primarily in neurosurgical planning. An important and very successful application of diffusion MRI has been brain ischemia, following the discovery that water diffusion drops immediately after the onset of an ischemic event, when brain cells undergo swelling through cytotoxic edema.

The increasing clinical importance of diffusion imaging has driven our interest to develop new processing tools for Diffusion Tensor MRI. Because of the complexity of the data, this imaging modality raises a large amount of mathematical and computational challenges. We have therefore developed original and efficient algorithms relying on Riemannian geometry, differential geometry, partial differential equations and front propagation techniques to correctly and efficiently estimate, regularize, segment and process Diffusion Tensor MRI (DT-MRI) (see [66] and [65]).

In DTI, the Gaussian assumption over-simplifies the diffusion of water molecules. While it is adequate for voxels in which there is only a single fiber orientation (or none), it breaks for voxels in which there are more complex internal structures and limitates the ability of the DTI to describe complex, singular and intricate fiber configurations (U-shape, kissing or crossing fibers). To overcome this limitation, so-called Diffusion Spectrum Imaging (DSI) [81] and High Angular Resolution Diffusion Imaging (HARDI) methods such as Q-ball imaging [79] and other multi-tensors and compartment models [74], [76], [38], [37], [71] were developed to resolve the orientationnality of more complicated fiber bundle configurations.

Q-Ball imaging (QBI) has been proven very successful in resolving multiple intravoxel fiber orientations in MR images, thanks to its ability to reconstruct the Orientation Distribution Function (ODF, the probability of diffusion in a given direction). These tools play a central role in our work related to the development of a robust and linear spherical harmonic estimation of the HARDI signal and to our development of a regularized, fast and robust analytical QBI solution that outperforms the state-of-the-art ODF numerical technique developed by Tuch [79]. Those contributions are fundamental and have already started to impact on the Diffusion MRI, HARDI and Q-Ball Imaging community [54]. They are at the core of our probabilistic and deterministic tractography algorithms devised to best exploit the full distribution of the fiber ODF (see [51], [3] and [52], [4]).

3.1.2. Beyond DTI with high order tensors

High Order Tensors (HOT) models to estimate the diffusion function while overcoming the shortcomings of the 2nd order tensor model have also been proposed such as the Generalized Diffusion Tensor Imaging (G-DTI) model developed by Ozarslan et al [85], [86] or 4th order Tensor Model [43]. For more details, we refer the reader to our articles in [57], [74] where we review HOT models and to our articles in [65], co-authored with some of our close collaborators, where we review recent mathematical models and computational methods for the processing of Diffusion Magnetic Resonance Images, including state-of-the-art reconstruction of diffusion models, cerebral white matter connectivity analysis, and segmentation techniques. We also worked on Diffusion Kurtosis Imaging (DKI), of great interest for the company OLEA MEDICAL(https://www.olea-medical.com/en). Indeed, DKI is fastly gaining popularity in the domain for characterizing the diffusion propagator or EAP by its deviation from Gaussianity. Hence it is an important clinical tool for characterizing the white-matter's integrity with biomarkers derived from the 3D 4th order kurtosis tensor (KT) [60].

All these powerful techniques are of utmost importance to acquire a better understanding of the CNS mechanisms and have helped to efficiently tackle and solve a number of important and challenging problems [37], [38]. They have also opened up a landscape of extremely exciting research fields for medicine and neuroscience. Hence, due to the complexity of the CNS data and as the magnetic field strength of scanners increases, as the strength and speed of gradients increase and as new acquisition techniques appear [2], these imaging modalities raise a large amount of mathematical and computational challenges at the core of the research we develop at ATHENA [59], [74].

3.1.3. Improving dMRI acquisitions

One of the most important challenges in diffusion imaging is to improve acquisition schemes and analyse approaches to optimally acquire and accurately represent diffusion profiles in a clinically feasible scanning time. Indeed, a very important and open problem in Diffusion MRI is related to the fact that HARDI scans generally require many times more diffusion gradient than traditional diffusion MRI scan times. This comes at the price of longer scans, which can be problematic for children and people with certain diseases. Patients are usually unable to tolerate long scans and excessive motion of the patient during the acquisition process can force a scan to be aborted or produce useless diffusion MRI images. We have developed novel methods for the acquisition and the processing of diffusion magnetic resonance images, to efficiently provide, with just few measurements, new insights into the structure and anatomy of the brain white matter in vivo.

First, we contributed developing real-time reconstruction algorithm based on the Kalman filter [50]. Then, we started to explore the utility of Compressive Sensing methods to enable faster acquisition of dMRI data by reducing the number of measurements, while maintaining a high quality for the results. Compressed Sensing (CS) is a relatively recent technique which has been proved to accurately reconstruct sparse signals from undersampled measurements acquired below the Shannon-Nyquist rate [69].

We have contributed to the reconstruction of the diffusion signal and its important features as the orientation distribution function and the ensemble average propagator, with a special focus on clinical setting in particular for single and multiple Q-shell experiments. Compressive sensing as well as the parametric reconstruction of the diffusion signal in a continuous basis of functions such as the Spherical Polar Fourier basis, have been proved through our contributions to be very useful for deriving simple and analytical closed formulae for many important dMRI features, which can be estimated via a reduced number of measurements [69], [48], [49].

We have also contributed to design optimal acquisition schemes for single and multiple Q-shell experiments. In particular, the method proposed in [2] helps generate sampling schemes with optimal angular coverage for multi-shell acquisitions. The cost function we proposed is an extension of the electrostatic repulsion to multi-shell and can be used to create acquisition schemes with incremental angular distribution, compatible with prematurely stopped scans. Compared to more commonly used radial sampling, our method improves the angular resolution, as well as fiber crossing discrimination. The optimal sampling schemes, freely available for download[0], have been selected for use in the HCP (Human Connectome Project)[0].

---

[0]http://www.emmanuelcaruyer.com/
[0]http://humanconnectome.org/documentation/Q1/imaging-protocols.html

We think that such kind of contributions open new perspectives for dMRI applications including, for example, tractography where the improved characterization of the fiber orientations is likely to greatly and quickly help tracking through regions with and/or without crossing fibers [58].

3.1.4. dMRI modelling, tissue microstructures features recovery & applications

The dMRI signal is highly complex, hence, the mathematical tools required for processing it have to be commensurate in their complexity. Overall, these last twenty years have seen an explosion of intensive scientific research which has vastly improved and literally changed the face of dMRI. In terms of dMRI models, two trends are clearly visible today: the parametric approaches which attempt to build models of the tissue to explain the signal based on model-parameters such as CHARMED [39], AxCaliber [40] and NODDI [82] to cite but a few, and the non-parametric approaches, which attempt to describe the signal in useful but generic functional bases such as the Spherical Polar Fourier (SPF) basis [42], [41], the Solid Harmonic (SoH) basis [53], the Simple Harmonic Oscillator based Reconstruction and Estimation (SHORE) basis [83] and more recent Mean Apparent Propagator or MAP-MRI basis [84].

We propose to investigate the feasibility of using our new models and methods to measure extremely important biological tissue microstructure quantities such as axonal radius and density in white matter. These parameters could indeed provide new insight to better understand the brain's architecture and more importantly could also provide new imaging bio-markers to characterize certain neurodegenerative diseases. This challenging scientific problem, when solved, will lead to direct measurements of important microstructural features that will be integrated in our analysis to provide much greater insight into disease mechanisms, recovery and development. These new microstructural parameters will open the road to go far beyond the limitations of the more simple bio-markers derived from DTI that are clinically used to this date – such as MD (Mean Diffusivity) and FA (Fractional Anisotropy) which are known to be extremely sensitive to confounding factors such as partial volume and axonal dispersion, non-specific and not able to capture any subtle effects that might be early indicators of diseases [5].

3.1.5. Towards microstructural based tractography

In order to go far beyond traditional fiber-tracking techniques, we believe that first order information, i.e. fiber orientations, has to be superseeded by second and third order information, such as microstructure details, to improve tractography. However, many of these higher order information methods are relatively new or unexplored and tractography algorithms based on these high order based methods have to be conceived and designed. In this aim, we propose to work with multiple-shells to reconstruct the Ensemble Average Propagator (EAP), which represents the whole 3D diffusion process and use the possibility it offers to deduce valuable insights on the microstructural properties of the white matter. Indeed, from a reconstructed EAP one can compute the angular features of the diffusion in an diffusion Orientation Distribution Function (ODF), providing insight in axon orientation, calculate properties of the entire diffusion in a voxel such as the Mean Squared Diffusivity (MSD) and Return-To-Origin Probability (RTOP), or come forth with bio-markers detailing diffusion along a particular white matter bundle direction such as the Return-to-Axis or Return-to-Plane Probability (RTAP or RTPP). This opens the way to a ground-breaking computational and unified framework for tractography based on EAP and microstructure features [6]. Using additional a priori anatomical and/or functional information, we could also constrain the tractography algorithm to start and terminate the streamlines only at valid processing areas of the brain.

This development of a computational and unified framework for tractography, based on EAP, microstructure and a priori anatomical and/or functional features, will open new perspectives in tractography, paving the way to a new generation of realistic and biologically plausible algorithms able to deal with intricate configurations of white matter fibers and to provide an exquisite and intrinsic brain connectivity quantification.

3.1.6. Going beyond the state-of-the-art dMRI

Overall, these last twenty years have seen an explosion of intensive scientific research which has vastly improved and literally changed the face of dMRI.

However, although great improvements have been made, major improvements are still required primarily to optimally acquire dMRI data, better understand the biophysics of the signal formation, recover high order invariant and intrinsic microstructure features, identify bio-physically important bio-markers and improve tractography.

Therefore, there is still considerable room for improvement when it comes to the concepts and tools able to efficiently acquire, process and analyze the complex structure of dMRI data. Develop ground-breaking dMRI tools and models for brain connectomics is one of the major objective we would like to achieve in order to take dMRI from the benchside to the bedside and lead to a decisive advance and breakthrough in this field.

3.2. MEG and EEG

Electroencephalography (EEG) and Magnetoencephalography (MEG) are two non-invasive techniques for measuring (part of) the electrical activity of the brain. While EEG is an old technique (Hans Berger, a German neuropsychiatrist, measured the first human EEG in 1929), MEG is a rather new one: the first measurements of the magnetic field generated by the electrophysiological activity of the brain were made in 1968 at MIT by D. Cohen. Nowadays, EEG is relatively inexpensive and is routinely used to detect and qualify neural activities (epilepsy detection and characterisation, neural disorder qualification, BCI, ...). MEG is, comparatively, much more expensive as SQUIDS (Superconducting QUantum Interference Device) only operate under very challenging conditions (at liquid helium temperature) and as a specially shielded room must be used to separate the signal of interest from the ambient noise. However, as it reveals a complementary vision to that of EEG and as it is less sensitive to the head structure, it also bears great hopes and an increasing number of MEG machines are being installed throughout the world. Inria and ODYSSÉE/ATHENA have participated in the acquisition of one such machine installed in the hospital "La Timone" in Marseille.

MEG and EEG can be measured simultaneously (M/EEG) and reveal complementary properties of the electrical fields. The two techniques have temporal resolutions of about the millisecond, which is the typical granularity of the measurable electrical phenomena that arise within the brain. This high temporal resolution makes MEG and EEG attractive for the functional study of the brain. The spatial resolution, on the contrary, is somewhat poor as only a few hundred data points can be acquired simultaneously (about 300-400 for MEG and up to 256 for EEG). MEG and EEG are somewhat complementary with fMRI (Functional MRI) and SPECT (Single-Photon Emission Computed Tomography) in that those provide a very good spatial resolution but a rather poor temporal resolution (of the order of a second for fMRI and a minute for SPECT). Also, contrarily to fMRI, which "only" measures an haemodynamic response linked to the metabolic demand, MEG and EEG measure a direct consequence of the electrical activity of the brain: it is acknowledged that the signals measured by MEG and EEG correspond to the variations of the post-synaptic potentials of the pyramidal cells in the cortex. Pyramidal neurons compose approximately 80% of the neurons of the cortex, and it requires at least about 50,000 active such neurons to generate some measurable signal.

While the few hundred temporal curves obtained using M/EEG have a clear clinical interest, they only provide partial information on the localisation of the sources of the activity (as the measurements are made on or outside of the head). Thus the practical use of M/EEG data raises various problems that are at the core of the ATHENA research in this topic:

- First, as acquisition is continuous and is run at a rate up to 1kHz, the amount of data generated by each experiment is huge. Data selection and reduction (finding relevant time blocks or frequency bands) and pre-processing (removing artifacts, enhancing the signal to noise ratio, ...) are largely done manually at present. Making a better and more systematic use of the measurements is an important step to optimally exploit the M/EEG data [1].

- With a proper model of the head and of the sources of brain electromagnetic activity, it is possible to simulate the electrical propagation and reconstruct sources that can explain the measured signal. Proposing better models [62], [7] and means to calibrate them [80] so as to have better reconstructions are other important aims of our work.

- Finally, we wish to exploit the temporal resolution of M/EEG and to apply the various methods we have developed to better understand some aspects of the brain functioning, and/or to extract more subtle information out of the measurements. This is of interest not only as a cognitive goal, but it also serves the purpose of validating our algorithms and can lead to the use of such methods in the field of Brain Computer Interfaces. To be able to conduct such kind of experiments, an EEG lab has been set up at ATHENA.

3.3. Combined M/EEG and dMRI

dMRI provides a global and systematic view of the long-range structural connectivity within the whole brain. In particular, it allows the recovery of the fiber structure of the white matter which can be considered as the wiring connections between distant cortical areas. These white matter based tractograms are analyzed e.g. to explore the differences in structural connectivity between pathological and normal populations. Moreover, as a by-product, the tractograms can be processed to reveal the nodes of the brain networks, i.e. by segregating together gray matter that share similar connections to the rest of the white matter. But dMRI does not provide information on:

- the cortico-cortical pathways (not passing through white matter) and to some extent, on the short-range connections in the white matter,
- the actual use of connections over time during a given brain activity.

On the opposite, M/EEG measures brain activation over time and provides, after source reconstruction (solving the so-called inverse problem of source reconstruction), time courses of the activity of the cortical areas. Unfortunately, deep brain structures have very little contribution to M/EEG measurements and are thus difficult to analyze. Consequently, M/EEG reveals information about the nodes of the network, but in a more blurry (because of the inverse problem) and fragmented view than dMRI (since it can only reveal brain areas measurable in M/EEG whose activity varies during the experimental protocol). Given its very high temporal resolution, the signal of reconstructed sources can be processed to reveal the functional connectivity between the nodes [75].

While dMRI and M/EEG have been the object of considerable research separately, there have been very few studies on combining the information they provide. Some existing studies deal with the localization of abnormal MEG signals, particularly in the case of epilepsy, and on studying the white matter fibers near the detected abnormal source [67], [70], but to our knowledge there are very few studies merging data coming both from M/EEG and dMRI at the analysis level [72], [56], [46], [73].

Combining the structural and functional information provided by dMRI and M/EEG is a difficult problem as the spatial and temporal resolutions of the two types of measures are extremely different. Still, combining the measurements obtained by these two types of techniques has the great potential of providing a detailed view both in space and time of the functioning brain at a macroscopic level. Consequently, it is a timely and extremely important objective to develop innovative computational tools and models that advance the dMRI and M/EEG state-of-the-art and combine these imaging modalities to build a comprehensive dynamical structural-functional brain connectivity network to be exploited in brain connectivities diseases.

The COBCOM ERC project aims to develop a joint dynamical structural-functional brain connectivity network built on advanced and integrated dMRI and M/EEG ground-breaking methods. To this end, COBCOM will provide new generation of computational dMRI and M/EEG models and methods for identifying and characterizing the connectivities on which the joint network is built. Capitalizing on the strengths of dMRI & M/EEG and building on the bio-physical and mathematical foundations of our models, COBCOM will contribute to create a joint and solid network which will be exploited to identify and characterize white matter abnormalities in some high-impact brain diseases such as Multiple Sclerosis (MS), Epilepsy and mild Traumatic Brain Injury (mTBI).

<div align="center">

BIOCORE Project-Team

</div>

3. Research Program

3.1. Mathematical and computational methods

BIOCORE's action is centered on the mathematical modeling of biological systems, more particularly of artificial ecosystems, that have been built or strongly shaped by human. Indeed, the complexity of such systems where life plays a central role often makes them impossible to understand, control, or optimize without such a formalization. Our theoretical framework of choice for that purpose is Control Theory, whose central concept is "the system", described by state variables, with inputs (action on the system), and outputs (the available measurements on the system). In modeling the ecosystems that we consider, mainly through ordinary differential equations, the state variables are often population, substrate and/or food densities, whose evolution is influenced by the voluntary or involuntary actions of man (inputs and disturbances). The outputs will be some product that one can collect from this ecosystem (harvest, capture, production of a biochemical product, etc), or some measurements (number of individuals, concentrations, etc). Developing a model in biology is however not straightforward: the absence of rigorous laws as in physics, the presence of numerous populations and inputs in the ecosystems, most of them being irrelevant to the problem at hand, the uncertainties and noise in experiments or even in the biological interactions require the development of dedicated techniques to identify and validate the structure of models from data obtained by or with experimentalists.

Building a model is rarely an objective in itself. Once we have checked that it satisfies some biological constraints (eg. densities stay positive) and fitted its parameters to data (requiring tailor-made methods), we perform a mathematical analysis to check that its behavior is consistent with observations. Again, specific methods for this analysis need to be developed that take advantage of the structure of the model (eg. the interactions are monotone) and that take into account the strong uncertainty that is linked to life, so that qualitative, rather than quantitative, analysis is often the way to go.

In order to act on the system, which often is the purpose of our modeling approach, we then make use of two strong points of Control Theory: 1) the development of observers, that estimate the full internal state of the system from the measurements that we have, and 2) the design of a control law, that imposes to the system the behavior that we want to achieve, such as the regulation at a set point or optimization of its functioning. However, due to the peculiar structure and large uncertainties of our models, we need to develop specific methods. Since actual sensors can be quite costly or simply do not exist, a large part of the internal state often needs to be re-constructed from the measurements and one of the methods we developed consists in integrating the large uncertainties by assuming that some parameters or inputs belong to given intervals. We then developed robust observers that asymptotically estimate intervals for the state variables [83]. Using the directly measured variables and those that have been obtained through such, or other, observers, we then develop control methods that take advantage of the system structure (linked to competition or predation relationships between species in bioreactors or in the trophic networks created or modified by biological control).

3.2. A methodological approach to biology: from genes to ecosystems

One of the objectives of BIOCORE is to develop a methodology that leads to the integration of the different biological levels in our modeling approach: from the biochemical reactions to ecosystems. The regulatory pathways at the cellular level are at the basis of the behavior of the individual organism but, conversely, the external stresses perceived by the individual or population will also influence the intracellular pathways. In a modern "systems biology" view, the dynamics of the whole biosystem/ecosystem emerge from the interconnections among its components, cellular pathways/individual organisms/population. The different scales of size and time that exist at each level will also play an important role in the behavior of the biosystem/ecosystem. We intend to develop methods to understand the mechanisms at play at each level,

from cellular pathways to individual organisms and populations; we assess and model the interconnections and influence between two scale levels (eg., metabolic and genetic; individual organism and population); we explore the possible regulatory and control pathways between two levels; we aim at reducing the size of these large models, in order to isolate subsystems of the main players involved in specific dynamical behaviors.

We develop a theoretical approach of biology by simultaneously considering different levels of description and by linking them, either bottom up (scale transfer) or top down (model reduction). These approaches are used on modeling and analysis of the dynamics of populations of organisms; modeling and analysis of small artificial biological systems using methods of systems biology; control and design of artificial and synthetic biological systems, especially through the coupling of systems.

The goal of this multi-level approach is to be able to design or control the cell or individuals in order to optimize some production or behavior at higher level: for example, control the growth of microalgae via their genetic or metabolic networks, in order to optimize the production of lipids for bioenergy at the photobioreactor level.

<p style="text-align:center;color:red;">**BIOVISION Project-Team**</p>

3. Research Program

3.1. Introduction

The Biovision team has started on January 1st, 2016 and became an Equipe Projet Inria on August 1st, 2018 . It aims at developing fundamental research as well as technological developments along two axes.

3.1.1. Axis 1: High tech vision aid-systems for low-vision patients

Visual impairment, also known as vision loss, is a decreased ability to see to a degree that causes problems not fixable by usual means, such as glasses or lenses. Low-vision is a condition caused by eye disease, in which visual acuity is 20/70, meaning that the person is able to see, at 20 meters from a chart, what a normal person would see at 70 meters. Visual impairment affects some 285 million humans in the world, mostly in developed countries where this number is going to increase rapidly due to aging. 85% have low-vision or poorer. [0] There is a strong need to conceive new aid-systems to help these people in their daily living activities. Such systems already exist and can be divided into two categories according to their function. The first category concerns aids that translate visual information into alternative sensory information, such as touch or sound, called Sensory Substitution Devices (SSDs) [45], [40]. The second category concerns aids that adapt visual information to render it more visible to the patients, using scene processing methods and suitable devices. These are based on technological and algorithmic solutions that enhance salient scene characteristics [60], [56]. In Biovision team, we focus on this second category by targeting new vision aid-systems helping patients in their daily life, adapting to their own pathology.

We have strong contacts and collaborations with low-vision centers and associations in order to better understand low-vision patients needs, and have feedback on our prototypes aimed to be distributed to patients via transfer or company creation (startup). With the fast-growing number of incurable eye diseases, crucial steps must be taken to increase visual accessibility by:

- Designing solutions for earlier and more decisive detection of visual pathologies,
- Developing efficient rehabilitation protocols, and,
- Designing innovative vision-aid systems to empower patients with improved perceptual capacities.

To do this, we need to work in synergy with patients to assess their needs, understand their pathologies at a perceptual level and design personalized solutions to create change and adoption. This will require developing state-of-the-art methods in computer science, necessitating skills from many areas such as artificial intelligence, virtual and augmented reality, human-machine interface, multimedia systems, etc. By doing so, we will leverage new technologies to offer life-changing solutions for people with visual impairment [12], [15].

3.1.2. Axis 2: Human vision understanding through joint experimental and modeling studies, for normal and distrophic retinas

A holistic point of view is emerging in neuroscience where one can observe simultaneously how vision works at different levels of the hierarchy in the visual system. Multiple scales functional analysis and connectomics are also exploding in brain science, and studies of visual systems are upfront on this fast move. These integrated studies call for new classes of theoretical and integrated models where the goal is the modeling of visual functions such as motion integration.

---

[0]Source: VisionAware

In Biovision we contribute to a better understanding of the visual system with those main goals:

1. Proposing simplified mathematical models characterizing how the retina converts a visual scene into **spike population coding**, in normal and under specific pathological conditions.
2. Designing biophysical models allowing to better understand the **multiscale dynamics** of the retina, from dynamics of individual cells to their collective activity, and how changes in biophysical parameters (development, pharmacology, pathology) impacts this dynamics.
3. Elaborating an **integrated mathematical and numerical model** of the visual stream, with a focus on motion integration, from retina to early visual cortex (V1).
4. Developing a **simulation platform** emulating the retinal response to visual and prosthetic simulations, enabling us to test hypotheses about the functioning of the early visual system, in normal, pharmacological or pathological conditions.

Finally, although this is not the main goal of our team, two other natural avenues of our research are (i) to develop novel synergistic solutions to solve computer vision tasks based on bio-inspired mechanisms [7]; (ii) collaborate with neuroscientists and neuronal modellers to address mathematical problems outside the scope of the retina or the early visual system.

3.2. Scientific methodology

In this section we briefly describe the scientific methods we use to achieve our research goals.

3.2.1. Adaptive image enhancement

Image enhancement is a natural type of image processing method to help low-vision people better understand visual scenes. An impressive number of techniques have been developed in the fields of computer vision and computer graphics to manipulate image content for a variety of applications. Some of these methods have a direct interest in the design of vision aid-systems. Only a few of them have been carefully evaluated with patients [36], [49], [50], [41], [37]. Our objective is to further exploit and evaluate them with patients, considering dedicated use-cases, using virtual and augmented reality technology (Sec. 3.2.2). We consider not only classical brightness manipulations (e.g., equalization, gamma correction, tone mapping, edge enhancement, image decomposition and cartoonization) but also more sophisticated approaches which can change the geometric information of the scene to highlight the most relevant information (e.g., scene retargeting and seam carving). In addition, we investigate how image enhancements could be adapted to patients needs by relating tuning parameters to the patient pathology.

3.2.2. Virtual, mixed and augmented reality

Virtual, mixed and augmented reality technology (VR/MR/AR) is based on the idea of combining digital worlds with physical realities in different ways. It encompasses a wide spectrum of hardware. It is our conviction that this technology will play a major role in the domain of low-vision. Not only can this technology be useful to design novel vision aid-systems and rehabilitation programs, but also it has the potential to revolutionize how we study the behaviour of low-vision people (controlled condition, free head, eye tracking, possibilities for large scale studies). These projects require a constant interaction with psychophysicists and ophtalmologists so as to design our solutions based on patients needs and capabilities.

3.2.3. Biophysical modeling

Modeling in neuroscience has to cope with several competing objectives. On one hand, describing the biological realm as close as possible, and, on the other hand, providing tractable equations at least at the descriptive level (simulation, qualitative description) and, when possible, at the mathematical level (i.e., affording a rigorous description). These objectives are rarely achieved simultaneously and most of the time one has to make compromises. In the Biovision team we adopt the point of view of a physicist: try to capture the phenomenological description of a biophysical mechanism, removing irrelevant details in the description, and try to have a qualitative description of equations behaviour at least at the numerical simulation level, and, when possible, obtain analytical results. We insist on the quality of the model in predicting and proposing new experiments. This requires a constant interaction with neuroscientists so as to keep the model on the tracks, warning of too crude approximation, still trying to construct equations from canonical principles [1], [2], [6].

3.2.4. Methods from theoretical physics

Biophysical models mainly consist of differential equations (ODEs or PDEs) or integro-differential equations (neural fields). We study them using dynamical systems and bifurcation theory as well as techniques coming from nonlinear physics (amplitude equations, stability analysis, Lyapunov spectrum, correlation analysis, multi-scales methods) [23].

For the study of large scale populations (e.g., when studying population coding) we use methods coming from statistical physics. This branch of physics gave birth to mean-field methods as well statistical methods for large population analysis. We use both of them. Mean-field methods are applied for large scale activity in the retina and in the cortex [4], [8], [39].

For the study of retina population coding we use the so-called Gibbs distribution, initially introduced by Boltzmann and Gibbs. This concept includes, but *is not limited to*, maximum entropy models [55] used by numerous authors in the context of the retina (see, e.g., [57], [59], [52], [51], [61]). These papers were restricted to a statistical description without memory neither causality: the time correlations between successive times is not considered. However, maximum entropy extends to spatio-temporal correlations as we have shown in, e.g., [2] [62], [43]. In this context, we study how the retina respond to transient stimuli (moving objects), i.e. how spatio-temporal correlations are modified when a moving object crosses the receptive fields of ganglion cells, taking into account the lateral connectivity due to amacrine cells [42], [20], [11], [21].

<p style="text-align:center; color:red;">**CAMIN Project-Team**</p>

3. Research Program

3.1. Exploration and understanding of the origins and control of movement

One of CAMIN's areas of expertise is **motion measurement, observation and modeling** in the context of **sensorimotor deficiencies**. The team has the capacity to design advanced protocols to explore motor control mechanisms in more or less invasive conditions in both animal and human.

Human movement can be assessed by several noninvasive means, from motion observation (MOCAP, IMU) to electrophysiological measurements (afferent ENG, EMG, see below). Our general approach is to develop solutions that are realistic in terms of clinical or home use by clinical staff and/or patients for diagnosis and assessment purposes. In doing so, we try to gain a better understanding of motor control mechanisms, including deficient ones, which in turn will give us greater insight into the basics of human motor control. Our ultimate goal is to optimally match a neuroprosthesis to the targeted sensorimotor deficiency.

The team is involved in research projects including:

- Peripheral nervous system (PNS) exploration, modeling and electrophysiology techniques
  Electroneurography (ENG) and electromyography (EMG) signals inform about neural and muscular activities. The team investigates both natural and evoked ENG/EMG through advanced and dedicated signal processing methods. Evoked responses to ES are very precious information for understanding neurophysiological mechanisms, as both the input (ES) and the output (evoked EMG/ENG) are controlled. CAMIN has the expertise to perform animal experiments (rabbits, rats, earthworms and big animals with partners), design hardware and software setups to stimulate and record in harsh conditions, process signals, analyze results and develop models of the observed mechanisms. Experimental surgery is mandatory in our research prior to invasive interventions in humans. It allows us to validate our protocols from theoretical, practical and technical aspects.

- Central nervous system (CNS) exploration
  Stimulating the CNS directly instead of nerves allows activation of the neural networks responsible for generating functions. Once again, if selectivity is achieved the number of implanted electrodes and cables would be reduced, as would the energy demand. We have investigated **spinal electrical stimulation** in animals (pigs) for urinary track and lower limb function management. This work is very important in terms of both future applications and the increase in knowledge about spinal circuitry. The challenges are technical, experimental and theoretical, and the preliminary results have enabled us to test some selectivity modalities through matrix electrode stimulation. This research area will be further intensified in the future as one of ways to improve neuroprosthetic solutions.
  We intend to gain a better understanding of the electrophysiological effects of DES through electroencephalographic (EEG) and electrocorticographic (ECoG) recordings in order to optimize anatomo-functional brain mapping, better understand brain dynamics and plasticity, and improve surgical planning, rehabilitation, and the quality of life of patients.

- Muscle models and fatigue exploration
  Muscle fatigue is one of the major limitations in all FES studies. Simply, the muscle torque varies over time even when the same stimulation pattern is applied. As there is also muscle recovery when there is a rest between stimulations, modeling the fatigue is almost an impossible task. Therefore, it is essential to monitor the muscle state and assess the expected muscle response by FES to improve the current FES system in the direction of greater adaptive force/torque control in the presence of muscle fatigue.

- Movement interpretation

We intend to develop ambulatory solutions to allow ecological observation. We have extensively investigated the possibility of using inertial measurement units (IMUs) within body area networks to observe movement and assess posture and gait variables. We have also proposed extracting gait parameters like stride length and foot-ground clearance for evaluation and diagnosis purposes.

3.2. Movement assistance and/or restoration

The challenges in movement restoration are: (i) improving nerve/muscle stimulation modalities and efficiency and (ii) global management of the function that is being restored in interaction with the rest of the body under voluntary control. For this, both local (muscle) and global (function) controls have to be considered.
Online modulation of ES parameters in the context of lower limb functional assistance requires the availability of information about the ongoing movement. Different levels of complexity can be considered, going from simple open-loop to complex control laws (Figure 2).
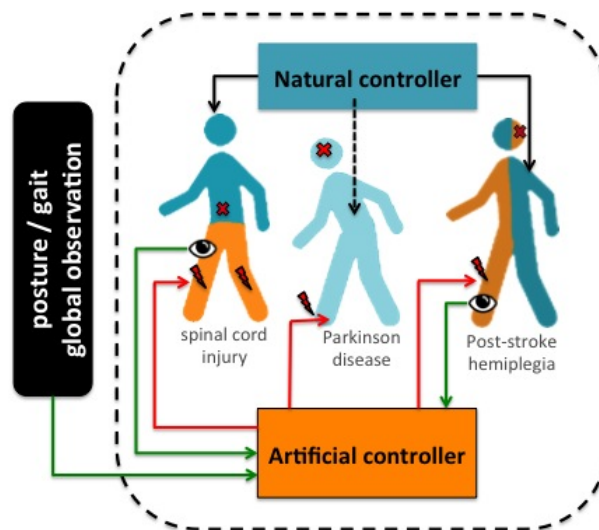

Figure 2. FES assistance should take into account the coexistence of artificial and natural controllers. Artificial controllers should integrate both global (posture/gait) and local (limb/joint) observations.

Real-time adaptation of the stimulation patterns is an important challenge in most of the clinical applications we consider. The modulation of ES parameters to adapt to the occurrence of muscular fatigue or to environement changes needs for advanced adaptive controllers based on sensory information. A special care in minimizing the number of sensors and their impact on patient motion should be taken.

3.3. On-going clinical protocols

One specificity of CAMIN team is to be involved in Clinical protocols. At the moment we are involved in the following protocols:

- CYCLOSEF: Training spinal cord injured people pedaling a tricycle assisted by electric stimulation of sublesional muscles: case study - Protocol RCB 2019-A00808-49. CRF La Châtaigneraie.
- AGILIS - Functional evaluation of the recovery of prehension in quadriplegics by implanted neural stimulation - Protocol RCB 2019-A02037-50. APHP (Paris)

- E-PREHENSTROKE - Evaluation of optimal piloting modalities and their impact on the grasping capacity in Functional Electrical Stimulation of Finger Extensor Muscles in the Hemiplegic Patient in Chronic Phase - Protocol RCB 2018-A02144-51. CHU Nîmes.

- PBREATHLOOP - Recording tracheal sounds for the purpose of developing a breath control algorithm - Protocol RCB 2019-A01813-54. ADOREPS

- Variability and evolution of the single fiber potentials of a spastic muscle treated with botulinum toxin - Protocol RCB 2019-A01863-52A. CHU Nîmes

- Pilot study: measurement of evoked potentials in electroencephalography and electrocorticography by electrical stimulation of the brain during awake neuro-surgery of low-grade infiltrating gliomas - Protocol RCB 2014-A00056-43. CHU Montpellier

CASTOR Project-Team

3. Research Program

3.1. Plasma Physics

Participants: Jacques Blum, Cédric Boulbe, Blaise Faugeras, Hervé Guillard, Holger Heumann, Sebastian Minjeaud, Boniface Nkonga, Richard Pasquetti, Afeintou Sangam.

The main reseach topics are:

1. Modelling and analysis
   – Fluid closure in plasma
   – Turbulence
   – Plasma anisotropy type instabilities
   – Free boundary equilibrium (FBE)
   – Coupling FBE – Transport
2. Numerical methods and simulations
   – High order methods
   – Curvilinear coordinate systems
   – Equilibrium simulation
   – Pressure correction scheme
   – Anisotropy
   – Solving methods and parallelism
3. Identification and control
   – Inverse problem: Equilibrium reconstruction
   – Open loop control
4. Applications
   – MHD instabilities : Edge-Localized Modes (ELMs)
   – Edge plasma turbulence
   – Optimization of scenarii

<p style="text-align:center; color:red;">**COFFEE Project-Team**</p>

3. Research Program

3.1. Research Program

Mathematical modeling and computer simulation are among the main research tools for environmental management, risks evaluation and sustainable development policy. Many aspects of the computer codes as well as the PDEs systems on which these codes are based can be considered as questionable regarding the established standards of applied mathematical modeling and numerical analysis. This is due to the intricate multiscale nature and tremendous complexity of those phenomena that require to set up new and appropriate tools. Our research group aims to contribute to bridging the gap by developing advanced abstract mathematical models as well as related computational techniques.

The scientific basis of the proposal is two–fold. On the one hand, the project is "technically–driven": it has a strong content of mathematical analysis and design of general methodology tools. On the other hand, the project is also "application–driven": we have identified a set of relevant problems motivated by environmental issues, which share, sometimes in a unexpected fashion, many common features. The proposal is precisely based on the conviction that these subjects can mutually cross-fertilize and that they will both be a source of general technical developments, and a relevant way to demonstrate the skills of the methods we wish to design.

To be more specific:

- We consider evolution problems describing highly heterogeneous flows (with different phases or with high density ratio). In turn, we are led to deal with non linear systems of PDEs of convection and/or convection–diffusion type.

- The nature of the coupling between the equations can be two–fold, which leads to different difficulties, both in terms of analysis and conception of numerical methods. For instance, the system can couple several equations of different types (elliptic/parabolic, parabolic/hyperbolic, parabolic or elliptic with algebraic constraints, parabolic with degenerate coefficients....). Furthermore, the unknowns can depend on different sets of variables, a typical example being the fluid/kinetic models for particulate flows. In turn, the simulation cannot use a single numerical approach to treat all the equations. Instead, hybrid methods have to be designed which raise the question of fitting them in an appropriate way, both in terms of consistency of the discretization and in terms of stability of the whole computation. For the problems under consideration, the coupling can also arises through interface conditions. It naturally occurs when the physical conditions are highly different in subdomains of the physical domain in which the flows takes place. Hence interface conditions are intended to describe the exchange (of mass, energy...) between the domains. Again it gives rise to rather unexplored mathematical questions, and for numerics it yields the question of defining a suitable matching at the discrete level, that is requested to preserve the properties of the continuous model.

- By nature the problems we wish to consider involve many different scales (of time or length basically). It raises two families of mathematical questions. In terms of numerical schemes, the multiscale feature induces the presence of stiff terms within the equations, which naturally leads to stability issues. A clear understanding of scale separation helps in designing efficient methods, based on suitable splitting techniques for instance. On the other hand asymptotic arguments can be used to derive hierarchy of models and to identify physical regimes in which a reduced set of equations can be used.

We can distinguish the following fields of expertise

- Numerical Analysis: Finite Volume Schemes, Well-Balanced and Asymptotic-Preserving Methods
  - Finite Volume Schemes for Diffusion Equations and Viscous Flows
  - Finite Volume Schemes for Conservation Laws
  - Well-Balanced and Asymptotic-Preserving Methods
  - Domain Decomposition Methods
- Modeling and Analysis of PDEs
  - Kinetic equations and hyperbolic systems
  - PDEs in random media
  - Interface problems

<p align="center">**EPIONE Project-Team**</p>

3. Research Program

3.1. Introduction

Our research objectives are organized along 5 scientific axes:

1. Biomedical Image Analysis & Machine Learning
2. Imaging & Phenomics, Biostatistics
3. Computational Anatomy, Geometric Statistics
4. Computational Physiology & Image-Guided Therapy
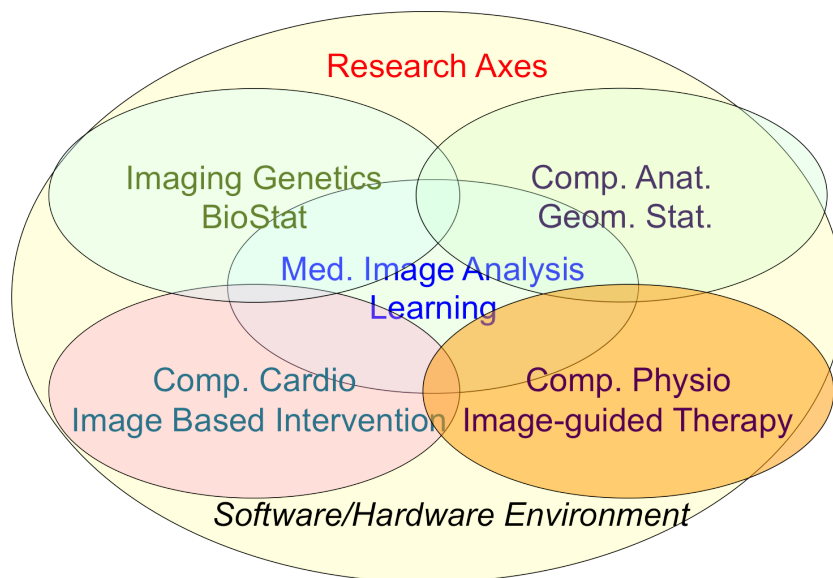5. Computational Cardiology & Image-Based Cardiac Interventions

Figure 3. Epione's five main research axes

For each scientific axis, we introduce the context and the long term vision of our research.

3.2. Biomedical Image Analysis & Machine Learning

The long-term objective of biomedical image analysis is to extract, from biomedical images, pertinent information for the construction of the e-patient and for the development of e-medicine. This relates to the development of advanced segmentation and registration of images, the extraction of image biomarkers of pathologies, the detection and classification of image abnormalities, the construction of temporal models of motion or evolution from time-series of images, etc.

A good illustration of the current state of the art and of the remaining challenges can be found in these recent publications which address for instance the extraction of quantitative biomarkers on static or time varying images, as well as image registration and deformation analysis problems. This also applies to the analysis of microscopic and multi-scale images.

In addition, the growing availability of very large databases of biomedical images, the growing power of computers and the progress of machine learning (ML) approaches have opened up new opportunities for biomedical image analysis.

This is the reason why we decided to revisit a number of biomedical image analysis problems with ML approaches, including segmentation and registration problems, automatic detection of abnormalities, prediction of a missing imaging modality, etc. Not only those ML approaches often outperform the previous state-of-the-art solutions in terms of performances (accuracy of the results, computing times), but they also tend to offer a higher flexibility like the possibility to be transferred from one problem to another one with a similar framework. However, even when successful, ML approaches tend to suffer from a lack of explanatory power, which is particularly annoying for medical applications. We also plan to work on methods that can interpret the results of the ML algorithms that we develop.

- **Revisiting Segmentation problems with Machine Learning:** Through a partnership with Microsoft Research in Cambridge (UK), we are studying new segmentation methods based on deep learning with *weakly annotated* data. In effect, a complete segmentation ground truth is costly to collect in medical image analysis, as it requires the tedious task of contouring regions of interest and their validation by an expert. On the other hand, the label "presence" or "absence" of a lesion for instance (weak annotation) can be obtained at a much lower cost.

  We also plan to explore the application of deep learning methods to the fast segmentation of static or deformable organs. For instance we plan to use deep learning methods for the 3D consistent segmentation of the myocardium tissue of the 2 cardiac ventricles, an important preliminary step to mesh the cardiac muscle for computational anatomy, physiology and cardiology projects.

- **Revisiting Registration problems with Machine Learning:** We are studying, through a partnership with Siemens (Princeton), the possibility to apply robust non-rigid registration through agent-based action learning. We propose a decision process where the objective simplifies to iteratively finding the strategically next best step. An artificial agent is driven to solve the task of non-rigid registration through exploring the parametric space of a statistical deformation model built from training data. Since it is difficult to extract trustworthy ground-truth deformation fields we propose a training scheme with synthetically deformed cases and few real inter-subject cases.

- **Prediction of an imaging modality from other imaging modalities with machine learning:** Through a partnership with the Brain and Stem Institute in Paris, we plan to develop deep learning approaches to quantify some brain alterations currently measured by an invasive nuclear medicine imaging modality (PET imaging with specific tracers), directly from a multi-sequence acquisition of a non-invasive imaging modality (MRI). This requires innovative approaches taking into account the relatively small size of the ground truth database (patients having undergone both PET and MR Image acquisitions) and exploiting the a priori knowledge on the brain anatomy. We believe that this approach could apply to other image prediction problems in the longer term.

- **Prediction of cardiac pathologies with machine learning and image simulation:** Following the important work on cardiac image simulation done during the ERC project MedYMA, we are currently able to simulate time-series of images of various cardiac pathologies for which we can vary the parameters of a generative electro-mechanical model. We plan to develop new deep learning methods exploiting both the *shape* and *motion* phenotypes present in the time-series of images to detect and characterize a number of cardiac pathologies, including subtle assynchronies, local ischemia or infarcts.

- **Measuring Brain, Cognition, Behaviour:** We developed a collaborative project MNC3 which is supported by the excellence initiative Idex UCA^{Jedi}. This project gathers partners from Inria, Nice Hospitals (physicians), Nice University, and IPMC (biologists). The goal is to provide a joint analysis of heterogenous data collected on patients with neurological and psychiatric diseases. Those data include medical imaging (mainly MRI), activity (measured by connected wrists or video or microphones), biology/genomics, and clinical information. We want to show the increase in the statistical power of a joint analysis of the data to classify a pathology and to quantify its evolution.

In addition to these mid-term goals, we have applied to two important projects with local clinicians. A project on "Lung cancer", headed by anatomopathologist P. Hofman, to better exploit the joint information coming from imaging and circulating tumoral cells (in collaboration with Median Tech company); and a project "Cluster headache", headed by neurosurgeon D. Fontaine, to better integrate and exploit information coming from imaging, genetics and clinic (in collaboration with Inria Team Athena).

3.3. Imaging & Phenomics, Biostatistics

The human phenotype is associated with a multitude of heterogeneous biomarkers quantified by imaging, clinical and biological measurements, reflecting the biological and patho-physiological processes governing the human body, and essentially linked to the underlying individual genotype.In order to deepen our understanding of these complex relationships and better identify pathological traits in individuals and clinical groups, a long-term objective of e-medicine is therefore to develop the tools for the joint analysis of this heterogeneous information, termed *Phenomics*, within the unified modeling setting of the e-patient.

Ongoing research efforts aim at investigating optimal approaches at the crossroad between biomedical imaging and bioinformatics to exploit this diverse information. This is an exciting and promising research avenue, fostered by the recent availability of large amounts of data from joint imaging and biological studies (such as the UK biobank [0], ENIGMA [0], ADNI [0],...). However, we currently face important methodological challenges, which limit the ability in detecting and understanding meaningful associations between phenotype and biological information.

To date the most common approach to the analysis of the joint variation between the structure and function of organs represented in medical images, and the classical -omics modalities from biology, such as genomics or lipidomics, is essentially based on the massive univariate statistical testing of single candidate features out of the many available. This is for example the case of genome-wide association studies (GWAS) aimed at identifying statistically significant effects in pools consisting of up to millions of genetics variants. Such approaches have known limitations such as multiple comparison problems, leading to underpowered discoveries of significant associations, and usually explain a rather limited amount of data variance. Although more sophisticated machine learning approaches have been proposed, the reliability and generalization of multivariate methods is currently hampered by the low sample size relatively to the usually large dimension of the parameters space.

To address these issues this research axis investigates novel methods for the integration of this heterogeneous information within a parsimonious and unified multivariate modeling framework. The cornerstone of the project consists in achieving an optimal trade-off between modeling flexibility and ability to generalize on unseen data by developing statistical learning methods informed by prior information, either inspired by "mechanistic" biological processes, or accounting for specific signal properties (such as the structured information from spatio-temporal image time series). Finally, particular attention will be paid to the effective exploitation of the methods in the growing Big Data scenario, either in the meta-analysis context, or for the application in large datasets and biobanks.

- **Modeling associations between imaging, clinical, and biological data.** The essential aspect of this research axis concerns the study of data regularization strategies encoding prior knowledge, for the identification of meaningful associations between biological information and imaging phenotype data. This knowledge can be represented by specific biological mechanisms, such as the complex non-local correlation patterns of the -omics encoded in genes pathways, or by known spatio-temporal relationship of the data (such as time series of biological measurements or images). This axis is based on the interaction with research partners in clinics and biology, such as IPMC (CNRS, France), the Lenval Children's Hospital (France), and University College London (UK). This kind of prior information can be used for defining scalable and parsimonious probabilistic regression models. For example, it can provide relational graphs of data interactions that can be modelled by means of

---

[0] http://www.ukbiobank.ac.uk/
[0] http://enigma.ini.usc.edu/.
[0] http://adni.loni.usc.edu/

Bayesian priors, or can motivate dimensionality reduction techniques and sparse frameworks to limit the effective size of the parameter space. Concerning the clinical application, an important avenue of research will come from the study of the *reduced* representations of the -omics data currently available in clinics, by focusing on the modeling of the disease variants reported in previous genetic findings. The combination of this kind of data with the information routinely available to clinicians, such as medical images and memory tests, has a great potential for leading to improved diagnostic instruments. The translation of this research into clinical practice is carried out thanks to the ongoing collaboration with primary clinical partners such as the University Hospital of Nice (MNC3 partner, France), the Dementia Research Centre of UCL (UK), and the Geneva University Hospital (CH).

- **Learning from collections of biomedical databases.** The current research scenario is characterised by medium/small scale (typically from 50 to 1000 patients) heterogeneous datasets distributed across centres and countries. The straightforward extension of learning algorithms successfully applied to big data problems is therefore difficult, and specific strategies need to be envisioned in order to optimally exploit the available information. To address this problem, we focus on learning approaches to jointly model clinical data localized in different centres. This is an important issue emerging from recent large-scale multi-centric imaging-genetics studies in which partners can only share model parameters (e.g. regression coefficients between specific genes and imaging features), as represented for example by the ENIGMA imaging-genetics study, led by the collaborators at University of Southern California. This problem requires the development of statistical methods for *federated* model estimation, in order to access data hosted in different clinical institutions by simply transmitting the model parameters, that will be in turn updated by using the local available data. This approach is extended to the definition of stochastic optimization strategies in which model parameters are optimized on local datasets, and then summarized in a meta-analysis context. Finally, this project studies strategies for aggregating the information from heterogeneous datasets, accounting for missing modalities due to different study design and protocols. The developed methodology finds important applications within the context of Big Data, for the development of effective learning strategies for massive datasets in the context of medical imaging (such as with the UK biobank), and beyond.

3.4. Computational Anatomy, Geometric Statistics

Computational anatomy is an emerging discipline at the interface of geometry, statistics and image analysis which aims at developing algorithms to model and analyze the biological shape of tissues and organs. The goal is not only to establish generative models of organ anatomies across diseases, populations, species or ages but also to model the organ development across time (growth or aging) and to estimate their variability and link to other functional, genetic or structural information. Computational anatomy is a key component to support computational physiology and is evidently crucial for building the e-patient and to support e-medicine. Pivotal applications include the spatial normalization of subjects in neuroscience (mapping all the anatomies into a common reference system) and atlas to patient registration to map generic knowledge to patient-specific data. Our objectives will be to develop new efficient algorithmic methods to address the emerging challenges described below and to generate precise specific anatomical model in particular for the brain and the heart, but also other organs and structures (e.g. auditory system, lungs, breasts, etc.).

The objects of computational anatomy are often shapes extracted from images or images of labels (segmentation). The observed organ images can also be modeled using registration as the random diffeomorphic deformation of an unknown template (i.e. an orbit).In these cases as in many other applications, invariance properties lead us to consider that these objects belong to non-linear spaces that have a geometric structure. Thus, the mathematical foundations of computational anatomy rely on statistics on non-linear spaces.

- **Geometric Statistics** aim at studying this abstracted problem at the theoretical level. Our goal is to advance the fundamental knowledge in this area, with potential applications to new areas outside of medical imaging. Several challenges which constitute shorter term objectives in this direction are described below.

- **Large databases and longitudinal evolution:** The emergence of larger databases of anatomical images (ADNI, UK biobank) and the increasing availability of temporal evolution drives the need for efficient and scalable statistical techniques. A key issue is to understand how to construct hierarchical models in a non-linear setting.

- **Non-parametric models of variability:** Despite important successes, anatomical data also tend to exhibit a larger variability than what can be modeled with a standard multivariate unimodal Gaussian model. This raises the need for new statistical models to describe the anatomical variability like Bayesian statistics or sample-based statistical model like multi-atlas and archetypal techniques. A second objective is thus to develop efficient algorithmic methods for encoding the statistical variability into models.

- **Intelligible reduced-order models:** Last but not least, these statistical models should live in low dimensional spaces with parameters that can be interpreted by clinicians. This requires of course dimension reduction and variable selection techniques. In this process, it is also fundamental to align the selected variable to a dictionary of clinically meaningful terms (an ontology), so that the statistical model can not only be used to predict but also to explain.

3.4.1. Geometric Statistics

- **Foundations of statistical estimation on geometric spaces:** Beyond the now classical Riemannian spaces, this axis will develop the foundations of statistical estimation on affine connection spaces (e.g. Lie groups), quotient and stratified metric spaces (e.g. orbifolds and tree spaces). In addition to the curvature, one of the key problem is the introduction of singularities at the boundary of the regular strata (non-smooth and non-convex analysis).

- **Parametric and non-parametric dimension reduction methods in non-linear spaces:** The goal is to extend what is currently done with the Fréchet mean (i.e. a 0-dimensional approximation space) to higher dimensional subspaces and finally to a complete hierarchy of embedded subspaces (flags) that iteratively model the data with more and more precision. The Barycentric Subspace Analysis (BSA) generalization of principal component analysis which was recently proposed in the team will of course be a tool of choice for that. In this process, a key issue is to estimate efficiently not only the model parameters (mean point, subspace, flag) but also their uncertainty. Here, we want to quantify the influence of curvature and singularities on non-asymptotic estimation theory since we always have a finite (and often too limited) number of samples. As the mean is generally not unique in curved spaces, this also leads to consider that the results of estimation procedures should be changed from points to singular distributions. A key challenge in developing such a geometrization of statistics will not only be to unify the theory for the different geometric structures, but also to provide efficient practical algorithms to implement them.

- **Learning the geometry from the data:** Data can be efficiently approximated with locally Euclidean spaces when they are very finely sampled with respect to the curvature (big data setting). In the high dimensional low sample size (small data) setting, we believe that invariance properties are essential to reasonably interpolate and approximate. New apparently antagonistic notions like approximate invariance could be the key to this interaction between geometry and learning.

  Beyond the traditional statistical survey of the anatomical shapes that is developed in computational anatomy above, we intend to explore other application fields exhibiting geometric but non-medical data. For instance, applications can be found in Brain-Computer Interfaces (BCI), tree-spaces in phylogenetics, Quantum Physics, etc.

3.5. Computational Physiology & Image-Guided Therapy

Computational Physiology aims at developing computational models of human organ *functions*, an important component of the e-patient, with applications in e-medicine and more specifically in computer-aided prevention, diagnosis, therapy planning and therapy guidance. The focus of our research is on *descriptive* (allowing

to reproduce available observations), *discriminative* (allowing to separate two populations), and above all *predictive models* which can be personalized from patient data including medical images, biosignals, biological information and other available metadata. A key aspect of this scientific axis is therefore the coupling of biophysical models with patient data which implies that we are mostly considering models with relatively few and identifiable parameters. To this end, *data assimilation* methods aiming at estimating biophysical model parameters in order to reproduce available patient data are preferably developed as they potentially lead to predictive models suitable for therapy planning.

Previous research projects in computational physiology have led us to develop biomechanical models representing quasi-static small or large soft tissue deformations (e.g. liver or breast deformation after surgery), mechanical growth or atrophy models (e.g. simulating brain atrophy related to neurodegenerative diseases), heat transfer models (e.g. simulating radiofrequency ablation of tumors), and tumor growth models (e.g. brain or lung tumor growth).

To improve the data assimilation of biophysical models from patient data, a long term objective of our research will be to develop *joint imaging and biophysical generative models in a probabilistic framework* which simultaneously describe the appearance and function of an organ (or its pathologies) in medical images. Indeed, current approaches for the personalization of biophysical models often proceed in two separate steps. In a first stage, geometric, kinematic or/ functional features are first extracted from medical images. In a second stage, they are used by personalization methods to optimize model parameters in order to match the extracted features. In this process, subtle information present in the image which could be informative for biophysical models is often lost which may lead to limited personalization results. Instead, we propose to develop more integrative approaches where the extraction of image features would be performed jointly with the model parameter fitting. Those imaging and biophysical generative models should lead to a *better understanding* of the content of images, to a *better personalization* of model parameters and also *better estimates of their uncertainty*.

This improved coupling between images and model should *help solving various practical problems* driven by clinical applications. Depending on available resources, datasets, and clinical problems, we wish to develop a new expertise for the simulation of *tissue perfusion* (e.g. to capture the uptake of contrast agent or radioactive tracers), or *blood flow in medium / small vessels* (e.g. to capture the transport of drugs or radioactive materials in interventional radiology).

- **Reduced Computational Biophysical Models**. Clinical constraint and uncertainty estimation inevitably lead to the requirement of relatively fast computation of biophysical models. In addition to hardware acceleration (GPU, multithreading) we will explore various ways to accelerate the computation of models through intrusive (e.g. proper orthogonal decomposition, computation of condensed stiffness matrices in non-linear mechanics) or non intrusive methods (e.g. polynomial chaos expansion, Gaussian processes).

- **Uncertainty estimation of Biophysical Models**. We will pursue our research on this topic by developing Bayesian methods to estimate the posterior probability of model parameters, initial and boundary conditions from image features or image voxels. Such approaches rely on the definition of relevant likelihood terms relating the model state variables to the observable quantities in images. When possible joint imaging and biophysical generative models will be developed to avoid to rely on intermediate image features. Approximate inference of uncertainty will be estimated through Variational Bayes approaches whose accuracy will be evaluated through a comparison with stochastic sampling methods (e.g. MCMC). Through this uncertainty estimation, we also aim at developing a reliable framework to select the most sensitive and discriminative parameters of a given model but also to select the biophysical model best suited to solve a given problem (e.g. prediction of therapy outcome).

- **High Order Finite Element Modeling**. Soft tissue biomechanical models have until now been formulated as linear elastic or hyperelastic materials discretized as linear tetrahedra finite elements. While being very generic, those elements are known to suffer from numerical locking for nearly

incompressible materials and lead to poor estimate of stress field. We will develop efficient implementation and assembly methods using high order tetrahedral (and possibly hexahedral) elements. To maintain the number of nodes relatively low while keeping a good accuracy, we intend to develop elements of adaptive degree (p-refinement) driven by local error indices. Solution for meshing surfaces or volumes with curved high order elements will be developed in collaboration with the Titane and Aromath Inria teams.

- **Clinical Applications**. We plan to develop new applications of therapy planning and therapy guidance through existing or emerging collaborations related to the following problems : breast reconstruction following insertion of breast implants (with Anatoscope), planning of cochlear electrodes implantation (with CHU Nice and Oticon Medical), lung deformation following COPD or pulmonary fibrosis (with CHU Nice), echography based elastometry (with CHU Nice).

3.6. Computational Cardiology & Image-Based Cardiac Interventions

Computational Cardiology has been an active research topic within the Computational Anatomy and Computational Physiology axes of the previous Asclepios project, leading to the development of personalized computational models of the heart designed to help characterizing the cardiac function and predict the effect of some device therapies like cardiac resynchronisation or tissue ablation. This axis of research has now gained a lot of maturity and a critical mass of involved scientists to justify an individualized research axis of the new project Epione, while maintaining many constructive interactions with the 4 other research axes of the project. This will develop all the cardiovascular aspects of the e-patient for cardiac e-medicine.

The new challenges we want to address in computational cardiology are related to the introduction of new levels of modeling and to new clinical and biological applications. They also integrate the presence of new sources of measurements and the potential access to very large multimodal databases of images and measurements at various spatial and temporal scales.

Our goal will be to combine two complementary computational approaches: *machine learning* and *biophysical modelling*. This research axis will leverage on the added value of such a combination. Also we will refine our biophysical modeling by the introduction of a pharmacokinetics/pharmacodynamics (PK/PD) component able to describe the effect of a drug on the cardiac function. This will come in complement to the current geometric, electrical, mechanical and hemodynamic components of our biophysical model of the heart. We will also carefully model the uncertainty in our modeling, and try to provide algorithms fast enough to allow future clinical translation.

- Physics of Ultrasound Images for Probe Design: we will design a digital phantom of the human torso in order to help the design of echocardiographic probes. This will be done in collaboration with GE Healthcare whose excellence centre for cardiac ultrasound probes is located in Sophia Antipolis.

- Cardiac Pharmacodynamics for Drug Personalisation: we will add to our biophysical cardiac model a pharmacodynamics model, coupled with a pharmacokinetics model and a personalisation framework in order to help the adjustment of drug therapy to a given patient. This will be done in collaboration with ExactCure, a start up company specialised on this topic.

- New Imaging Modality Coupling MRI and Electrodes: we will use our fast models in order to regularize the ill-posed inverse problem of cardiac electrocardiography in order to estimate cardiac electrical activity from body surface potentials. This will be done within the ERC Starting Grant ECSTATIC coordinated by Hubert Cochet from the IHU Liryc, Bordeaux.

- Cardiac Imaging during Exercise: a particular aspect of the cardiac function is its constant adaptation to satisfy the needs of the human body. This dynamic aspect provides important information on the cardiac function but is challenging to measure. We will set up exercise protocols with Nice University Hospital and STAPS in order to model and quantify such an adaptation of the cardiac function.

- Sudden Cardiac Death is the cause of important mortality (300 000 per year in Europe, same in US) and it is difficult to identify people at risk. Based on a large multi-centric database of images, we will learn the image features correlated with a high risk of arrhythmia, with the IHU Liryc.

- Personalising models from connected objects: with the Internet of Things and the pletora of sensors available today, the cardiac function can be monitored almost continuously. Such new data open up possibilities for novel methods and tools for diagnosis, prognosis and therapy.

LEMON Project-Team

3. Research Program

3.1. Foreword

The team has three main scientific objectives. The first is to develop new models and advanced mathematical methods for inland flow processes. The second is to investigate the derivation and use of coupled models for marine and coastal processes (mainly hydrodynamics, but not only). The third is to develop theoretical methods to be used in the mathematical models serving the first two objectives. As mentioned above, the targeted applications cover PDE models and related extreme events using a hierarchy of models of increasing complexity. LEMON members also contribute to research projects that are not in the core of the team topics and that correspond to external collaborations: they are mentioned in the fourth section below.

In every section, people involved in the project are listed in alphabetical order, except for the first one (underlined) which corresponds to the leading scientist on the corresponding objective.

3.2. Inland flow processes

3.2.1. Shallow water models with porosity

3.2.1.1. State of the Art

Simulating urban floods and free surface flows in wetlands requires considerable computational power. Two-dimensional shallow water models are needed. Capturing the relevant hydraulic detail often requires computational cell sizes smaller than one meter. For instance, meshing a complete urban area with a sufficient accuracy would require 10^6 to 10^8 cells, and simulating one second often requires several CPU seconds. This makes the use of such model for crisis management impossible. Similar issues arise when modelling wetlands and coastal lagoons, where large areas are often connected by an overwhelming number of narrow channels, obstructed by vegetation and a strongly variable bathymetry. Describing such channels with the level of detail required in a 2D model is impracticable. A new generation of models overcoming this issue has emerged over the last 20 years: porosity-based shallow water models. They are obtained by averaging the two-dimensional shallow water equations over large areas containing both water and a solid phase [29]. The size of a computational cell can be increased by a factor 10 to 50 compared to a 2D shallow water model, with CPU times reduced by 2 to 3 orders of magnitude [48]. While the research on porosity-based shallow water models has accelerated over the past decade [43], [59], [62], [39], [38], [48], [73], [74], [68], [69], a number of research issues remain pending.

3.2.1.2. Four year research objectives

The research objectives are (i) to improve the upscaling of the flux and source term models to be embedded in porosity shallow water models, (ii) to validate these models against laboratory and in situ measurements. Improving the upscaled flux and source term models for urban applications requires that description of anisotropy in porosity models be improved to account for the preferential flows induced by building and street alignment. The description of the porosity embedded in the most widespread porosity approach, the so-called Integral Porosity model [59], [41], has been shown to provide an incomplete description of the connectivity properties of the urban medium. Firstly, the governing equations are strongly mesh-dependent because of consistency issues [41]. Secondly, the flux and source term models fail to reproduce the alignment with the main street axes in a number of situations [40]. Another path for improvement concerns the upscaling of obstacle-induced drag terms in the presence of complex geometries. Recent upscaling research results obtained by the LEMON team in collaboration with Tour du Valat suggest that the effects of microtopography on the flow cannot be upscaled using "classical" equation-of-state approaches, as done in most hydraulic models. A totally different approach must be proposed. The next four years will be devoted to the development and validation of improved flux and source term closures in the presence of strongly anisotropic urban geometries

and in the presence of strongly variable topography. Validation will involve not only the comparison of porosity model outputs with refined flow simulation results, but also the validation against experimental data sets. No experimental data set allowing for a sound validation of flux closures in porosity models can be found in the literature. Laboratory experiments will be developed specifically in view of the validation of porosity models. Such experiments will be set up and carried out in collaboration with the Université Catholique de Louvain (UCL), that has an excellent track record in experimental hydraulics and the development of flow monitoring and data acquisition equipment. These activities will take place in the framework of the PoroCity Associate International Laboratory (see next paragraph).

3.2.1.3. People

Vincent Guinot, Carole Delenne, Pascal Finaud-Guyot, Antoine Rousseau.

3.2.1.4. External collaborations

- Tour du Valat (O. Boutron): the partnership with TdV focuses on the development and application of depth-dependent porosity models to the simulation of coastal lagoons, where the bathymetry and geometry is too complex to be represented using refined flow models.

- University of California Irvine (B. Sanders): the collaboration with UCI started in 2014 with research on the representation of urban anisotropic features in integral porosity models [48]. It has led to the development of the Dual Integral Porosity model [42]. Ongoing research focuses on improved representations of urban anisotropy in urban floods modelling.

- Université Catholique de Louvain - UCL (S. Soares-Frazão): UCL is one of the few places with experimental facilities allowing for the systematic, detailed validation of porosity models. The collaboration with UCL started in 2005 and will continue with the PoroCity Associate International Laboratory proposal. In this proposal, a four year research program is set up for the validation, development and parametrization of shallow water models with porosity.

- Luxembourg Institute of Technology (R. Hostache): the collaboration with LIST started in 2018 with the project CASCADE funded by the Fond National de la Recherche du Luxembourg, and the co-direction of Vita Ayoub. The depth-dependant porosity model is applied to simulate the flooding of the Severn river (UK).

3.2.2. Forcing

3.2.2.1. State of the Art

Reproducing optimally realistic spatio-temporal rainfall fields is of salient importance to the forcing of hydrodynamic models. This challenging task requires combining intense, usual and dry weather events. Far from being straightforward, this combination of extreme and non-extreme scenarii requires a realistic modelling of the transitions between normal and extreme periods. [52] have proposed in a univariate framework a statistical model that can serve as a generator and that takes into account low, moderate and intense precipitation. In the same vein, [70] developed a bivariate model. However, its extension to a spatial framework remains a challenge. Existing spatial precipitation stochastic generators are generally based on Gaussian spatial processes [15], [50], that are not adapted to generate extreme rainfall events. Recent advances in spatio-temporal extremes modelling based on generalized Pareto processes [32], [65] and semi-parametric simulation techniques [21] are very promising and could form the base for relevant developments in our framework.

3.2.2.2. Four year research objectives

The purpose is to develop stochastic methods for the simulation of realistic spatio-temporal processes integrating extreme events. Two steps are identified. The first one is about the simulation of extreme events and the second one concerns the combination of extreme and non extreme events in order to build complete, realistic precipitations time series. As far as the first step is concerned, a first task is to understand and to model the space-time structure of hydrological extremes such as those observed in the French Mediterranean basin, that is known for its intense rainfall events (Cevenol episodes), which have recently received increased attention. We will propose modelling approaches based on the exceedance, which allows the simulated fields to be interpreted as events. Parametric, semi-parametric and non-parametric approaches are currently under

consideration. They would allow a number of scientific locks to be removed. Examples of such locks are e.g. accounting for the temporal dimension and for various dependence structures (asymptotic dependence or asymptotic independence possibly depending on the dimension and/or the distance considered). Methodological aspects are detailed in Section 3.4.1 . The second step, which is not straightforward, consists in combining different spatio-temporal simulations in order to help to ultimately develop a stochastic precipitation generator capable of producing full precipitation fields, including dry and non-extreme wet periods.

3.2.2.3. People

Gwladys Toulemonde, Carole Delenne, Vincent Guinot.

3.2.2.4. External collaborations

The Cerise (2016-2018) and Fraise (2019-2021) projects (see 8.2), led by Gwladys Toulemonde, are funded by the action MANU (MAthematical and Numerical methods) of the CNRS LEFE program [0]. Among others, they aim to propose methods for simulating scenarii integrating spatio-temporal extremes fields with a possible asymptotic independence for impact studies in environmental sciences. Among the members of this project, Jean-Noel Bacro (IMAG, UM), Carlo Gaetan (DAIS, Italy) and Thomas Opitz (BioSP, MIA, INRA) are involved in the first step as identified in the research objectives of the present sub-section. Denis Allard (BioSP, MIA, INRA), Julie Carreau (IRD, HSM) and Philippe Naveau (CNRS, LSCE) will be involved in the second one.

3.2.3. Parametrization of shallow water models with porosity

3.2.3.1. State of the Art

Numerical modelling requires data acquisition, both for model validation and for parameter assessment. Model benchmarking against laboratory experiments is an essential step and is an integral part of the team's strategy. However, scale model experiments may have several drawbacks: (i) experiments are very expensive and extremely time-consuming, (ii) experiments cannot always be replicated, and measurement have precision and reliability limitations, (iii) dimensional similarity (in terms of geometry and flow characteristic variables such as Froude or Reynolds numbers) cannot always be preserved.

An ideal way to obtain data would be to carry out in situ measurements. But this would be too costly at the scale of studied systems, not to mention the fact that field may become impracticable during flood periods.

Geographical and remote sensing data are becoming widely available with high spatial and temporal resolutions. Several recent studies have shown that flood extends can be extracted from optical or radar images [35], for example: to characterize the flood dynamics of great rivers [53], to monitor temporary ponds [63], but also to calibrate hydrodynamics models and assess roughness parameters (e.g. [72]).

Upscaled models developed in LEMON (see 3.2.1) embed new parameters that reflect the statistical properties of the medium geometry and the subgrid topography. New methods are thus to be developed to characterize such properties from remote sensing and geographical data.

3.2.3.2. Four year research objectives

This research line consists in deriving methods and algorithms for the determination of upscaled model parameters from geodata.

For applications in urban areas, it is intended to extract information on the porosity parameters from National geographical survey databases largely available in developed countries. Such databases usually incorporate separate layers for roads, buildings, parking lots, yards, etc. Most of the information is stored in vector form, which can be expected to make the treatment of urban anisotropic properties easier than with a raster format. In developing countries, data is made increasingly available over the world thanks to crowdsourcing (e.g. OpenStreetMap) the required level of detail sometimes not available in vector format, especially in suburban areas, where lawns, parks and other vegetated areas, that may also contribute to flood propagation and storage, are not always mapped. In this context, the necessary information can be extracted from aerial and/or satellite images, that are widely available and the spatial resolution of which improves constantly, using supervised classification approaches.

---

[0]Les Enveloppes Fluides et l'Environnement

For applications in great rivers the main objective is to develop an efficient framework for optimally integrating remote sensing derived flood information to compensate the lack of observation related to riverbed bathymetry and river discharge. The effective integration of such remote sensing-derived flood information into hydraulic models remains a critical issue. In partnership with R. Hostache (LIST), we will investigate new ways for making use of SEO data (i.e. flooded areas and water level estimates derived from SAR data collections) for retrieving uncertain model parameters and boundary conditions. The method will be developed and validated using synthetically generated data sets as well as real-event data retrieved from the European Space Agency's archives. Extensive testing will be carried out in a number of high magnitude events recorded over the Severn (United Kingdom) and Zambezi (Mozambique) floodplain areas.

In wetlands applications, connectivity between different ponds is highly dependent on the free surface elevation, thus conditioning the presence of a flow. Characterizing such connectivity requires that topographical variations be known with high accuracy. Despite the increased availability of direct topographic measurements from LiDARS on riverine systems, data collection remains costly when wide areas are involved. Data acquisition may also be difficult when poorly accessible areas are dealt with. If the amount of topographic points is limited, information on elevation contour lines can be easily extracted from the flood dynamics visible in simple SAR or optical images. A challenge is thus to use such data in order to estimate continuous topography on the floodplain combining topographic sampling points and located contour lines the levels of which are unknown or uncertain.

3.2.3.3. People

Carole Delenne, Vincent Guinot, Antoine Rousseau, Pascal Finaud-Guyot

3.2.3.4. External collaborations

- A first attempt for topography reconstruction in wetlands was done in collaboration with J.-S. Bailly (LISAH) in 2016 [30]. It is intended to reactivate this topic in the coming years.
- Porosity model calibration for application on great rivers will be done in the framework of CAS-CADE project in collaboration with R. Hostache (LIST).
- A collaboration started with the LISAH laboratory to investigate the feasibility of depth-dependent porosity laws reconstruction over cultivates areas. LISAH personel involved: D. Feurer, D. Raclot.

3.3. Marine and coastal systems

3.3.1. Multi-scale ocean modelling

The expertise of LEMON in this scientific domain is more in the introduction and analysis of new boundary conditions for ocean modelling systems, that can be tested on academical home-designed test cases. This is in the core of Antoine Rousseau's contributions over the past years. The real implementation, within operational ocean models, has to be done thanks to external collaborations which have already started with LEMON (see below).

3.3.1.1. State of the Art

In physical oceanography, all operational models - regardless of the scale they apply to - are derived from the complete equations of geophysical fluid dynamics. Depending on the considered process properties (nonlinearity, scale) and the available computational power, the original equations are adapted with some simplifying hypotheses. The reader can refer to [58], [51] for a hierarchical presentation of such models.
In the nearshore area, the hydrostatic approximation that is used is most large scales models (high sea) cannot be used without a massive loss of accuracy. In particular, shallow water models are inappropriate to describe the physical processes that occur in this zone (see Figure 1). This is why Boussinesq-type models are prefered: see [49]. They embed dispersive terms that allow for shoaling and other bathymetry effects. Since the pioneering works of Green and Naghdi (see [36]), numerous theoretical and numerical studies have been delivered by the "mathematical oceanography" community, more specifically in France (see the works of Lannes, Marche, Sainte-Marie, Bresch, etc.). The corresponding numerical models (BOSZ, WaveBox) must thus be integrated in any reasonable nearshore modelling platform.
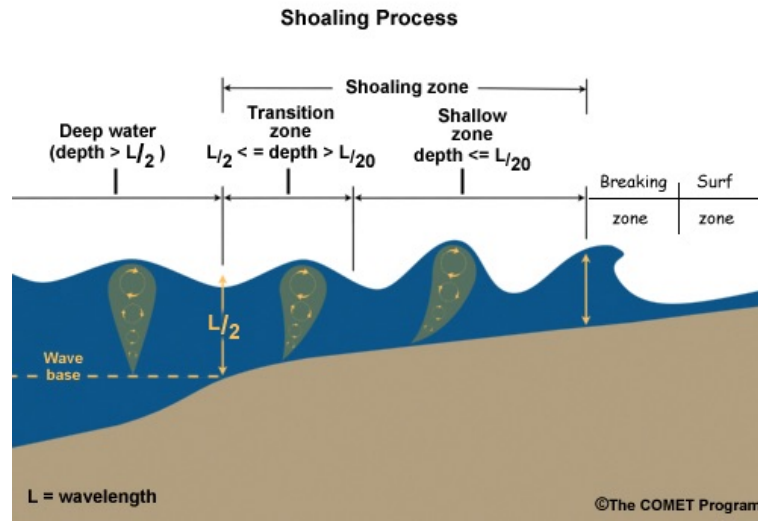
Shoaling Process

Figure 1. Deep sea, shoaling, and breaking zones.

However, these models cannot simply replace all previous models everywhere in the ocean: dispersive models are useless away from the shore and it is known that wave breaking cannot be simulated using Boussinesq-type equations. Hence the need to couple these models with others. Some work has been done in this direction with a multi-level nesting using software packages such as ROMS, but to the best of our knowledge, all the "boxes" rely on the same governing equations with different grid resolutions. A real coupling between different models is a more difficult task since different models may have different mathematical properties, as shown in the work by Eric Blayo and Antoine Rousseau on shallow water modelling (see [17]).

3.3.1.2. Four year research objectives

Starting from the knowledge acquired in the collaboration with Eric Blayo on model coupling using domain decomposition techniques, our ambition is to propose theoretical and numerical tools in order to incorporate nearshore ocean models into large complex systems including several space and time scales. Two complementary research directions are considered:

- **Dispersive *vs* non-dispersive shallow water models**. As depicted in Figure 1 above, Boussinesq-type models (embedding dispersive effects) should be used in the so-called shoaling zone. The coupling with classical deep-sea / shallow water models has to be done such that all the processes in Figure 1 are correctly modelled (by different equations), with a reduced numerical cost. As a first guess, we think that Schwarz-type methods (widely used by the DDM community) could be good candidates, in particular when the interface locations are well-known. Moving interfaces (depending on the flow, the bathymetry and naturally the wind and all external forcings) is a more challenging objective that will be tackled after the first step (known interface) is achieved.

- **spectral *vs* time-domain models**. In the context of mathematical modelling and numerical simulation for the marine energy, we want to build a coupled numerical model that would be able to simulate wave propagation in domains covering both off-shore regions, where spectral models are used, and nearshore regions, better described by nonlinear dispersive (Boussinesq-type) models.
  While spectral models work with a statistical and phase-averaged description of the waves, solving the evolution of its energy spectrum, Boussinesq-type models are phase-resolving and solves nonlinear dispersive shallow water equations for physical variables (surface elevation and velocity)

in the time domain. Furthermore, the time and space scales are very different: they are much larger in the case of spectral models, which justifies their use for modelling off-shore propagation over large time frames. Moreover, important small scale phenomena in nearshore areas are better captured by Boussinesq models, in which the time step is limited by the CFL condition.

From a mathematical and modelling point of view, this task mainly consists in working on the boundary conditions of each model, managing the simultaneous use of spectral and time series data, while studying transparent boundary conditions for the models and developing domain decomposition approaches to improve the exchange of information.

3.3.1.3. People

Antoine Rousseau, Joao Guilherme Caldas Steinstraesser

3.3.1.4. External collaborations

- **Eric Blayo** is the former scientific leader of team MOISE in Grenoble, where Antoine Rousseau was first recruited. Eric Blayo and Antoine Rousseau have co-advised 3 PhDs and continue to work together on coupling methods in hydrodynamics, especially in the framework of the COMODO ANR network.

- **Fabien Marche** (at IMAG, Montpellier, currently on leave in Bordeaux) is an expert in numerical modelling and analysis of Boussinesq-type models. He is the principal investigator of the WaveBox software project, to be embedded in the national scale Uhaina initiative.

- In the framework of its collaboration with **MERIC**, Antoine Rousseau and Joao Guilherme Caldas Steinstraesser collaborate with the consortium DiMe (ANR-FEM project), and more particularly with Jean-François Filipot ans Volker Roeber for the coupling of spectral and time-domain methods.

3.3.2. Data-model interactions

3.3.2.1. State of the Art

An alternative to direct observations is the chaining of numerical models, which for instance represent the physic from offshore to coastal areas. Typically, output data from atmospheric and ocean circulation models are used as forcings for a wave model, which in turn feeds a littoral model. In the case of extreme events, their numerical simulation from physical models is generally unreachable. This is due to a lack of knowledge on boundary conditions and on their physical reliability for such extreme quantities. Based on numerical simulated data, an alternative is to use statistical approaches. [21] proposed such an approach. They first produced and studied a 52-year hindcast using the WW3 wave model [19], [22], [20], [66]. Then stemming from parts of the original work of [18], [37], [32], [21] proposed a semi-parametric approach which aims to simulate extreme space-time waves processes to, in turn, force a littoral hazard model. Nevertheless their approach allows only a very small number of scenarii to be simulated.

3.3.2.2. Four year research objectives

A first objective is to establish the link between the simulation approach proposed by [21] and the Pareto Processes [32]. This will allow the work of [21] to be generalized, thus opening up the possibility of generating an infinity of extreme scenarii. While continuing to favor the semi- or non-parametric approaches made possible by the access to high spatial resolution calculations, we will try to capture the strength of potentially decreasing extremal dependence when moving towards higher values, which requires the development of models that allow for so-called asymptotic independence.

3.3.2.3. People

Gwladys Toulemonde, Fátima Palacios Rodríguez, Antoine Rousseau

3.3.2.4. External collaborations

- since late 2019, LEMON has started a collaboration with IRT Saint-Exupéry on the hybridization of models and large amounts of data for the modelling of urban floods
- The collaboration with Romain Chailan (IMAG, UM, CNRS) and Frédéric Bouchette (Geosciences, UM) started in 2012 during the PhD of Romain entitled Application of scientific computing and statistical analysis to address coastal hazards.
- During her post doctoral position, Fátima Palacios Rodríguez with her co-advisors will considered a generalization of the proposed simulation method by [21].

3.4. Methodological developments

In addition to the application-driven sections, the team also works on the following theoretical questions. They are clearly connected to the abovementioned scientific issues but do not correspond to a specific application or process.

3.4.1. Stochastic models for extreme events

3.4.1.1. State of the Art

Max-stable random fields [61], [60], [46], [26], [54] are the natural limit models for spatial maximum data and have spawned a very rich literature. An overview of typical approaches to modelling maxima is due to [28]. Physical interpretation of simulated data from such models can be discussed. An alternative to the max-stable framework are models for threshold exceedances. Processes called GPD processes, which appear as a generalization of the univariate formalism of the high thresholds exceeding a threshold based on the GPD, have been proposed [32], [65]. Strong advantages of these thresholding techniques are their capability to exploit more information from the data and explicitly model the original event data. However, the asymptotic dependence stability in these limiting processes for maximum and threshold exceedance tends to be overly restrictive when asymptotic dependence strength decreases at high levels and may ultimately vanish in the case of asymptotic independence. Such behaviours appear to be characteristic for many real-world data sets such as precipitation fields [27], [64]. This has motivated the development of more flexible dependence models such as max-mixtures of max-stable and asymptotically independent processes [71], [13] for maxima data, and Gaussian scale mixture processes [55], [45] for threshold exceedances. These models can accommodate asymptotic dependence, asymptotic independence and Gaussian dependence with a smooth transition. Extreme events also generally present a temporal dependence [67] . Developing flexible space-time models for extremes is crucial for characterizing the temporal persistence of extreme events spanning several time steps; such models are important for short-term prediction in applications such as the forecasting of wind power and for extreme event scenario generators providing inputs to impact models, for instance in hydrology and agriculture. Currently, only few models are available from the statistical literature (see for instance [24], [25], [44]) and remain difficult to interpret.

3.4.1.2. Four year research objectives

The objective is to extend state-of-the-art methodology with respect to three important aspects: 1) adapting well-studied spatial modelling techniques for extreme events based on asymptotically justified models for threshold exceedances to the space-time setup; 2) replacing restrictive parametric dependence modelling by semiparametric or nonparametric approaches; 3) proposing more flexible spatial models in terms of asymmetry or in terms of dependence. This means being able to capture the strength of potentially decreasing extremal dependence when moving towards higher values, which requires developing models that allow for so-called asymptotic independence.

3.4.1.3. People

Gwladys Toulemonde, Fátima Palacios Rodríguez

3.4.1.4. External collaborations

In a natural way, the Cerise and Fraise project members are the main collaborators for developing and studying new stochastic models for extremes.

- More specifically, research with Jean-Noel Bacro (IMAG, UM), Carlo Gaetan (DAIS, Italy) and Thomas Opitz (BioSP, MIA, INRA) focuses on relaxing dependence hypothesis.

- The asymmetry issue and generalization of some Copula-based models are studied with Julie Carreau (IRD, HydroSciences, UM).

3.4.2. Integrating heterogeneous data

3.4.2.1. State of the Art

Assuming that a given hydrodynamic models is deemed to perform satisfactorily, this is far from being sufficient for its practical application. Accurate information is required concerning the overall geometry of the area under study and model parametrization is a necessary step towards the operational use. When large areas are considered, data acquisition may turn out prohibitive in terms of cost and time, not to mention the fact that information is sometimes not accessible directly on the field. To give but one example, how can the roughness of an underground sewer pipe be measured? A strategy should be established to benefit from all the possible sources of information in order to gather data into a geographical database, along with confidence indexes.

The assumption is made that even hardly accessible information often exists. This stems from the increasing availability of remote-sensing data, to the crowd-sourcing of geographical databases, including the inexhaustible source of information provided by the Internet. However, information remains quite fragmented and stored in various formats: images, vector shapes, texts, etc.

This path of research begun with the Cart'Eaux project (2015-2018), that aims to produce regular and complete mapping of urban wastewater system. Contrary to drinkable water networks, the knowledge of sewer pipe location is not straightforward, even in developed countries. Over the past century, it was common practice for public service providers to install, operate and repair their networks separately [57]. Now local authorities are confronted with the task of combining data produced by different parts, having distinct formats, variable precision and granularity [23].

3.4.2.2. Four year research objectives

The overall objective of this research line is to develop methodologies to gather various types of data in the aim of producing an accurate mapping of the studied systems for hydrodynamics models.

Concerning wastewater networks, the methodology applied consists in inferring the shape of the network from a partial dataset of manhole covers that can be detected from aerial images [56]. Since manhole covers positions are expected to be known with low accuracy (positional uncertainty, detection errors), a stochastic algorithm is set up to provide a set of probable network geometries [4]. As more information is required for hydraulic modelling than the simple mapping of the network (slopes, diameters, materials, etc.), text mining techniques such as used in [47] are particularly interesting to extract characteristics from data posted on the Web or available through governmental or specific databases. Using an appropriate keyword list, thematic entities are identified and linked to the surrounding spatial and temporal entities in order to ease the burden of data collection. It is clear at this stage that obtaining numerical values on specific pipes will be challenging. Thus, when no information is found, decision rules will be used to assign acceptable numerical values to enable the final hydraulic modelling.

In any case, the confidence associated to each piece of data, be it directly measured or reached from a roundabout route, should be assessed and taken into account in the modelling process. This can be done by generating a set of probable inputs (geometry, boundary conditions, forcing, etc.) yielding simulation results along with the associated uncertainty.

Combining heterogeneous data for a better knowledge of studied systems raises the question of data fusion. What is the reality when contradictory information is collected from different sources? Dealing with spatial information, offset are quite frequent between different geographical data layers; pattern comparison approaches should be developed to judge whether two pieces of information represented by two elements close to each other are in reality identical, complementary, or contradictory.

3.4.2.3. People

<u>Carole Delenne</u>, Vincent Guinot, Antoine Rousseau, Gwladys Toulemonde

3.4.2.4. External collaborations

The Cart'Eaux project has been a lever to develop a collaboration with Berger-Levrault company and several multidisciplinary collaborations for image treatment (LIRMM), text analysis (LIRMM and TETIS) and network cartography (LISAH, IFSTTAR).

- The MeDo project lead by N. Chahinian (HSM) in collaboration with linguists of UMR Praxiling, uses data mining and text analysis approaches to retrieve information on wastewater networks from the Web. Carole Delenne has a slight implication in this project, as domain expert to guide the text annotations and for the uncertainties definition and representation in the mapping of the data collected.
- Concerning geographical data fusion for the wastewater network cartograhpy, the Phd thesis of Yassine Bel-Ghaddar has been funded by the French Association of Research and Technology (ANRT) in collaboration with Berger-Levrault company and in co-direction with A. Begdouri (LSIA Fès, Morocco).

3.4.3. Numerical methods for porosity models

3.4.3.1. State of the Art

Porosity-based shallow water models are governed by hyperbolic systems of conservation laws. The most widespread method used to solve such systems is the finite volume approach. The fluxes are computed by solving Riemann problems at the cell interfaces. This requires that the wave propagation properties stemming from the governing equations be known with sufficient accuracy. Most porosity models, however, are governed by non-standard hyperbolic systems.

Firstly, the most recently developed DIP models include a momentum source term involving the divergence of the momentum fluxes [42]. This source term is not active in all situations but takes effect only when positive waves are involved [39], [40]. The consequence is a discontinuous flux tensor and discontinuous wave propagation properties. The consequences of this on the existence and uniqueness of solutions to initial value problems (especially the Riemann problem) are not known, or are the consequences on the accuracy of the numerical methods used to solve this new type of equations.

Secondly, most applications of these models involve anisotropic porosity fields [48], [59]. Such anisotropy can be modelled using 2×2 porosity tensors, with principal directions that are not aligned with those of the Riemann problems in two dimensions of space. The solution of such Riemann problems has not been investigated yet. Moreover, the governing equations not being invariant by rotation, their solution on unstructured grids is not straightforward.

Thirdly, the Riemann-based, finite volume solution of the governing equations require that the Riemann problem be solved in the presence of a porosity discontinuity. While recent work [31] has addressed the issue for the single porosity equations, similar work remains to be done for integral- and multiple porosity-based models.

3.4.3.2. Four year research objectives

The four year research objectives are the following:

- investigate the properties of the analytical solutions of the Riemann problem for a continuous, anisotropic porosity field,
- extend the properties of such analytical solutions to discontinuous porosity fields,
- derive accurate and CPU-efficient approximate Riemann solvers for the solution of the conservation form of the porosity equations.

3.4.3.3. People

Vincent Guinot, Pascal Finaud-Guyot

3.4.3.4. External collaborations

Owing to the limited staff of the LEMON team, external collaborations will be sought with researchers in applied mathematics. Examples of researchers working in the field are

- Minh Le, Saint Venant laboratory, Chatou (France): numerical methods for shallow water flows, experience with the 2D, finite element/finte volume-based Telemac2D system.

- M.E. Vazquez-Cendon, Univ. Santiago da Compostela (Spain): finite volume methods for shallow water hydrodynamics and transport, developed Riemann solvers for the single porosity equations.

- A. Ferrari, R. Vacondio, S. Dazzi, P. Mignosa, Univ. Parma (Italy): applied mathematics, Riemann solvers for the single porosity equations.

- O. Delestre, Univ. Nice-Sophia Antipolis (France): development of numerical methods for shallow water flows (source term treatment, etc.)

- F. Benkhaldoun, Univ. Paris 13 (France): development of Riemann solvers for the porous shallow water equations.

.

3.4.4. Inland hydrobiological systems

3.4.4.1. State of the Art

Water bodies such as lakes or coastal lagoons (possibly connected to the sea) located in high human activity areas are subject to various kinds of stress such as industrial pollution, high water demand or bacterial blooms caused by freshwater over-enrichment. For obvious environmental reasons, these water resources have to be protected, hence the need to better understand and possibly control such fragile ecosystems to eventually develop decision-making tools. From a modelling point of view, they share a common feature in that they all involve interacting biological and hydrological processes. According to [33], models may be classified into two main types: "minimal dynamic models" and "complex dynamic models". These two model types do not have the same objectives. While the former are more heuristic and rather depict the likelihood of considered processes, the latter are usually derived from fundamental laws of biochemistry or fluid dynamics. Of course, the latter necessitate much more computational resources than the former. In addition, controlling such complex systems (usually governed by PDEs) is by far more difficult that controlling the simpler ODE-driven command systems.

LEMON has already contributed both to the reduction of PDE models for the simulation of water confinement in coastal lagoons [34], [16] and to the improvement of ODE models in order to account for space-heterogeneity of bioremediation processes in water resources [14].

3.4.4.2. Four year research objectives

In collaboration with colleagues from the ANR-ANSWER project and colleagues from INRA, our ambition is to improve existing models of lagoon/marine ecosystems by integrating both accurate and numerically affordable coupled hydrobiological systems. A major challenge is to find an optimal trade-off between the level of detail in the description of the ecosystem and the level of complexity in terms of number of parameters (in particular regarding the governing equations for inter-species reactions). The model(s) should be able to reproduce the inter-annual variability of the observed dynamics of the ecosystem in response to meteorological forcing. This will require the adaptation of hydrodynamics equations to such time scales (reduced/upscaled models such as porosity shallow water models (see Section 3.2.1) will have to be considered) together with the coupling with the ecological models. At short time scales (i.e. the weekly time scale), accurate (but possibly CPU-consuming) 3D hydrodynamic models processes (describing thermal stratification, mixing, current velocity, sediment resuspension, wind waves...) are needed. On the longer term, it is intended to develop reduced models accounting for spatial heterogeneity.

The team will focus on two main application projects in the coming years:

- the ANR ANSWER project (2017-2021, with INRA Montpellier and LEESU) focusing on the cyanobacteria dynamics in lagoons and lakes. A PhD student is co-advised by Antoine Rousseau in collaboration with Céline Casenave (INRA, Montpellier).

- the long term collaboration with Alain Rapaport (INRA Montpellier) will continue both on the bioremediation of water resources such as the Tunquen lagoon in Chile and with a new ongoing project on water reuse (converting wastewater into water that can be reused for other purposes such as irrigation of agricultural fields). Several projects are submitted to the ANR and local funding structures in Montpellier.

3.4.4.3. People

Céline Casenave (INRA Montpellier), Antoine Rousseau, Vincent Guinot, Joseph Luis Kahn Casapia.

3.4.4.4. External collaborations

- ANR ANSWER consortium: Céline Casenave (UMR MISTEA, INRA Montpellier), Brigitte Vinçon-Leite (UM LEESU, ENPC), Jean-François Humbert (UMR IEES, UPMC). ANSWER is a French-Chinese collaborative project that focuses on the modelling and simulation of eutrophic lake ecosystems to study the impact of anthropogenic environmental changes on the proliferation of cyanobacteria. Worldwide the current environmental situation is preoccupying: man-driven water needs increase, while the quality of the available resources is deteriorating due to pollution of various kinds and to hydric stress. In particular, the eutrophication of lentic ecosystems due to excessive inputs of nutrients (phosphorus and nitrogen) has become a major problem because it promotes cyanobacteria blooms, which disrupt the functioning and the uses of the ecosystems.

- A. Rousseau has a long lasting collaboration with Alain Rapaport (UMR MISTEA, INRA Montpellier) and Héctor Ramirez (CMM, Université du Chili).

<div align="center">

MATHNEURO Project-Team

</div>

3. Research Program

3.1. Neural networks dynamics

The study of neural networks is certainly motivated by the long term goal to understand how brain is working. But, beyond the comprehension of brain or even of simpler neural systems in less evolved animals, there is also the desire to exhibit general mechanisms or principles at work in the nervous system. One possible strategy is to propose mathematical models of neural activity, at different space and time scales, depending on the type of phenomena under consideration. However, beyond the mere proposal of new models, which can rapidly result in a plethora, there is also a need to understand some fundamental keys ruling the behaviour of neural networks, and, from this, to extract new ideas that can be tested in real experiments. Therefore, there is a need to make a thorough analysis of these models. An efficient approach, developed in our team, consists of analysing neural networks as dynamical systems. This allows to address several issues. A first, natural issue is to ask about the (generic) dynamics exhibited by the system when control parameters vary. This naturally leads to analyse the bifurcations [70] [71] occurring in the network and which phenomenological parameters control these bifurcations. Another issue concerns the interplay between the neuron dynamics and the synaptic network structure.

3.2. Mean-field and stochastic approaches

Modeling neural activity at scales integrating the effect of thousands of neurons is of central importance for several reasons. First, most imaging techniques are not able to measure individual neuron activity (microscopic scale), but are instead measuring mesoscopic effects resulting from the activity of several hundreds to several hundreds of thousands of neurons. Second, anatomical data recorded in the cortex reveal the existence of structures, such as the cortical columns, with a diameter of about $50\mu m$ to $1mm$, containing of the order of one hundred to one hundred thousand neurons belonging to a few different species. The description of this collective dynamics requires models which are different from individual neurons models. In particular, when the number of neurons is large enough averaging effects appear, and the collective dynamics is well described by an effective mean-field, summarizing the effect of the interactions of a neuron with the other neurons, and depending on a few effective control parameters. This vision, inherited from statistical physics requires that the space scale be large enough to include a large number of microscopic components (here neurons) and small enough so that the region considered is homogeneous.

Our group is developing mathematical and numerical methods allowing on one hand to produce dynamic mean-field equations from the physiological characteristics of neural structure (neurons type, synapse type and anatomical connectivity between neurons populations), and on the other so simulate these equations; see Figure 1 . These methods use tools from advanced probability theory such as the theory of Large Deviations [59] and the study of interacting diffusions [3].

3.3. Neural fields

Neural fields are a phenomenological way of describing the activity of population of neurons by delayed integro-differential equations. This continuous approximation turns out to be very useful to model large brain areas such as those involved in visual perception. The mathematical properties of these equations and their solutions are still imperfectly known, in particular in the presence of delays, different time scales and noise.
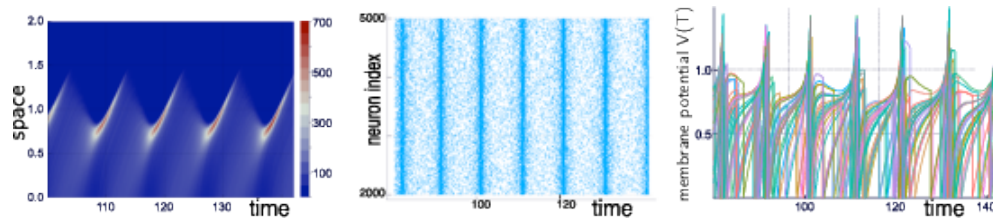
Figure 1. Simulations of the quasi-synchronous state of a stochastic neural network with $N = 5000$ neurons. Left: empirical distribution of membrane potential as a function (t, v). Middle: (raster plot) spiking times as a function of neuron index and time. Right: several membrane potentials $v_i(t)$ as a function of time for $i \in [1, 100]$. Simulated with the Julia Package PDMP.jl from [17]. This figure has been slightly modified from [11].

Our group is developing mathematical and numerical methods for analysing these equations. These methods are based upon techniques from mathematical functional analysis, bifurcation theory [14], [72], equivariant bifurcation analysis, delay equations, and stochastic partial differential equations. We have been able to characterize the solutions of these neural fields equations and their bifurcations, apply and expand the theory to account for such perceptual phenomena as edge, texture [50], and motion perception. We have also developed a theory of the delayed neural fields equations, in particular in the case of constant delays and propagation delays that must be taken into account when attempting to model large size cortical areas [16], [73]. This theory is based on center manifold and normal forms ideas [15].

3.4. Slow-fast dynamics in neuronal models

Neuronal rhythms typically display many different timescales, therefore it is important to incorporate this slow-fast aspect in models. We are interested in this modeling paradigm where slow-fast point models, using Ordinary Differential Equations (ODEs), are investigated in terms of their bifurcation structure and the patterns of oscillatory solutions that they can produce. To insight into the dynamics of such systems, we use a mix of theoretical techniques — such as geometric desingularisation and centre manifold reduction [63] — and numerical methods such as pseudo-arclength continuation [54]. We are interested in families of complex oscillations generated by both mathematical and biophysical models of neurons. In particular, so-called *mixed-mode oscillations (MMOs)* [9], [52], [62], which represent an alternation between subthreshold and spiking behaviour, and *bursting oscillations* [53], [60], also corresponding to experimentally observed behaviour [51]; see Figure 2 . We are working on extending these results to spatio-temporal neural models [2].

3.5. Modeling neuronal excitability

Excitability refers to the all-or-none property of neurons [58], [61]. That is, the ability to respond nonlinearly to an input with a dramatic change of response from "none" — no response except a small perturbation that returns to equilibrium — to "all" — large response with the generation of an action potential or spike before the neuron returns to equilibrium. The return to equilibrium may also be an oscillatory motion of small amplitude; in this case, one speaks of resonator neurons as opposed to integrator neurons. The combination of a spike followed by subthreshold oscillations is then often referred to as mixed-mode oscillations (MMOs) [52]. Slow-fast ODE models of dimension at least three are well capable of reproducing such complex neural oscillations. Part of our research expertise is to analyse the possible transitions between different complex oscillatory patterns of this sort upon input change and, in mathematical terms, this corresponds to understanding the bifurcation structure of the model. Furthermore, the shape of time series of this sort with a given oscillatory pattern can be analysed within the mathematical framework of dynamic bifurcations; see the section on slow-fast dynamics in Neuronal Models. The main example of abnormal neuronal excitability is hyperexcitability
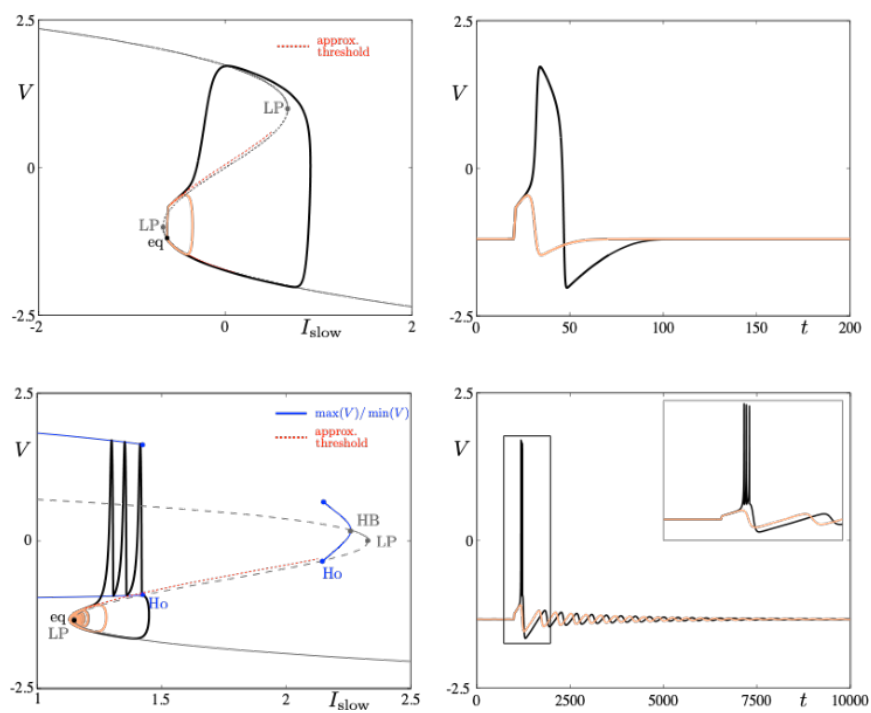
Figure 2. Excitability threshold as slow manifolds in a simple spiking model, namely the FitzHugh-Nagumo model, (top panels) and in a simple bursting model, namely the Hindmarsh-Rose model (bottom panels). This figure is unpublished.

and it is important to understand the biological factors which lead to such excess of excitability and to identify (both in detailed biophysical models and reduced phenomenological ones) the mathematical structures leading to these anomalies. Hyperexcitability is one important trigger for pathological brain states related to various diseases such as chronic migraine [65], epilepsy [75] or even Alzheimer's Disease [64]. A central central axis of research within our group is to revisit models of such pathological scenarios, in relation with a combination of advanced mathematical tools and in partnership with biological labs.

3.6. Synaptic Plasticity

Neural networks show amazing abilities to evolve and adapt, and to store and process information. These capabilities are mainly conditioned by plasticity mechanisms, and especially synaptic plasticity, inducing a mutual coupling between network structure and neuron dynamics. Synaptic plasticity occurs at many levels of organization and time scales in the nervous system [49]. It is of course involved in memory and learning mechanisms, but it also alters excitability of brain areas and regulates behavioral states (e.g., transition between sleep and wakeful activity). Therefore, understanding the effects of synaptic plasticity on neurons dynamics is a crucial challenge.

Our group is developing mathematical and numerical methods to analyse this mutual interaction. On the one hand, we have shown that plasticity mechanisms [8], [13], Hebbian-like or STDP, have strong effects on neuron dynamics complexity, such as synaptic and propagation delays [16], dynamics complexity reduction, and spike statistics.

<p align="center">**MORPHEME Project-Team**</p>

3. Research Program

3.1. Research program

The recent advent of an increasing number of new microscopy techniques giving access to high throughput screenings and micro or nano-metric resolutions provides a means for quantitative imaging of biological structures and phenomena. To conduct quantitative biological studies based on these new data, it is necessary to develop non-standard specific tools. This requires using a multi-disciplinary approach. We need biologists to define experiment protocols and interpret the results, but also physicists to model the sensors, computer scientists to develop algorithms and mathematicians to model the resulting information. These different expertises are combined within the Morpheme team. This generates a fecund frame for exchanging expertise, knowledge, leading to an optimal framework for the different tasks (imaging, image analysis, classification, modeling). We thus aim at providing adapted and robust tools required to describe, explain and model fundamental phenomena underlying the morphogenesis of cellular and supra-cellular biological structures. Combining experimental manipulations, in vivo imaging, image processing and computational modeling, we plan to provide methods for the quantitative analysis of the morphological changes that occur during development. This is of key importance as the morphology and topology of mesoscopic structures govern organ and cell function. Alterations in the genetic programs underlying cellular morphogenesis have been linked to a range of pathologies.

Biological questions we will focus on include:

1. what are the parameters and the factors controlling the establishment of ramified structures? (Are they really organize to ensure maximal coverage? How are genetic and physical constraints limiting their morphology?),

2. how are newly generated cells incorporated into reorganizing tissues during development? (is the relative position of cells governed by the lineage they belong to?)

Our goal is to characterize different populations or development conditions based on the shape of cellular and supra-cellular structures, e.g. micro-vascular networks, dendrite/axon networks, tissues from 2D, 2D+t, 3D or 3D+t images (obtained with confocal microscopy, video-microscopy, photon-microscopy or micro-tomography). We plan to extract shapes or quantitative parameters to characterize the morphometric properties of different samples. On the one hand, we will propose numerical and biological models explaining the temporal evolution of the sample, and on the other hand, we will statistically analyze shapes and complex structures to identify relevant markers for classification purposes. This should contribute to a better understanding of the development of normal tissues but also to a characterization at the supra-cellular scale of different pathologies such as Alzheimer, cancer, diabetes, or the Fragile X Syndrome. In this multidisciplinary context, several challenges have to be faced. The expertise of biologists concerning sample generation, as well as optimization of experimental protocols and imaging conditions, is of course crucial. However, the imaging protocols optimized for a qualitative analysis may be sub-optimal for quantitative biology. Second, sample imaging is only a first step, as we need to extract quantitative information. Achieving quantitative imaging remains an open issue in biology, and requires close interactions between biologists, computer scientists and applied mathematicians. On the one hand, experimental and imaging protocols should integrate constraints from the downstream computer-assisted analysis, yielding to a trade-off between qualitative optimized and quantitative optimized protocols. On the other hand, computer analysis should integrate constraints specific to the biological problem, from acquisition to quantitative information extraction. There is therefore a need of specificity for embedding precise biological information for a given task. Besides, a level of generality is also desirable for addressing data from different teams acquired with different protocols and/or sensors. The mathematical modeling of the physics of the acquisition system will yield higher performance reconstruction/restoration algorithms in terms of accuracy. Therefore, physicists and computer scientists have to work together. Quantitative information extraction also has to deal with both the complexity of the structures of interest (e.g., very

dense network, small structure detection in a volume, multiscale behavior, ...) and the unavoidable defects of in vivo imaging (artifacts, missing data, ...). Incorporating biological expertise in model-based segmentation methods provides the required specificity while robustness gained from a methodological analysis increases the generality. Finally, beyond image processing, we aim at quantifying and then statistically analyzing shapes and complex structures (e.g., neuronal or vascular networks), static or in evolution, taking into account variability. In this context, learning methods will be developed for determining (dis)similarity measures between two samples or for determining directly a classification rule using discriminative models, generative models, or hybrid models. Besides, some metrics for comparing, classifying and characterizing objects under study are necessary. We will construct such metrics for biological structures such as neuronal or vascular networks. Attention will be paid to computational cost and scalability of the developed algorithms: biological experimentations generally yield huge data sets resulting from high throughput screenings. The research of Morpheme will be developed along the following axes:

- **Imaging:** this includes i) definition of the studied populations (experimental conditions) and preparation of samples, ii) definition of relevant quantitative characteristics and optimized acquisition protocol (staining, imaging, ...) for the specific biological question, and iii) reconstruction/restoration of native data to improve the image readability and interpretation.

- **Feature extraction:** this consists in detecting and delineating the biological structures of interest from images. Embedding biological properties in the algorithms and models is a key issue. Two main challenges are the variability, both in shape and scale, of biological structures and the huge size of data sets. Following features along time will allow to address morphogenesis and structure development.

- **Classification/Interpretation:** considering a database of images containing different populations, we can infer the parameters associated with a given model on each dataset from which the biological structure under study has been extracted. We plan to define classification schemes for characterizing the different populations based either on the model parameters, or on some specific metric between the extracted structures.

- **Modeling:** two aspects will be considered. This first one consists in modeling biological phenomena such as axon growing or network topology in different contexts. One main advantage of our team is the possibility to use the image information for calibrating and/or validating the biological models. Calibration induces parameter inference as a main challenge. The second aspect consists in using a prior based on biological properties for extracting relevant information from images. Here again, combining biology and computer science expertise is a key point.

<p align="center" style="color:red">COATI Project-Team</p>

3. Research Program

3.1. Research Program

Members of COATI have a strong expertise in the design and management of wired and wireless backbone, backhaul, broadband, software defined and complex networks. On the one hand, we cope with specific problems such as energy efficiency in backhaul and backbone networks, routing reconfiguration in connection oriented networks (MPLS, WDM), traffic aggregation in SONET networks, compact routing in large-scale networks, survivability to single and multiple failures, etc. These specific problems often come from questions of our industrial partners. On the other hand, we study fundamental problems mainly related to routing and reliability that appear in many networks (not restricted to our main fields of applications) and that have been widely studied in the past. However, previous solutions do not take into account the constraints of current networks/traffic such as their huge size and their dynamics. COATI thus puts a significant research effort in the following directions:

- **Service Function Chains (SFC):** we study the placement of Service Function Chains within the network considering the ordering constraints. Then, we focus firstly on energy efficiency and secondly on reliability and protection mechanisms. In a last step, we study reconfiguration of the SFCs in case of dynamic traffic with a make-before-break approach.

- **Larger networks:** Another challenge one has to face is the increase in size of practical instances. It is already difficult, if not impossible, to solve practical instances optimally using existing tools. Therefore, we have to find new ways to solve problems using reduction and decomposition methods, characterization of polynomial instances (which are surprisingly often the practical ones), or algorithms with acceptable practical performances.

- **Stochastic behaviors:** Larger topologies mean frequent changes due to traffic and radio fluctuations, failures, maintenance operations, growth, routing policy changes, etc. We aim at including these stochastic behaviors in our combinatorial optimization process to handle the dynamics of the system and to obtain robust designs of networks.

The methods and tools used in our studies come from discrete mathematics and combinatorial optimization, and COATI contributes to their improvements. Also, COATI works on graph-decomposition methods and various games on graphs which are essential for a better understanding of the structural and combinatorial properties of the problems, but also for the design of efficient exact or approximate algorithms. We contribute to the modelling of optimization problems in terms of graphs, study the complexity of the problems, and then we investigate the structural or metric properties of graphs that make these problems hard or easy. We exploit these properties in the design of algorithms in order to find the most efficient ways for solving the problems.

COATI also focuses on the theory of *directed graphs*. Indeed, graph theory can be roughly partitioned into two branches: the areas of undirected graphs and directed graphs. Even though both areas have numerous important applications, for various reasons, undirected graphs have been studied much more extensively than directed graphs. It is worth noticing that many telecommunication problems are modelled with directed graphs. Therefore, a deeper understanding of the theory of directed graphs will benefit to the resolution of telecommunication networks problems. For instance, the problem of finding disjoint paths becomes much more difficult in directed graphs and understanding the underlying structures of actual directed networks would help us to propose solutions.

Last, we have recently started investigating how tools from multi-agents based systems and machine learning theory could help solving some optimization problems in networks. The arrival of Emanuele Natale as a Junior Researcher (CNRS) in the team and of two new PhD students (Franceso D'Amore and Hicham Lesfari) will foster these investigations.

<p style="text-align:center; color:red">DIANA Project-Team</p>

3. Research Program

3.1. Service Transparency

Transparency is to provide network users and application developers with reliable information about the current or predicted quality of their communication services, and about potential leakages of personal information, or of other information related to societal interests of the user as a "connected citizen" (e.g. possible violation of network neutrality, opinion manipulation). Service transparency therefore means to provide information meaningful to users and application developers, such as quality of experience, privacy leakages, or opinion manipulation, etc. rather than network-level metrics such as available bandwidth, loss rate, delay or jitter.

The Internet is built around a best effort routing service that does not provide any guarantee to end users in terms of quality of service (QoS). The simplicity of the Internet routing service is at the root of its huge success. Unfortunately, a simple service means unpredicted quality at the access. Even though a considerable effort is done by operators and content providers to optimise the Internet content delivery chain, mainly by over-provisioning and sophisticated engineering techniques, service degradation is still part of the Internet. The proliferation of wireless and mobile access technologies, and the versatile nature of Internet traffic, make end users quality of experience (QoE) forecast even harder. As a matter of fact, the Internet is missing a dedicated measurement plane that informs the end users on the quality they obtain and in case of substantial service degradation, on the origin of this degradation. Current state of the art activities are devoted to building a distributed measurement infrastructure to perform active, passive and hybrid measurements in the wired Internet. However, the problem is exacerbated with modern terminals such as smartphones or tablets that do not facilitate the task for end users (they even make it harder) as they focus on simplifying the interface and limiting the control on the network, whereas the Internet behind is still the same in terms of the quality it provides. Interestingly, this same observation explains the existing difficulty to detect and prevent privacy leaks. We argue that the lack of transparency for diagnosing QoE and for detecting privacy leaks have the same root causes and can be solved using common primitives. For instance, in both cases, it is important to be able to link data packets to an application. Indeed, as the network can only access data packets, there must be a way to bind these packets to an application (to understand users QoE for this application or to associate a privacy leak to an application). This is however a complex task as the traffic might be obfuscated or encrypted. Our objectives in the research direction are the following:

- Design and develop measurement tools providing transparency, in spite of current complexity
- Deploy those measurement tools at the Internet's edge and make them useful for end users
- Propose measurements plane as an overlay or by exploiting in-network functionalities
- Adapt measurements techniques to network architectural change
- Provide measurements as native functionality in future network architecture

3.2. Open network architecture

We are surrounded by personal content of all types: photos, videos, documents, etc. The volume of such content is increasing at a fast rate, and at the same time, the spread of such content among all our connected devices (mobiles, storage devices, set-top boxes, etc) is also increasing. All this complicates the control of personal content by the user both in terms of access and sharing with other users. The access of the personal content in a seamless way independently of its location is a key challenge for the future of networks. Proprietary solutions exist, but apart from fully depending on one of them, there is no standard plane in the Internet for a seamless access to personal content. Therefore, providing network architectural support to design and develop content access and sharing mechanisms is crucial to allow users control their own data over heterogeneous underlying network or cloud services.

On the other hand, privacy is a growing concern for states, administrations, and companies. Indeed, for instance the French CNIL (entity in charge of citizens privacy in computer systems) puts privacy at the core of its activities by defining rules on any stored and collected private data. Also, companies start to use privacy preserving solutions as a competitive advantage. Therefore, understanding privacy leaks and preventing them is a problem that can already find support. However, all end-users do not *currently* put privacy as their first concern. Indeed, in face of two services with one of higher quality, they usually prefer the highest quality one whatever the privacy implication. This was, for instance, the case concerning the Web search service of Google that is more accurate but less privacy preserving than Bing or Qwant. This is also the case for cloud services such as iCloud or Dropbox that are much more convenient than open source solutions, but very bad in terms of privacy. Therefore, to reach end-users, any privacy preserving solutions must offer a service equivalent to the best existing services.

We consider that it will be highly desirable for Internet users to be able to *easily* move their content from a provider to another and therefore not to depend on a content provider or a social network monopoly. This requires that the network provides built-in architectural support for content networking.

In this research direction, we will define a new *service abstraction layer* (SAL) that could become the new waist of the network architecture with network functionalities below (IP, SDN, cloud) and applications on top. SAL will define different services that are of use to all Internet users for accessing and sharing data (seamless content localisation and retrieval, privacy leakage protection, transparent vertical and horizontal handover, etc.). The biggest challenge here is to cope in the same time with large number of content applications requirements and high underlying networks heterogeneity while still providing efficient applications performance. This requires careful definition of the services primitives and the parameters to be exchanged through the service abstraction layer.

Two concurring factors make the concept behind SAL feasible and relevant today. First, the notion of scalable network virtualization that is a required feature to deploy SAL in real networks today has been discussed recently only. Second, the need for new services abstraction is recent. Indeed, more than fifteen years ago the Internet for the end-users was mostly the Web. Only ten years ago smartphones came into the picture of the Internet boosting the number of applications with new functionalities and risks. Since a few years, many discussions in the network communities took place around the actual complexity of the Internet and the difficulty to develop applications. Many different approaches have been discussed (such as CCN, SDN) that intend to solve only part of the complexity. SAL takes a broader architectural look at the problem and considers solutions such as CCN as mere use cases. Our objectives in this research direction include the following:

- Identify common key networking services required for content access and sharing
- Detect and prevent privacy leaks for content communication
- Enhance software defined networks for large scale heterogeneous environments
- Design and develop open Content Networking architecture
- Define a service abstraction layer as the thin waist for the future content network architecture
- Test and deploy different applications using SAL primitives on heterogeneous network technologies

3.3. Methodology

We follow an experimental approach that can be described in the following techniques:

- Measurements: the aim is to get a better view of a problem in quantifiable terms. Depending on the field of interest, this may involve large scale distributed systems crawling tools; active probing techniques to infer the status and properties of a complex and non controllable system as the Internet; or even crowdsourcing-based deployments for gathering data on real-users environments or behaviours.
- Experimental evaluation: once a new idea has been designed and implemented, it is of course very desirable to assess and quantify how effective it can be, before being able to deploy it on any realistic scale. This is why a wide range of techniques can be considered for getting early, yet as significant as possible, feedback on a given paradigm or implementation. The spectrum for such techniques span from simulations to real deployments in protected and/or controlled environments.

FOCUS Project-Team

3. Research Program

3.1. Foundations 1: Models

The objective of Focus is to develop concepts, techniques, and possibly also tools, that may contribute to the analysis and synthesis of CBUS. Fundamental to these activities is *modeling*. Therefore designing, developing and studying computational models appropriate for CBUS is a central activity of the project. The models are used to formalise and verify important computational properties of the systems, as well as to propose new linguistic constructs.

The models we study are in the process calculi (e.g., the π-calculus) and λ-calculus tradition. Such models, with their emphasis on algebra, well address compositionality—a central property in our approach to problems. Accordingly, the techniques we employ are mainly operational techniques based on notions of behavioural equivalence, and techniques based on algebra, mathematical logics, and type theory.

3.2. Foundations 2: Foundational calculi and interaction

Modern distributed systems have witnessed a clear shift towards interaction and conversations as basic building blocks for software architects and programmers. The systems are made by components, that are supposed to interact and carry out dialogues in order to achieve some predefined goal; Web services are a good example of this. Process calculi are models that have been designed precisely with the goal of understanding interaction and composition. The theory and tools that have been developed on top of process calculi can set a basis with which CBUS challenges can be tackled. Indeed industrial proposals of languages for Web services such as BPEL are strongly inspired by process calculi, notably the π-calculus.

3.3. Foundations 3: Type systems and logics

Type systems and logics for reasoning on computations are among the most successful outcomes in the history of the research in λ-calculus and (more recently) in process calculi. Type systems can also represent a powerful means of specifying dialogues among components of CBUS. For instance—again referring to Web services—current languages for specifying interactions only express basic connectivity, ignoring causality and timing aspects (e.g., an intended order on the messages), and the alternative is to use Turing Complete languages that are however undecidable. Types can come at hand here: they can express causality and order information on messages [53], [49], [54], while remaining decidable systems.

3.4. Foundations 4: Implicit computational complexity

A number of elegant and powerful results have been recently obtained in implicit computational complexity in the λ-calculus in which ideas from Linear Logics enable a fine-grained control over computations. This experience can be profitable when tackling issues of CBUS related to resource consumption, such as resources allocation, access to resources, certification of bounds on resource consumption (e.g., ensuring that a service will answer to a request in time polynomial with respect to the size of the input data).

<h1 style="color:red; text-align:center">INDES Project-Team</h1>

3. Research Program

3.1. Parallelism, concurrency, and distribution

Concurrency management is at the heart of diffuse programming. Since the execution platforms are highly heterogeneous, many different concurrency principles and models may be involved. Asynchronous concurrency is the basis of shared-memory process handling within multiprocessor or multicore computers, of direct or fifo-based message passing in distributed networks, and of fifo- or interrupt-based event handling in web-based human-machine interaction or sensor handling. Synchronous or quasi-synchronous concurrency is the basis of signal processing, of real-time control, and of safety-critical information acquisition and display. Interfacing existing devices based on these different concurrency principles within *Hop* or other diffuse programming languages will require better understanding of the underlying concurrency models and of the way they can nicely cooperate, a currently ill-resolved problem.

3.2. Web, functional, and reactive programming

We are studying new paradigms for programming Web applications that rely on multi-tier functional programming. We have created a Web programming environment named *Hop*. It relies on a single formalism for programming the server-side and the client-side of the applications as well as for configuring the execution engine.

Hop is a functional language based on the SCHEME programming language. That is, it is a strict functional language, fully polymorphic, supporting side effects, and dynamically type-checked. *Hop* is implemented as an extension of the BIGLOO compiler that we develop. In the past, we have extensively studied static analyses (type systems and inference, abstract interpretations, as well as classical compiler optimizations) to improve the efficiency of compilation in both space and time.

As a *Hop* DSL, we have created *HipHop*, a synchronous orchestration language for web and IoT applications. *HipHop* facilitates the design and programming of complex web/IoT applications by smoothly integrating three computation models and programming styles that have been historically developed in different communities and for different purposes: *i) Transformational programs* that simply compute output values from input values, with comparatively simple interaction with their environment; *ii)* asynchronous concurrent programs that perform interactions between their components or with their environment with uncontrollable timing, using typically network-based communication; and *iii)* synchronous reactive programs that react to external events in a conceptually instantaneous and deterministic way.

3.3. Security of diffuse programs

The main goal of our security research is to provide scalable and rigorous language-based techniques that can be integrated into multi-tier compilers to enforce the security of diffuse programs. Research on language-based security has been carried on before in former Inria teams. In particular previous research has focused on controlling information flow to ensure confidentiality.

Typical language-based solutions to these problems are founded on static analysis, logics, provable cryptography, and compilers that generate correct code by construction. Relying on the multi-tier programming language *Hop* that tames the complexity of writing and analysing secure diffuse applications, we are studying language-based solutions to prominent web security problems such as code injection and cross-site scripting, to name a few.

<p style="text-align: center; color: red; font-weight: bold;">NEO Project-Team</p>

3. Research Program

3.1. Stochastic Operations Research

Stochastic Operations Research is a collection of modeling, optimization and numerical computation techniques, aimed at assessing the behavior of man-made systems driven by random phenomena, and at helping to make decisions in such a context.

The discipline is based on applied probability and focuses on effective computations and algorithms. Its core theory is that of Markov chains over discrete state spaces. This family of stochastic processes has, at the same time, a very large modeling capability and the potential of efficient solutions. By "solution" is meant the calculation of some *performance metric*, usually the distribution of some random variable of interest, or its average, variance, etc. This solution is obtained either through exact "analytic" formulas, or numerically through linear algebra methods. Even when not analytically or numerically tractable, Markovian models are always amenable to "Monte-Carlo" simulations with which the metrics can be statistically measured.

An example of this is the success of classical Queueing Theory, with its numerous analytical formulas. Another important derived theory is that of the Markov Decision Processes, which allows to formalize *optimal* decision problems in a random environment. This theory allows to characterize the optimal decisions, and provides algorithms for calculating them.

Strong trends of Operations Research are: a) an increasing importance of multi-criteria multi-agent optimization, and the correlated introduction of Game Theory in the standard methodology; b) an increasing concern of (deterministic) Operations Research with randomness and risk, and the consequent introduction of topics like Chance Constrained Programming and Stochastic Optimization. Data analysis is also more and more present in Operations Research: techniques from statistics, like filtering and estimation, or Artificial Intelligence like clustering, are coupled with modeling in Machine Learning techniques like Q-Learning.

<p style="color:red;text-align:center;">**CHORALE Team**</p>

3. Research Program

3.1. Task based world modeling and understanding

Executing a robotic task needs to specify a task space and a set of objective functions to be optimized. One research issue will be to define a framework allowing to represent the tasks in a generic canonic space in order to make their design and their analysis easier thanks to the control theory tools (observability, controllability, robustness...). All along the execution of the task, autonomous robotics systems have to acquire and maintain a model of the world and of the interactions between the different components involved in the task (heterogeneous robots, human beings, changes in the environment...). This model evolves in time and in space. In this research axis, we will investigate novel task-oriented world multi-layers representations (photometry, geometry, semantic) embedded in a short/long term memory framework able to handle static and dynamic events (long term mapping). A particular attention will be also paid to integrate human-robot interactions in shared environment (social skills). Another ambition of the project will be to build a bridge between model-based and machine learning methods. Understanding the world evolution is one of the key of autonomy. In this aim, we will focus on situation awareness.

3.2. Multi-sensory perception and control

Multi-sensory based perception and control is an area that starts from one single robot evolving in the environment with a set of sensors, up to a set of heterogeneous robots collaborating for the execution of a global shared task. We will address problems such as the active selection of the most suitable source of information (e.g. sensors and features) during the execution of the task and the active sensing control in order to maximize the collected information about the world modeling (including calibration and environment parameters, exogenous disturbances), allowing the task-driven sensor-based control framework to be more efficiently and robustly executed. Another issue will be the execution of a task defined by another robot or human, and to be replicated with a robot with different capabilities in perception, control and level of autonomy (i.e. heterogeneous robots). Last issues will come from the collaboration of different autonomous and heterogeneous robots in order to accomplish a shared task (mapping, robust localization, calibration, tracking, transporting, moving, ...)

<p align="center" style="color:red">GRAPHDECO Project-Team</p>

3. Research Program

3.1. Introduction

Our research program is oriented around two main axes: 1) Computer-Assisted Design with Heterogeneous Representations and 2) Graphics with Uncertainty and Heterogeneous Content. These two axes are governed by a set of common fundamental goals, share many common methodological tools and are deeply intertwined in the development of applications.

3.1.1. *Computer-Assisted Design with Heterogeneous Representations*

Designers use a variety of visual representations to explore and communicate about a concept. Fig. 2 illustrates some typical representations, including sketches, hand-made prototypes, 3D models, 3D printed prototypes or instructions.

(a) Ideation sketch	(b) Presentation sketch	(c) Coarse prototype	(d) 3D model
(f) Simulation	(e) 3D Printing	(g) Technical diagram	(h) Instructions

Figure 2. Various representations of a hair dryer at different stages of the design process. Image source, in order: c-maeng on deviantart.com, shauntur on deviantart.com, "Prototyping and Modelmaking for Product Design" Hallgrimsson, B., Laurence King Publishers, 2012, samsher511 on turbosquid.com, my.solidworks.com, weilung tseng on cargocollective.com, howstuffworks.com, u-manual.com

The early representations of a concept, such as rough sketches and hand-made prototypes, help designers formulate their ideas and test the form and function of multiple design alternatives. These low-fidelity representations are meant to be cheap and fast to produce, to allow quick exploration of the *design space* of the concept. These representations are also often approximate to leave room for subjective interpretation and to stimulate imagination; in this sense, these representations can be considered *uncertain*. As the concept gets more finalized, time and effort are invested in the production of more detailed and accurate representations, such as high-fidelity 3D models suitable for simulation and fabrication. These detailed models can also be used to create didactic instructions for assembly and usage.

Producing these different representations of a concept requires specific skills in sketching, modeling, manufacturing and visual communication. For these reasons, professional studios often employ different experts to produce the different representations of the same concept, at the cost of extensive discussions and numerous iterations between the actors of this process. The complexity of the multi-disciplinary skills involved in the design process also hinders their adoption by laymen.

Existing solutions to facilitate design have focused on a subset of the representations used by designers. However, no solution considers all representations at once, for instance to directly convert a series of sketches into a set of physical prototypes. In addition, all existing methods assume that the concept is unique rather than ambiguous. As a result, rich information about the variability of the concept is lost during each conversion step.

We plan to facilitate design for professionals and laymen by adressing the following objectives:

- We want to assist designers in the exploration of the *design space* that captures the possible variations of a concept. By considering a concept as a *distribution* of shapes and functionalities rather than a single object, our goal is to help designers consider multiple design alternatives more quickly and effectively. Such a representation should also allow designers to preserve multiple alternatives along all steps of the design process rather than committing to a single solution early on and pay the price of this decision for all subsequent steps. We expect that preserving alternatives will facilitate communication with engineers, managers and clients, accelerate design iterations and even allow mass personalization by the end consumers.

- We want to support the various representations used by designers during concept development. While drawings and 3D models have received significant attention in past Computer Graphics research, we will also account for the various forms of rough physical prototypes made to evaluate the shape and functionality of a concept. Depending on the task at hand, our algorithms will either analyse these prototypes to generate a virtual concept, or assist the creation of these prototypes from a virtual model. We also want to develop methods capable of adapting to the different drawing and manufacturing techniques used to create sketches and prototypes. We envision design tools that conform to the habits of users rather than impose specific techniques to them.

- We want to make professional design techniques available to novices. Affordable software, hardware and online instructions are democratizing technology and design, allowing small businesses and individuals to compete with large companies. New manufacturing processes and online interfaces also allow customers to participate in the design of an object via mass personalization. However, similarly to what happened for desktop publishing thirty years ago, desktop manufacturing tools need to be simplified to account for the needs and skills of novice designers. We hope to support this trend by adapting the techniques of professionals and by automating the tasks that require significant expertise.

3.1.2. Graphics with Uncertainty and Heterogeneous Content

Our research is motivated by the observation that traditional CG algorithms have not been designed to account for uncertain data. For example, global illumination rendering assumes accurate virtual models of geometry, light and materials to simulate light transport. While these algorithms produce images of high realism, capturing effects such as shadows, reflections and interreflections, they are not applicable to the growing mass of uncertain data available nowadays.

The need to handle uncertainty in CG is timely and pressing, given the large number of *heterogeneous sources of 3D content* that have become available in recent years. These include data from cheap depth+image sensors (e.g., Kinect or the Tango), 3D reconstructions from image/video data, but also data from large 3D geometry databases, or casual 3D models created using simplified sketch-based modeling tools. Such alternate content has varying levels of *uncertainty* about the scene or objects being modelled. This includes uncertainty in geometry, but also in materials and/or lights – which are often not even available with such content. Since CG algorithms cannot be applied directly, visual effects artists spend hundreds of hours correcting inaccuracies and completing the captured data to make them useable in film and advertising.

Figure 3. Image-Based Rendering (IBR) techniques use input photographs and approximate 3D to produce new synthetic views.

We identify a major scientific bottleneck which is the need to treat *heterogeneous* content, i.e., containing both (mostly captured) uncertain and perfect, traditional content. Our goal is to provide solutions to this bottleneck, by explicitly and formally modeling uncertainty in CG, and to develop new algorithms that are capable of mixed rendering for this content.

We strive to develop methods in which heterogeneous – and often uncertain – data can be handled automatically in CG with a principled methodology. Our main focus is on *rendering* in CG, including dynamic scenes (video/animations).

Given the above, we need to address the following challenges:

- Develop a theoretical model to handle uncertainty in computer graphics. We must define a new formalism that inherently incorporates uncertainty, and must be able to express traditional CG rendering, both physically accurate and approximate approaches. Most importantly, the new formulation must elegantly handle mixed rendering of perfect synthetic data and captured uncertain content. An important element of this goal is to incorporate *cost* in the choice of algorithm and the optimizations used to obtain results, e.g., preferring solutions which may be slightly less accurate, but cheaper in computation or memory.

- The development of rendering algorithms for heterogeneous content often requires preprocessing of image and video data, which sometimes also includes depth information. An example is the decomposition of images into intrinsic layers of reflectance and lighting, which is required to perform relighting. Such solutions are also useful as image-manipulation or computational photography techniques. The challenge will be to develop such "intermediate" algorithms for the uncertain and heterogeneous data we target.

- Develop efficient rendering algorithms for uncertain and heterogeneous content, reformulating rendering in a probabilistic setting where appropriate. Such methods should allow us to develop approximate rendering algorithms using our formulation in a well-grounded manner. The formalism should include probabilistic models of how the scene, the image and the data interact. These models should be data-driven, e.g., building on the abundance of online geometry and image databases, domain-driven, e.g., based on requirements of the rendering algorithms or perceptually guided, leading to plausible solutions based on limitations of perception.

<p style="text-align:center; color:red;">**GRAPHIK Project-Team**</p>

3. Research Program

3.1. Logic-based Knowledge Representation and Reasoning

We follow the mainstream *logic-based* approach to knowledge representation (KR). First-order logic (FOL) is the reference logic in KR and most formalisms in this area can be translated into fragments (i.e., particular subsets) of FOL. This is in particular the case for description logics and existential rules, two well-known KR formalisms studied in the team.

A large part of research in this domain can be seen as studying the *trade-off* between the expressivity of languages and the complexity of (sound and complete) reasoning in these languages. The fundamental problem in KR languages is entailment checking: is a given piece of knowledge entailed by other pieces of knowledge, for instance from a knowledge base (KB)? Another important problem is *consistency* checking: is a set of knowledge pieces (for instance the knowledge base itself) consistent, i.e., is it sure that nothing absurd can be entailed from it? The *ontology-mediated query answering* problem is a topical problem (see Section 3.3). It asks for the set of answers to a query in the KB. In the case of Boolean queries (i.e., queries with a yes/no answer), it can be recast as entailment checking.

3.2. Graph-based Knowledge Representation and Reasoning

Besides logical foundations, we are interested in KR formalisms that comply, or aim at complying with the following requirements: to have good *computational* properties and to allow users of knowledge-based systems to have a maximal *understanding and control* over each step of the knowledge base building process and use.

These two requirements are the core motivations for our graph-based approach to KR. We view labelled graphs as an *abstract representation* of knowledge that can be expressed in many KR languages (different kinds of conceptual graphs —historically our main focus— the Semantic Web language RDF (Resource Description Framework), its extension RDFS (RDF Schema), expressive rules equivalent to the so-called tuple-generating-dependencies in databases, some description logics dedicated to query answering, etc.). For these languages, reasoning can be based on the structure of objects, thus based on graph-theoretic notions, while staying logically founded.

More precisely, our basic objects are labelled graphs (or hypergraphs) representing entities and relationships between these entities. These graphs have a natural translation in first-order logic. Our basic reasoning tool is graph homomorphism. The fundamental property is that graph homomorphism is sound and complete with respect to logical entailment *i.e.*, given two (labelled) graphs G and H, there is a homomorphism from G to H *if and only if* the formula assigned to G is entailed by the formula assigned to H. In other words, logical reasoning on these graphs can be performed by graph mechanisms. These knowledge constructs and the associated reasoning mechanisms can be extended (to represent rules for instance) while keeping this fundamental correspondence between graphs and logics.

3.3. Ontology-Mediated Query Answering

Querying knowledge bases has become a central problem in knowledge representation and in databases. A knowledge base (KB) is classically composed of a terminological part (metadata, ontology) and an assertional part (facts, data). Queries are supposed to be at least as expressive as the basic queries in databases, i.e., conjunctive queries, which can be seen as existentially closed conjunctions of atoms or as labelled graphs. The challenge is to define good trade-offs between the expressivity of the ontological language and the complexity of querying data in presence of ontological knowledge. Description logics have been so far the prominent family of formalisms for representing and reasoning with ontological knowledge. However, classical description logics were not designed for efficient data querying. On the other hand, database languages are able to process complex queries on huge databases, but without taking the ontology into account. There is thus a need for new languages and mechanisms, able to cope with the ever growing size of knowledge bases in the Semantic Web or in scientific domains.

This problem is related to two other problems identified as fundamental in KR:

- *Query answering with incomplete information.* Incomplete information means that it might be unknown whether a given assertion is true or false. Databases classically make the so-called closed-world assumption: every fact that cannot be retrieved or inferred from the base is assumed to be false. Knowledge bases classically make the open-world assumption: if something cannot be inferred from the base, and neither can its negation, then its truth status is unknown. The need of coping with incomplete information is a distinctive feature of querying knowledge bases with respect to querying classical databases (however, as explained above, this distinction tends to disappear). The presence of incomplete information makes the query answering task much more difficult.

- *Reasoning with rules.* Researching types of rules and adequate manners to process them is a mainstream topic in the Semantic Web, and, more generally a crucial issue for knowledge-based systems. For several years, we have been studying rules, both in their logical and their graph form, which are syntactically very simple but also very expressive. These rules, known as existential rules or Datalog+, can be seen as an abstraction of ontological knowledge expressed in the main languages used in the context of KB querying.

3.4. Inconsistency and Decision Making

While classical FOL is the kernel of many KR languages, to solve real-world problems we often need to consider features that cannot be expressed purely (or not naturally) in classical logic. The logic and graph-based formalisms used for previous points have thus to be extended with such features. The following requirements have been identified from scenarios in decision making, privileging the agronomy domain:

- to cope with inconsistency;
- to cope with defeasible knowledge;
- to take into account different and potentially conflicting viewpoints;
- to integrate decision notions (priorities, gravity, risk, benefit).

Although the solutions we develop require to be validated on the applications that motivated them, we also want them to be sufficiently generic to be applied in other contexts. One angle of attack (but not the only possible one) consists in increasing the expressivity of our core languages, while trying to preserve their essential combinatorial properties, so that algorithmic optimizations can be transferred to these extensions.

HEPHAISTOS Project-Team

3. Research Program

3.1. Interval analysis

We are interested in real-valued system solving ($f(X) = 0$, $f(X) \leq 0$), in optimization problems, and in the proof of the existence of properties (for example, it exists X such that $f(X) = 0$ or it exist two values X_1, X_2 such that $f(X_1) > 0$ and $f(X_2) < 0$). There are few restrictions on the function f as we are able to manage explicit functions using classical mathematical operators (e.g. $\sin(x + y) + \log(\cos(e^x) + y^2)$ as well as implicit functions (e.g. determining if there are parameter values of a parametrized matrix such that the determinant of the matrix is negative, without calculating the analytical form of the determinant).

Solutions are searched within a finite domain (called a *box*) which may be either continuous or mixed (i.e. for which some variables must belong to a continuous range while other variables may only have values within a discrete set). An important point is that we aim at finding all the solutions within the domain whenever the computer arithmetic will allow it: in other words we are looking for *certified* solutions. For example, for 0-dimensional system solving, we will provide a box that contains one, and only one, solution together with a numerical approximation of this solution. This solution may further be refined at will using multi-precision.

The core of our methods is the use of *interval analysis* that allows one to manipulate mathematical expressions whose unknowns have interval values. A basic component of interval analysis is the *interval evaluation* of an expression. Given an analytical expression F in the unknowns $\{x_1, x_2, ..., x_n\}$ and ranges $\{X_1, X_2, ..., X_n\}$ for these unknowns we are able to compute a range $[A, B]$, called the interval evaluation, such that

$$\forall \{x_1, x_2, ..., x_n\} \in \{X_1, X_2, ..., X_n\}, A \leq F(x_1, x_2, ..., x_n) \leq B \qquad (6)$$

In other words the interval evaluation provides a lower bound of the minimum of F and an upper bound of its maximum over the box.

For example if $F = x \, sin(x + x^2)$ and $x \in [0.5, 1.6]$, then $F([0.5, 1.6]) = [-1.362037441, 1.6]$, meaning that for any x in [0.5,1.6] we guarantee that $-1.362037441 \leq f(x) \leq 1.6$.

The interval evaluation of an expression has interesting properties:

- it can be implemented in such a way that the results are guaranteed with respect to round-off errors i.e. property 1 is still valid in spite of numerical errors induced by the use of floating point numbers

- if $A > 0$ or $B < 0$, then no values of the unknowns in their respective ranges can cancel F

- if $A > 0$ ($B < 0$), then F is positive (negative) for any value of the unknowns in their respective ranges

A major drawback of the interval evaluation is that $A(B)$ may be overestimated i.e. values of $x_1, x_2, ..., x_n$ such that $F(x_1, x_2, ..., x_n) = A(B)$ may not exist. This overestimation occurs because in our calculation each occurrence of a variable is considered as an independent variable. Hence if a variable has multiple occurrences, then an overestimation may occur. Such phenomena can be observed in the previous example where $B = 1.6$ while the real maximum of F is approximately 0.9144. The value of B is obtained because we are using in our calculation the formula $F = x sin(y + z^2)$ with y, z having the same interval value as x.

Fortunately there are methods that allow one to reduce the overestimation and the overestimation amount decreases with the width of the ranges. The latter remark leads to the use of a branch-and-bound strategy in which for a given box a variable range will be bisected, thereby creating two new boxes that are stored in a list and processed later on. The algorithm is complete if all boxes in the list have been processed, or if during the process a box generates an answer to the problem at hand (e.g. if we want to prove that $F(X) < 0$, then the algorithm stops as soon as $F(\mathcal{B}) \geq 0$ for a certain box $\mathcal{B}$).

A generic interval analysis algorithm involves the following steps on the current box [8], [4]:

1. *exclusion operators*: these operators determine that there is no solution to the problem within a given box. An important issue here is the extensive and smart use of the monotonicity of the functions

2. *filters*: these operators may reduce the size of the box i.e. decrease the width of the allowed ranges for the variables

3. *existence operators*: they allow one to determine the existence of a unique solution within a given box and are usually associated with a numerical scheme that allows for the computation of this solution in a safe way

4. *bisection*: choose one of the variable and bisect its range for creating two new boxes

5. *storage*: store the new boxes in the list

The scope of the HEPHAISTOS project is to address all these steps in order to find the most efficient procedures. Our efforts focus on mathematical developments (adapting classical theorems to interval analysis, proving interval analysis theorems), the use of symbolic computation and formal proofs (a symbolic pre-processing allows one to automatically adapt the solver to the structure of the problem), software implementation and experimental tests (for validation purposes).

Important note: We have insisted on interval analysis because this is a **major component** or our robotics activity. Our theoretical work in robotics is an analysis of the robotic environment in order to exhibit proofs on the behavior of the system that may be qualitative (e.g. the proof that a cable-driven parallel robot with more than 6 non-deformable cables will have at most 6 cables under tension simultaneously) or quantitative. In the quantitative case as we are dealing with realistic and not toy examples (including our own prototypes that are developed whenever no equivalent hardware is available or to very our assumptions) we have to manage problems that are so complex that analytical solutions are probably out of reach (e.g. the direct kinematics of parallel robots) and we have to resort to algorithms and numerical analysis. We are aware of different approaches in numerical analysis (e.g. some team members were previously involved in teams devoted to computational geometry and algebraic geometry) but interval analysis provides us another approach with high flexibility, the possibility of managing non algebraic problems (e.g. the kinematics of cable-driven parallel robots with sagging cables, that involves inverse hyperbolic functions) and to address various types of issues (system solving, optimization, proof of existence ...). However whenever needed we will rely as well on continuation, algebraic geometry, geometry or learning.

3.2. Robotics

HEPHAISTOS, as a follow-up of COPRIN, has a long-standing tradition of robotics studies, especially for closed-loop robots [3], especially cable-driven parallel robots. We address theoretical issues with the purpose of obtaining analytical and theoretical solutions, but in many cases only numerical solutions can be obtained due to the complexity of the problem. This approach has motivated the use of interval analysis for two reasons:

1. the versatility of interval analysis allows us to address issues (e.g. singularity analysis) that cannot be tackled by any other method due to the size of the problem

2. uncertainties (which are inherent to a robotic device) have to be taken into account so that the *real* robot is guaranteed to have the same properties as the *theoretical* one, even in the worst case. This is a crucial issue for many applications in robotics (e.g. medical or assistance robot)

Our field of study in robotics focuses on *kinematic* issues such as workspace and singularity analysis, positioning accuracy, trajectory planning, reliability, calibration, modularity management and, prominently, *appropriate design*, i.e. determining the dimensioning of a robot mechanical architecture that guarantees that the real robot satisfies a given set of requirements. The methods that we develop can be used for other robotic problems, see for example the management of uncertainties in aircraft design [6].

Our theoretical work must be validated through experiments that are essential for the sake of credibility. A contrario, experiments will feed theoretical work. Hence HEPHAISTOS works with partners on the development of real robots but also develops its own prototypes. In the last years we have developed a large number of prototypes and we have extended our development to devices that are not strictly robots but are part of an overall environment for assistance. We benefit here from the development of new miniature, low energy computers with an interface for analog and logical sensors such as the Arduino or the Phidgets. The web pages http://www-sop.inria.fr/hephaistos/mediatheque/index.html presents all of our prototypes and experimental work.

<p style="text-align:center;">**Stars Project-Team**</p>

3. Research Program

3.1. Introduction

Stars follows three main research directions: perception for activity recognition, action recognition and semantic activity recognition. **These three research directions are organized following the workflow of activity recognition systems**: First, *the perception* and *the action recognition* directions provide new techniques to extract powerful features, whereas *the semantic activity recognition* research direction provides new paradigms to match these features with concrete video analytic and healthcare applications.

Transversely, we consider a *new research axis in machine learning*, combining a priori knowledge and learning techniques, to set up the various models of an activity recognition system. A major objective is to automate model building or model enrichment at the perception level and at the understanding level.

3.2. Perception for Activity Recognition

Participants: François Brémond, Antitza Dantcheva, Sabine Moisan, Monique Thonnat.

Activity Recognition, Scene Understanding, Machine Learning, Computer Vision, Cognitive Vision Systems, Software Engineering

3.2.1. *Introduction*

Our main goal in perception is to develop vision algorithms able to address the large variety of conditions characterizing real world scenes in terms of sensor conditions, hardware requirements, lighting conditions, physical objects, and application objectives. We have also several issues related to perception which combine machine learning and perception techniques: learning people appearance, parameters for system control and shape statistics.

3.2.2. *Appearance Models and People Tracking*

An important issue is to detect in real-time physical objects from perceptual features and predefined 3D models. It requires finding a good balance between efficient methods and precise spatio-temporal models. Many improvements and analysis need to be performed in order to tackle the large range of people detection scenarios.

Appearance models. In particular, we study the temporal variation of the features characterizing the appearance of a human. This task could be achieved by clustering potential candidates depending on their position and their reliability. This task can provide any people tracking algorithms with reliable features allowing for instance to (1) better track people or their body parts during occlusion, or to (2) model people appearance for re-identification purposes in mono and multi-camera networks, which is still an open issue. The underlying challenge of the person re-identification problem arises from significant differences in illumination, pose and camera parameters. The re-identification approaches have two aspects: (1) establishing correspondences between body parts and (2) generating signatures that are invariant to different color responses. As we have already several descriptors which are color invariant, we now focus more on aligning two people detection and on finding their corresponding body parts. Having detected body parts, the approach can handle pose variations. Further, different body parts might have different influence on finding the correct match among a whole gallery dataset. Thus, the re-identification approaches have to search for matching strategies. As the results of the re-identification are always given as the ranking list, re-identification focuses on learning to rank. "Learning to rank" is a type of machine learning problem, in which the goal is to automatically construct a ranking model from a training data.

Therefore, we work on information fusion to handle perceptual features coming from various sensors (several cameras covering a large scale area or heterogeneous sensors capturing more or less precise and rich information). New 3D RGB-D sensors are also investigated, to help in getting an accurate segmentation for specific scene conditions.

Long term tracking. For activity recognition we need robust and coherent object tracking over long periods of time (often several hours in video surveillance and several days in healthcare). To guarantee the long term coherence of tracked objects, spatio-temporal reasoning is required. Modeling and managing the uncertainty of these processes is also an open issue. In Stars we propose to add a reasoning layer to a classical Bayesian framework modeling the uncertainty of the tracked objects. This reasoning layer can take into account the a priori knowledge of the scene for outlier elimination and long-term coherency checking.

Controlling system parameters.   Another research direction is to manage a library of video processing programs. We are building a perception library by selecting robust algorithms for feature extraction, by insuring they work efficiently with real time constraints and by formalizing their conditions of use within a program supervision model. In the case of video cameras, at least two problems are still open: robust image segmentation and meaningful feature extraction. For these issues, we are developing new learning techniques.

3.3. Action Recognition

Participants: François Brémond, Antitza Dantcheva, Monique Thonnat.

Machine Learning, Computer Vision, Cognitive Vision Systems

3.3.1. Introduction

Due to the recent development of high processing units, such as GPU, this is now possible to extract meaningful features directly from videos (e.g. video volume) to recognize reliably short actions. Action Recognition benefits also greatly from the huge progress made recently in Machine Learning (e.g. Deep Learning), especially for the study of human behavior. For instance, Action Recognition enables to measure objectively the behavior of humans by extracting powerful features characterizing their everyday activities, their emotion, eating habits and lifestyle, by learning models from a large number of data from a variety of sensors, to improve and optimize for example, the quality of life of people suffering from behavior disorders. However, Smart Homes and Partner Robots have been well advertised but remain laboratory prototypes, due to the poor capability of automated systems to perceive and reason about their environment. A hard problem is for an automated system to cope 24/7 with the variety and complexity of the real world. Another challenge is to extract people fine gestures and subtle facial expressions to better analyze behavior disorders, such as anxiety or apathy. Taking advantage of what is currently studied for self-driving cars or smart retails, there is a large avenue to design ambitious approaches for the healthcare domain. In particular, the advance made with Deep Learning algorithms has already enabled to recognize complex activities, such as cooking interactions with instruments, and from this analysis to differentiate healthy people from the ones suffering from dementia.

To address these issues, we propose to tackle several challenges:

3.3.2. Action recognition in the wild

The current Deep Learning techniques are mostly developed to work on few clipped videos, which have been recorded with students performing a limited set of predefined actions in front of a camera with high resolution. However, real life scenarios include actions performed in a spontaneous manner by older people (including people interactions with their environment or with other people), from different viewpoints, with varying framerate, partially occluded by furniture at different locations within an apartment depicted through long untrimmed videos. Therefore, a new dedicated dataset should be collected in a real-world setting to become a public benchmark video dataset and to design novel algorithms for ADL activity recognition. A special attention should be taken to anonymize the videos.

3.3.3. Attention mechanisms for action recognition

Activities of Daily Living (ADL) and video-surveillance activities are different from internet activities (e.g. Sports, Movies, YouTube), as they may have very similar context (e.g. same background kitchen) with high intra-variation (different people performing the same action in different manners), but in the same time low inter-variation, similar ways to perform two different actions (e.g. eating and drinking a glass of water). Consequently, fine-grained actions are badly recognized. So, we will design novel attention mechanisms for action recognition, for the algorithm being able to focus on a discriminative part of the person conducting the action. For instance, we will study attention algorithms, which could focus on the most appropriate body parts (e.g. full body, right hand). In particular, we plan to design a soft mechanism, learning the attention weights directly on the feature map of a 3DconvNet, a powerful convolutional network, which takes as input a batch of videos.

3.3.4. Action detection for untrimmed videos

Many approaches have been proposed to solve the problem of action recognition in short clipped 2D videos, which achieved impressive results with hand-crafted and deep features. However, these approaches cannot address real life situations, where cameras provide online and continuous video streams in applications such as robotics, video surveillance, and smart-homes. Here comes the importance of action detection to help recognizing and localizing each action happening in long videos. Action detection can be defined as the ability to localize starting and ending of each human action happening in the video, in addition to recognizing each action label. There have been few action detection algorithms designed for untrimmed videos, which are based on either sliding window, temporal pooling or frame-based labeling. However, their performance is too low to address real-word datasets. A first task consists in benchmarking the already published approaches to study their limitations on novel untrimmed video datasets, recorded following real-world settings. A second task could be to propose a new mechanism to improve either 1) the temporal pooling directly from the 3DconvNet architecture using for instance Temporal Convolution Networks (TCNs) or 2) frame-based labeling with a clustering technique (e.g. using Fisher Vectors) to discover the sub-activities of interest.

3.3.5. View invariant action recognition

The performance of current approaches strongly relies on the used camera angle: enforcing that the camera angle used in testing is the same (or extremely close to) as the camera angle used in training, is necessary for the approach performs well. On the contrary, the performance drops when a different camera view-point is used. Therefore, we aim at improving the performance of action recognition algorithms by relying on 3D human pose information. For the extraction of the 3D pose information, several open-source algorithms can be used, such as openpose or videopose3D (from CMU or Facebook research, https://github.com/CMU-Perceptual-Computing-Lab/openpose. Also, other algorithms extracting 3d meshes can be used. To generate extra views, Generative Adversarial Network (GAN) can be used together with the 3D human pose information to complete the training dataset from the missing view.

3.3.6. Uncertainty and action recognition

Another challenge is to combine the short-term actions recognized by powerful Deep Learning techniques with long-term activities defined by constraint-based descriptions and linked to user interest. To realize this objective, we have to compute the uncertainty (i.e. likelihood or confidence), with which the short-term actions are inferred. This research direction is linked to the next one, to Semantic Activity Recognition.

3.4. Semantic Activity Recognition

Participants: François Brémond, Sabine Moisan, Monique Thonnat.

Activity Recognition, Scene Understanding, Computer Vision

3.4.1. Introduction

Semantic activity recognition is a complex process where information is abstracted through four levels: signal (e.g. pixel, sound), perceptual features, physical objects and activities. The signal and the feature levels are characterized by strong noise, ambiguous, corrupted and missing data. The whole process of scene understanding consists in analyzing this information to bring forth pertinent insight of the scene and its dynamics while handling the low level noise. Moreover, to obtain a semantic abstraction, building activity models is a crucial point. A still open issue consists in determining whether these models should be given a priori or learned. Another challenge consists in organizing this knowledge in order to capitalize experience, share it with others and update it along with experimentation. To face this challenge, tools in knowledge engineering such as machine learning or ontology are needed.

Thus we work along the following research axes: high level understanding (to recognize the activities of physical objects based on high level activity models), learning (how to learn the models needed for activity recognition) and activity recognition and discrete event systems.

3.4.2. High Level Understanding

A challenging research axis is to recognize subjective activities of physical objects (i.e. human beings, animals, vehicles) based on a priori models and objective perceptual measures (e.g. robust and coherent object tracks).

To reach this goal, we have defined original activity recognition algorithms and activity models. Activity recognition algorithms include the computation of spatio-temporal relationships between physical objects. All the possible relationships may correspond to activities of interest and all have to be explored in an efficient way. The variety of these activities, generally called video events, is huge and depends on their spatial and temporal granularity, on the number of physical objects involved in the events, and on the event complexity (number of components constituting the event).

Concerning the modeling of activities, we are working towards two directions: the uncertainty management for representing probability distributions and knowledge acquisition facilities based on ontological engineering techniques. For the first direction, we are investigating classical statistical techniques and logical approaches. For the second direction, we built a language for video event modeling and a visual concept ontology (including color, texture and spatial concepts) to be extended with temporal concepts (motion, trajectories, events ...) and other perceptual concepts (physiological sensor concepts ...).

3.4.3. Learning for Activity Recognition

Given the difficulty of building an activity recognition system with a priori knowledge for a new application, we study how machine learning techniques can automate building or completing models at the perception level and at the understanding level.

At the understanding level, we are learning primitive event detectors. This can be done for example by learning visual concept detectors using SVMs (Support Vector Machines) with perceptual feature samples. An open question is how far can we go in weakly supervised learning for each type of perceptual concept (i.e. leveraging the human annotation task). A second direction is to learn typical composite event models for frequent activities using trajectory clustering or data mining techniques. We name composite event a particular combination of several primitive events.

3.4.4. Activity Recognition and Discrete Event Systems

The previous research axes are unavoidable to cope with the semantic interpretations. However they tend to let aside the pure event driven aspects of scenario recognition. These aspects have been studied for a long time at a theoretical level and led to methods and tools that may bring extra value to activity recognition, the most important being the possibility of formal analysis, verification and validation.

We have thus started to specify a formal model to define, analyze, simulate, and prove scenarios. This model deals with both absolute time (to be realistic and efficient in the analysis phase) and logical time (to benefit from well-known mathematical models providing re-usability, easy extension, and verification). Our purpose is to offer a generic tool to express and recognize activities associated with a concrete language to specify activities in the form of a set of scenarios with temporal constraints. The theoretical foundations and the tools being shared with Software Engineering aspects.

The results of the research performed in perception and semantic activity recognition (first and second research directions) produce new techniques for scene understanding and contribute to specify the needs for new software architectures (third research direction).

<p style="text-align:center">**TITANE Project-Team**</p>

3. Research Program

3.1. Context

Geometric modeling and processing revolve around three main end goals: a computerized shape representation that can be visualized (creating a realistic or artistic depiction), simulated (anticipating the real) or realized (manufacturing a conceptual or engineering design). Aside from the mere editing of geometry, central research themes in geometric modeling involve conversions between physical (real), discrete (digital), and mathematical (abstract) representations. Going from physical to digital is referred to as shape acquisition and reconstruction; going from mathematical to discrete is referred to as shape approximation and mesh generation; going from discrete to physical is referred to as shape rationalization.

Geometric modeling has become an indispensable component for computational and reverse engineering. Simulations are now routinely performed on complex shapes issued not only from computer-aided design but also from an increasing amount of available measurements. The scale of acquired data is quickly growing: we no longer deal exclusively with individual shapes, but with entire *scenes*, possibly at the scale of entire cities, with many objects defined as structured shapes. We are witnessing a rapid evolution of the acquisition paradigms with an increasing variety of sensors and the development of community data, as well as disseminated data.

In recent years, the evolution of acquisition technologies and methods has translated in an increasing overlap of algorithms and data in the computer vision, image processing, and computer graphics communities. Beyond the rapid increase of resolution through technological advances of sensors and methods for mosaicing images, the line between laser scan data and photos is getting thinner. Combining, e.g., laser scanners with panoramic cameras leads to massive 3D point sets with color attributes. In addition, it is now possible to generate dense point sets not just from laser scanners but also from photogrammetry techniques when using a well-designed acquisition protocol. Depth cameras are getting increasingly common, and beyond retrieving depth information we can enrich the main acquisition systems with additional hardware to measure geometric information about the sensor and improve data registration: e.g., accelerometers or GPS for geographic location, and compasses or gyrometers for orientation. Finally, complex scenes can be observed at different scales ranging from satellite to pedestrian through aerial levels.

These evolutions allow practitioners to measure urban scenes at resolutions that were until now possible only at the scale of individual shapes. The related scientific challenge is however more than just dealing with massive data sets coming from increase of resolution, as complex scenes are composed of multiple objects with structural relationships. The latter relate i) to the way the individual shapes are grouped to form objects, object classes or hierarchies, ii) to geometry when dealing with similarity, regularity, parallelism or symmetry, and iii) to domain-specific semantic considerations. Beyond reconstruction and approximation, consolidation and synthesis of complex scenes require rich structural relationships.

The problems arising from these evolutions suggest that the strengths of geometry and images may be combined in the form of new methodological solutions such as photo-consistent reconstruction. In addition, the process of measuring the geometry of sensors (through gyrometers and accelerometers) often requires both geometry process and image analysis for improved accuracy and robustness. Modeling urban scenes from measurements illustrates this growing synergy, and it has become a central concern for a variety of applications ranging from urban planning to simulation through rendering and special effects.

3.2. Analysis

Complex scenes are usually composed of a large number of objects which may significantly differ in terms of complexity, diversity, and density. These objects must be identified and their structural relationships must be recovered in order to model the scenes with improved robustness, low complexity, variable levels of details and ultimately, semantization (automated process of increasing degree of semantic content).

Object classification is an ill-posed task in which the objects composing a scene are detected and recognized with respect to predefined classes, the objective going beyond scene segmentation. The high variability in each class may explain the success of the stochastic approach which is able to model widely variable classes. As it requires a priori knowledge this process is often domain-specific such as for urban scenes where we wish to distinguish between instances as ground, vegetation and buildings. Additional challenges arise when each class must be refined, such as roof super-structures for urban reconstruction.

Structure extraction consists in recovering structural relationships between objects or parts of object. The structure may be related to adjacencies between objects, hierarchical decomposition, singularities or canonical geometric relationships. It is crucial for effective geometric modeling through levels of details or hierarchical multiresolution modeling. Ideally we wish to learn the structural rules that govern the physical scene manufacturing. Understanding the main canonical geometric relationships between object parts involves detecting regular structures and equivalences under certain transformations such as parallelism, orthogonality and symmetry. Identifying structural and geometric repetitions or symmetries is relevant for dealing with missing data during data consolidation.

Data consolidation is a problem of growing interest for practitioners, with the increase of heterogeneous and defect-laden data. To be exploitable, such defect-laden data must be consolidated by improving the data sampling quality and by reinforcing the geometrical and structural relations sub-tending the observed scenes. Enforcing canonical geometric relationships such as local coplanarity or orthogonality is relevant for registration of heterogeneous or redundant data, as well as for improving the robustness of the reconstruction process.

3.3. Approximation

Our objective is to explore the approximation of complex shapes and scenes with surface and volume meshes, as well as on surface and domain tiling. A general way to state the shape approximation problem is to say that we search for the shape discretization (possibly with several levels of detail) that realizes the best complexity / distortion trade-off. Such a problem statement requires defining a discretization model, an error metric to measure distortion as well as a way to measure complexity. The latter is most commonly expressed in number of polygon primitives, but other measures closer to information theory lead to measurements such as number of bits or minimum description length.

For surface meshes we intend to conceive methods which provide control and guarantees both over the global approximation error and over the validity of the embedding. In addition, we seek for resilience to heterogeneous data, and robustness to noise and outliers. This would allow repairing and simplifying triangle soups with cracks, self-intersections and gaps. Another exploratory objective is to deal generically with different error metrics such as the symmetric Hausdorff distance, or a Sobolev norm which mixes errors in geometry and normals.

For surface and domain tiling the term meshing is substituted for tiling to stress the fact that tiles may be not just simple elements, but can model complex smooth shapes such as bilinear quadrangles. Quadrangle surface tiling is central for the so-called *resurfacing* problem in reverse engineering: the goal is to tile an input raw surface geometry such that the union of the tiles approximates the input well and such that each tile matches certain properties related to its shape or its size. In addition, we may require parameterization domains with a simple structure. Our goal is to devise surface tiling algorithms that are both reliable and resilient to defect-laden inputs, effective from the shape approximation point of view, and with flexible control upon the structure of the tiling.

3.4. Reconstruction

Assuming a geometric dataset made out of points or slices, the process of shape reconstruction amounts to recovering a surface or a solid that matches these samples. This problem is inherently ill-posed as infinitely-many shapes may fit the data. One must thus regularize the problem and add priors such as simplicity or smoothness of the inferred shape.

The concept of geometric simplicity has led to a number of interpolating techniques commonly based upon the Delaunay triangulation. The concept of smoothness has led to a number of approximating techniques that commonly compute an implicit function such that one of its isosurfaces approximates the inferred surface. Reconstruction algorithms can also use an explicit set of prior shapes for inference by assuming that the observed data can be described by these predefined prior shapes. One key lesson learned in the shape problem is that there is probably not a single solution which can solve all cases, each of them coming with its own distinctive features. In addition, some data sets such as point sets acquired on urban scenes are very domain-specific and require a dedicated line of research.

In recent years the *smooth, closed case* (i.e., shapes without sharp features nor boundaries) has received considerable attention. However, the state-of-the-art methods have several shortcomings: in addition to being in general not robust to outliers and not sufficiently robust to noise, they often require additional attributes as input, such as lines of sight or oriented normals. We wish to devise shape reconstruction methods which are both geometrically and topologically accurate without requiring additional attributes, while exhibiting resilience to defect-laden inputs. Resilience formally translates into stability with respect to noise and outliers. Correctness of the reconstruction translates into convergence in geometry and (stable parts of) topology of the reconstruction with respect to the inferred shape known through measurements.

Moving from the smooth, closed case to the *piecewise smooth case* (possibly with boundaries) is considerably harder as the ill-posedness of the problem applies to each sub-feature of the inferred shape. Further, very few approaches tackle the combined issue of robustness (to sampling defects, noise and outliers) and feature reconstruction.

WIMMICS Project-Team

3. Research Program

3.1. Users Modeling and Designing Interaction on the Web

Wimmics focuses on interactions of ordinary users with ontology-based knowledge systems, with a preference for semantic Web formalisms and Web 2.0 applications. We specialize interaction design and evaluation methods to Web application tasks such as searching, browsing, contributing or protecting data. The team is especially interested in using semantics in assisting the interactions. We propose knowledge graph representations and algorithms to support interaction adaptation, for instance for context-awareness or intelligent interactions with machine. We propose and evaluate Web-based visualization techniques for linked data, querying, reasoning, explaining and justifying. Wimmics also integrates natural language processing approaches to support natural language based interactions. We rely on cognitive studies to build models of the system, the user and the interactions between users through the system, in order to support and improve these interactions. We extend the user modeling technique known as *Personas* where user models are represented as specific, individual humans. *Personas* are derived from significant behavior patterns (i.e., sets of behavioral variables) elicited from interviews with and observations of users (and sometimes customers) of the future product. Our user models specialize *Personas* approaches to include aspects appropriate to Web applications. Wimmics also extends user models to capture very different aspects (e.g. emotional states).

3.2. Communities and Social Interactions Analysis

The domain of social network analysis is a whole research domain in itself and Wimmics targets what can be done with typed graphs, knowledge representations and social models. We also focus on the specificity of social Web and semantic Web applications and in bridging and combining the different social Web data structures and semantic Web formalisms. Beyond the individual user models, we rely on social studies to build models of the communities, their vocabularies, activities and protocols in order to identify where and when formal semantics is useful. We propose models of collectives of users and of their collaborative functioning extending the collaboration personas and methods to assess the quality of coordination interactions and the quality of coordination artifacts. We extend and compare community detection algorithms to identify and label communities of interest with the topics they share. We propose mixed representations containing social semantic representations (e.g. folksonomies) and formal semantic representations (e.g. ontologies) and propose operations that allow us to couple them and exchange knowledge between them. Moving to social interaction we develop models and algorithms to mine and integrate different yet linked aspects of social media contributions (opinions, arguments and emotions) relying in particular on natural language processing and argumentation theory. To complement the study of communities we rely on multi-agent systems to simulate and study social behaviors. Finally we also rely on Web 2.0 principles to provide and evaluate social Web applications.

3.3. Vocabularies, Semantic Web and Linked Data Based Knowledge Representation and Artificial Intelligence Formalisms on the Web

For all the models we identified in the previous sections, we rely on and evaluate knowledge representation methodologies and theories, in particular ontology-based modeling. We also propose models and formalisms to capture and merge representations of different levels of semantics (e.g. formal ontologies and social folksonomies). The important point is to allow us to capture those structures precisely and flexibly and yet create as many links as possible between these different objects. We propose vocabularies and semantic Web formalizations for all the aspects that we model and we consider and study extensions of these formalisms when needed. The results have all in common to pursue the representation and publication of our models

as linked data. We also contribute to the transformation and linking of existing resources (informal models, databases, texts, etc.) to be published on the Semantic Web and as Linked Data. Examples of aspects we formalize include: user profiles, social relations, linguistic knowledge, business processes, derivation rules, temporal descriptions, explanations, presentation conditions, access rights, uncertainty, emotional states, licenses, learning resources, etc. At a more conceptual level we also work on modeling the Web architecture with philosophical tools so as to give a realistic account of identity and reference and to better understand the whole context of our research and its conceptual cornerstones.

3.4. Artificial Intelligence Processing: Learning, Analyzing and Reasoning on Heterogeneous Semantic Graphs

One of the characteristics of Wimmics is to rely on graph formalisms unified in an abstract graph model and operators unified in an abstract graph machine to formalize and process semantic Web data, Web resources, services metadata and social Web data. In particular Corese, the core software of Wimmics, maintains and implements that abstraction. We propose algorithms to process the mixed representations of the previous section. In particular we are interested in allowing cross-enrichment between them and in exploiting the life cycle and specificity of each one to foster the life-cycles of the others. Our results all have in common to pursue analyzing and reasoning on heterogeneous semantic graphs issued from social and semantic Web applications. Many approaches emphasize the logical aspect of the problem especially because logics are close to computer languages. We defend that the graph nature of Linked Data on the Web and the large variety of types of links that compose them call for typed graphs models. We believe the relational dimension is of paramount importance in these representations and we propose to consider all these representations as fragments of a typed graph formalism directly built above the Semantic Web formalisms. Our choice of a graph based programming approach for the semantic and social Web and of a focus on one graph based formalism is also an efficient way to support interoperability, genericity, uniformity and reuse.

ZENITH Project-Team

3. Research Program

3.1. Distributed Data Management

Data management is concerned with the storage, organization, retrieval and manipulation of data of all kinds, from small and simple to very large and complex. It has become a major domain of computer science, with a large international research community and a strong industry. Continuous technology transfer from research to industry has led to the development of powerful DBMS, now at the heart of any information system, and of advanced data management capabilities in many kinds of software products (search engines, application servers, document systems, etc.).

To deal with the massive scale of scientific data, we exploit large-scale distributed systems, with the objective of making distribution transparent to the users and applications. Thus, we capitalize on the principles of large-scale distributed systems such as clusters, peer-to-peer (P2P) and cloud.

Data management in distributed systems has been traditionally achieved by distributed database systems which enable users to transparently access and update several databases in a network using a high-level query language (e.g. SQL). Transparency is achieved through a global schema which hides the local databases' heterogeneity. In its simplest form, a distributed database system supports a global schema and implements distributed database techniques (query processing, transaction management, consistency management, etc.). This approach has proved to be effective for applications that can benefit from centralized control and full-fledge database capabilities, e.g. information systems. However, it cannot scale up to more than tens of databases.

Parallel database systems extend the distributed database approach to improve performance (transaction throughput or query response time) by exploiting database partitioning using a multiprocessor or cluster system. Although data integration systems and parallel database systems can scale up to hundreds of data sources or database partitions, they still rely on a centralized global schema and strong assumptions about the network.

In contrast, peer-to-peer (P2P) systems adopt a completely decentralized approach to data sharing. By distributing data storage and processing across autonomous peers in the network, they can scale without the need for powerful servers. P2P systems typically have millions of users sharing petabytes of data over the Internet. Although very useful, these systems are quite simple (e.g. file sharing), support limited functions (e.g. keyword search) and use simple techniques (e.g. resource location by flooding) which have performance problems. A P2P solution is well-suited to support the collaborative nature of scientific applications as it provides scalability, dynamicity, autonomy and decentralized control. Peers can be the participants or organizations involved in collaboration and may share data and applications while keeping full control over their (local) data sources. But for very-large scale scientific data analysis, we believe cloud computing (see next section), is the right approach as it can provide virtually infinite computing, storage and networking resources. However, current cloud architectures are proprietary, ad-hoc, and may deprive users of the control of their own data. Thus, we postulate that a hybrid P2P/cloud architecture is more appropriate for scientific data management, by combining the best of both approaches. In particular, it will enable the clean integration of the users' own computational resources with different clouds.

3.2. Big Data

Big data (like its relative, data science) has become a buzz word, with different meanings depending on your perspective, e.g. 100 terabytes is big for a transaction processing system, but small for a web search engine. It is also a moving target, as shown by two landmarks in DBMS products: the Teradata database machine in the 1980's and the Oracle Exadata database machine in 2010.

Although big data has been around for a long time, it is now more important than ever. We can see overwhelming amounts of data generated by all kinds of devices, networks and programs, e.g. sensors, mobile devices, connected objects (IoT), social networks, computer simulations, satellites, radiotelescopes, etc. Storage capacity has doubled every 3 years since 1980 with prices steadily going down (e.g. 1 Gigabyte of Hard Disk Drive for: 1M$ in 1982, 1K$ in 1995, 0.02$ in 2015), making it affordable to keep more data around. And massive data can produce high-value information and knowledge, which is critical for data analysis, decision support, forecasting, business intelligence, research, (data-intensive) science, etc.

The problem of big data has three main dimensions, quoted as the three big V's:

- Volume: refers to massive amounts of data, making it hard to store, manage, and analyze (big analytics);
- Velocity: refers to continuous data streams being produced, making it hard to perform online processing and analysis;
- Variety: refers to different data formats, different semantics, uncertain data, multiscale data, etc., making it hard to integrate and analyze.

There are also other V's such as: validity (is the data correct and accurate?); veracity (are the results meaningful?); volatility (how long do you need to store this data?).

Many different big data management solutions have been designed, primarily for the cloud, as cloud and big data are synergistic. They typically trade consistency for scalability, simplicity and flexibility, hence the new term Data-Intensive Scalable Computing (DISC). Examples of DISC systems include data processing frameworks (e.g. Hadoop MapReduce, Apache Spark, Pregel), file systems (e.g. Google GFS, HDFS), NoSQL systems (Google BigTable, Hbase, MongoDB), NewSQL systems (Google F1, CockroachDB, LeanXcale). In Zenith, we exploit or extend DISC technologies to fit our needs for scientific workflow management and scalable data analysis.

3.3. Data Integration

Scientists can rely on web tools to quickly share their data and/or knowledge. Therefore, when performing a given study, a scientist would typically need to access and integrate data from many data sources (including public databases). Data integration can be either physical or logical. In the former, the source data are integrated and materialized in a data warehouse. In logical integration, the integrated data are not materialized, but accessed indirectly through a global (or mediated) schema using a data integration system. These two approaches have different trade-offs, e.g. efficient analytics but only on historical data for data warehousing versus real-time access to data sources for data integration systems (e.g. web price comparators).

In both cases, to understand a data source content, metadata (data that describe the data) is crucial. Metadata can be initially provided by the data publisher to describe the data structure (e.g. schema), data semantics based on ontologies (that provide a formal representation of the domain knowledge) and other useful information about data provenance (publisher, tools, methods, etc.). Scientific metadata is very heterogeneous, in particular because of the autonomy of the underlying data sources, which leads to a large variety of models and formats. Thus, it is necessary to identify semantic correspondences between the metadata of the related data sources. This requires the matching of the heterogeneous metadata, by discovering semantic correspondences between ontologies, and the annotation of data sources using ontologies. In Zenith, we rely on semantic web techniques (e.g. RDF and SparkQL) to perform these tasks and deal with high numbers of data sources.

Scientific workflow management systems (SWfMS) are also useful for data integration. They allow scientists to describe and execute complex scientific activities, by automating data derivation processes, and supporting various functions such as provenance management, queries, reuse, etc. Some workflow activities may access or produce huge amounts of distributed data. This requires using distributed and parallel execution environments. However, existing workflow management systems have limited support for data parallelism. In Zenith, we use an algebraic approach to describe data-intensive workflows and exploit parallelism.

3.4. Data Analytics

Data analytics refers to a set of techniques to draw conclusions through data examination. It involves data mining, statistics, and data management, and is applied to categorical and continuous data. In the Zenith team, we are interested in both of these data types. Categorical data designates a set of data that can be described as "check boxes". It can be names, products, items, towns, etc. A common illustration is the market basket data, where each item bought by a client is recorded and the set of items is the basket. The typical data mining problems with this kind of data are:

- **Frequent itemsets and association rules**. In this case, the data is usually a table with a high number of rows and the data mining algorithm extracts correlations between column values. A typical example of frequent itemset from a sensor network in a smart building would say that "in 20% rooms, the door is closed, the room is empty, and lights are on."

- **Frequent sequential pattern extraction.** This problem is very similar to frequent itemset discovery but considering the order between. In the smart building example, a frequent sequence could say that "in 40% of rooms, lights are on at time i, the room is empty at time $i + j$ and the door is closed at time $i + j + k$".

- **Clustering**. The goal of clustering is to group together similar data while ensuring that dissimilar data will not be in the same cluster. In our example of smart buildings, we could find clusters of rooms, where offices will be in one category and copy machine rooms in another because of their differences (hours of people presence, number of times lights are turned on/off, etc.).

Continuous data are numeric records that can have an infinite number of values between any two values. A temperature value or a timestamp are examples of such data. They are involved in a widely used type of data known as time series: a series of values, ordered by time, and giving a measure, e.g. coming from a sensor. There is a large number of problems that can apply to this kind of data, including:

- **Indexing and retrieval.** The goal, here, is usually to find, given a query q and a time series dataset D, the records of D that are most similar to q. This may involve any transformation of D by means of an index or an alternative representation for faster execution.

- **Pattern and outlier detection.** The discovery of recurrent patterns or atypical sub-windows in a time series has applications in finance, industrial manufacture or seismology, to name a few. It calls for techniques that avoid pairwise comparisons of all the sub-windows, which would lead to prohibitive response times.

- **Clustering**. The goal is the same as categorical data clustering: group similar time series and separate dissimilar ones.

One main problem in data analytics is to deal with data streams. Existing methods have been designed for very large data sets where complex algorithms from artificial intelligence were not efficient because of data size. However, we now must deal with data streams, sequences of data events arriving at high rate, where traditional data analytics techniques cannot complete in real-time, given the infinite data size. In order to extract knowledge from data streams, the data mining community has investigated approximation methods that could yield good result quality.

3.5. High dimensional data processing and search

High dimensionality is inherent in applications involving images, audio and text as well as in many scientific applications involving raster data or high-throughput data. Because of the *dimensionality curse*, technologies for processing and analyzing such data cannot rely on traditional relational DBMS or data mining methods. It rather requires to employ machine learning methods such as dimensionality reduction, representation learning or random projection. The activity of Zenith in this domain focuses on methods that permit data processing and search at scale, in particular in the presence of strong uncertainty and/or ambiguity. Actually, while small datasets are often characterized by a careful collection process, massive amounts of data often come with outliers and spurious items, because it appears impossible to guarantee faultless collection at massive

bandwidth. Another source of noise is often the sensor itself, that may be of low quality but of high sampling rate, or even the actual content, e.g. in cultural heritage applications when historical content appears seriously damaged by time. To attack these difficult problems, we focus on the following research topics:

- **Uncertainty estimation**. Items in massive datasets may either be uncertain, e.g. for automatically annotated data as in image analysis, or be more or less severely corrupted by noise, e.g. in noisy audio recordings or in the presence of faulty sensors. In both cases, the concept of *uncertainty* is central for the end-user to exploit the content. In this context, we use probability theory to quantify uncertainty and propose machine learning algorithms that may operate robustly, or at least assess the quality of their output. This vast topic of research is guided by large-scale applications (both data search and data denoising), and our research is oriented towards computationally effective methods.

- **Deep neural networks**. A major breakthrough in machine learning performance has been the advent of deep neural networks, which are characterized by high numbers (millions) of parameters and scalable learning procedures. We are striving towards original architectures and methods that are theoretically grounded and offer state-of-the-art performance for data search and processing. The specific challenges we investigate are: very high dimensionality for static data and very long-term dependency for temporal data, both in the case of possibly strong uncertainty or ambiguity (e.g. hundreds of thousands of classes).

- **Community service**. Research in machine learning is guided by applications. In Zenith, two main communities are targeted: botany, and digital humanities. In both cases, our observation is that significant breakthroughs may be achieved by connecting these communities to machine learning researchers. This may be achieved through wording application-specific problems in classical machine learning parlance. Thus, the team is actively involved in the organization of international evaluation campaigns that allow machine learning researchers to propose new methods while solving important application problems.