Activity Report 2019

Section Application Domains

ABS Project-Team (section vide)

<div align="center">

ACUMES Project-Team

</div>

4. Application Domains

4.1. Active flow control for vehicles

The reduction of CO_2 emissions represents a great challenge for the automotive and aeronautic industries, which committed respectively a decrease of 20% for 2020 and 75% for 2050. This goal will not be reachable, unless a significant improvement of the aerodynamic performance of cars and aircrafts is achieved (e.g. aerodynamic resistance represents 70% of energy losses for cars above 90 km/h). Since vehicle design cannot be significantly modified, due to marketing or structural reasons, active flow control technologies are one of the most promising approaches to improve aerodynamic performance. This consists in introducing micro-devices, like pulsating jets or vibrating membranes, that can modify vortices generated by vehicles. Thanks to flow non-linearities, a small energy expense for actuation can significantly reduce energy losses. The efficiency of this approach has been demonstrated, experimentally as well as numerically, for simple configurations [167].

However, the lack of efficient and flexible numerical tools, that allow to simulate and optimize a large number of such devices on realistic configurations, is still a bottleneck for the emergence of this technology in industry. The main issue is the necessity of using high-order schemes and complex models to simulate actuated flows, accounting for phenomena occurring at different scales. In this context, we intend to contribute to the following research axes:

- *Sensitivity analysis for actuated flows.* Adjoint-based (reverse) approaches, classically employed in design optimization procedure to compute functional gradients, are not well suited to this context. Therefore, we propose to explore the alternative (direct) formulation, which is not so much used, in the perspective of a better characterization of actuated flows and optimization of control devices.

- *Isogeometric simulation of control devices.* To simulate flows perturbed by small-scale actuators, we investigate the use of isogeometric analysis methods, which allow to account exactly for CAD-based geometries in a high-order hierarchical representation framework. In particular, we try to exploit the features of the method to simulate more accurately complex flows including moving devices and multiscale phenomena.

4.2. Vehicular and pedestrian traffic flows

Intelligent Transportation Systems (ITS) is nowadays a booming sector, where the contribution of mathematical modeling and optimization is widely recognized. In this perspective, traffic flow models are a commonly cited example of "complex systems", in which individual behavior and self-organization phenomena must be taken into account to obtain a realistic description of the observed macroscopic dynamics [125]. Further improvements require more advanced models, keeping into better account interactions at the microscopic scale, and adapted control techniques, see [71] and references therein. In particular, we will focus on the following aspects:

- *Junction models.* We are interested in designing a general junction model both satisfying basic analytical properties guaranteeing well-posedness and being realistic for traffic applications. In particular, the model should be able to overcome severe drawbacks of existing models, such as restrictions on the number of involved roads and prescribed split ratios [86], [114], which limit their applicability to real world situations. Hamilton-Jacobi equations could be also an interesting direction of research, following the recent results obtained in [130].

- *Data assimilation.* In traffic flow modeling, the capability of correctly estimating and predicting the state of the system depends on the availability of rich and accurate data on the network. Up to now, the most classical sensors are fixed ones. They are composed of inductive loops (electrical wires) that are installed at different spatial positions of the network and that can measure the traffic flow,

the occupancy rate (i.e. the proportion of time during which a vehicle is detected to be over the loop) and the speed (in case of a system of two distant loops). These data are useful / essential to calibrate the phenomenological relationship between flow and density which is known in the traffic literature as the Fundamental Diagram. Nowadays, thanks to the wide development of mobile internet and geolocalization techniques and its increasing adoption by the road users, smartphones have turned into perfect mobile sensors in many domains, including in traffic flow management. They can provide the research community with a large database of individual trajectory sets that are known as Floating Car Data (FCD), see [127] for a real field experiment. Classical macroscopic models, say (hyperbolic systems of) conservation laws, are not designed to take into account this new kind of microscopic data. Other formulations, like Hamilton-Jacobi partial differential equations, are most suited and have been intensively studied in the past five years (see [79], [80]), with a stress on the (fixed) Eulerian framework. Up to our knowledge, there exist a few studies in the time-Lagrangian as well as space-Lagrangian frameworks, where data coming from mobile sensors could be easily assimilated, due to the fact that the Lagrangian coordinate (say the label of a vehicle) is fixed.

- *Control of autonomous vehicles.* Traffic flow is usually controlled via traffic lights or variable speed limits, which have fixed space locations. The deployment of autonomous vehicles opens new perspectives in traffic management, as the use of a small fraction of cars to optimize the overall traffic. In this perspective, the possibility to track vehicles trajectories either by coupled micro-macro models [94], [115] or via the Hamilton-Jacobi approach [79], [80] could allow to optimize the flow by controlling some specific vehicles corresponding to internal conditions.

4.3. Virtual Fractional Flow Reserve in coronary stenting

Atherosclerosis is a chronic inflammatory disease that affects the entire arterial network and especially the coronary arteries. It is an accumulation of lipids over the arterial surface due to a dysfunction of this latter. The objective of clinical intervention, in this case, is to establish a revascularization using different angioplasty techniques, among which the implantation of stents is the most widespread. This intervention involves introducing a stent into the damaged portion in order to allow the blood to circulate in a normal way over all the vessels. Revascularization is based on the principle of remedying ischemia, which is a decrease or an interruption of the supply of oxygen to the various organs. This anomaly is attenuated by the presence of several lesions (multivessel disease patients), which can lead to several complications. The key of a good medical intervention is the fact of establishing a good diagnosis, in order to decide which lesion requires to be treated. In the diagnosis phase, the clinician uses several techniques, among which angiography is the most popular. Angiography is an X-ray technique to show the inside (the lumen) of blood vessels, in order to identify vessel narrowing: stenosis. Despite its widespread use, angiography is often imperfect in determining the physiological significance of coronary stenosis. If the problem remains simple for non significant lesions ($\leq 40\%$) or very severe ($\geq 70\%$), a very important category of intermediate lesions must benefit from a functional evaluation which will determine the strategy of treatment [90].

The technique of the Fractional Flow Reserve (FFR) has derived from the initial coronary physical approaches decades ago. Since then, many studies have demonstrated its effectiveness in improving the patients prognosis, by applying the appropriate approach. Its contribution in the reduction of mortality was statistically proved by the FAME (Fractional Flow Reserve Versus Angiography for Multivessel Evaluation) study [169]. It is established that the FFR can be easily measured during coronary angiography by calculating the ratio of distal coronary pressure P_d to aortic pressure P_a. These pressures are measured simultaneously with a special guide-wire. FFR in a normal coronary artery equals to 1.0. FFR value of 0.80 or less identifies ischemia-causing coronary lesions with an accuracy of more than 90% [169].

Obviously, from an interventional point of view, the FFR is binding since it is invasive. It should also be noted that this technique induces additional costs, which are not covered by insurances in several countries. For these reasons, it is used only in less than 10% of the cases.

In this perspective, a new virtual version of the FFR, entitled VFFR, has emerged as an attractive and non-invasive alternative to standard FFR, see [161], [144]. VFFR is based on computational modeling, mainly fluid and fluid-structural dynamics. However, there are key scientific, logistic and commercial challenges that need to be overcome before VFFR can be translated into routine clinical practice.

While most of the studies related to VFFR use Navier-Stokes models, we focus on the non-newtonian case, starting with a generalized fluid flow approach. These models are more relevant for the coronary arteries, and we expect that the computation of the FFR should then be more accurate. We are also leading numerical studies to assess the impact (on the FFR) of the interaction of the physical devices (catheter, optical captors, spheroids) with the blood flow.

4.4. Other application fields

Besides the above mentioned axes, which constitute the project's identity, the methodological tools described in Section have a wider range of application. We currently carry on also the following research actions, in collaboration with external partners.

- **Modeling cell dynamics.** Migration and proliferation of epithelial cell sheets are the two keystone aspects of the collective cell dynamics in most biological processes such as morphogenesis, embryogenesis, cancer and wound healing. It is then of utmost importance to understand their underlying mechanisms.

  Semilinear reaction-diffusion equations are widely used to give a phenomenological description of the temporal and spatial changes occurring within cell populations that undergo scattering (moving), spreading (expanding cell surface) and proliferation. We have followed the same methodology and contributed to assess the validity of such approaches in different settings (cell sheets [122], dorsal closure [58], actin organization [57]). However, epithelial cell-sheet movement is complex enough to undermine most of the mathematical approaches based on *locality*, that is mainly traveling wavefront-like partial differential equations. In [109] it is shown that Madin-Darby Canine Kidney (MDCK) cells extend cryptic lamellipodia to drive the migration, several rows behind the wound edge. In [149] MDCK monolayers are shown to exhibit similar non-local behavior (long range velocity fields, very active border-localized leader cells).

  Our aim is to start from a mesoscopic description of cell interaction: considering cells as independent anonymous agents, we plan to investigate the use of mathematical techniques adapted from the mean-field game theory. Otherwise, looking at them as interacting particles, we will use a multi-agent approach (at least for the actin dynamics). We intend also to consider approaches stemming from compartment-based simulation in the spirit of those developed in [106], [111], [113].

- **Game strategies for thermoelastography.** Thermoelastography is an innovative non-invasive control technology, which has numerous advantages over other techniques, notably in medical imaging [142]. Indeed, it is well known that most pathological changes are associated with changes in tissue stiffness, while remaining isoechoic, and hence difficult to detect by ultrasound techniques. Based on elastic waves and heat flux reconstruction, thermoelastography shows no destructive or aggressive medical sequel, unlike X-ray and comparables techniques, making it a potentially prominent choice for patients.

  Physical principles of thermoelastography originally rely on dynamical structural responses of tissues, but as a first approach, we only consider static responses of linear elastic structures.
  The mathematical formulation of the thermoelasticity reconstruction is based on data completion and material identification, making it a harsh ill posed inverse problem. In previous works [123], [132], we have demonstrated that Nash game approaches are efficient to tackle ill-posedness. We intend to extend the results obtained for Laplace equations in [123], and the algorithms developed in Section 3.1.2.4 to the following problems (of increasing difficulty):
  - Simultaneous data and parameter recovery in linear elasticity, using the so-called Kohn and Vogelius functional (ongoing work, some promising results obtained).

- Data recovery in coupled heat-thermoelasticity systems.
- Data recovery in linear thermoelasticity under stochastic heat flux, where the imposed flux is stochastic.
- Data recovery in coupled heat-thermoelasticity systems under stochastic heat flux, formulated as an incomplete information Nash game.
- Application to robust identification of cracks.

- **Constraint elimination in Quasi-Newton methods.** In single-objective differentiable optimization, Newton's method requires the specification of both gradient and Hessian. As a result, the convergence is quadratic, and Newton's method is often considered as the target reference. However, in applications to distributed systems, the functions to be minimized are usually "functionals", which depend on the optimization variables by the solution of an often complex set of PDE's, through a chain of computational procedures. Hence, the exact calculation of the full Hessian becomes a complex and costly computational endeavor.

  This has fostered the development of *quasi-Newton's methods* that mimic Newton's method but use only the gradient, the Hessian being iteratively constructed by successive approximations inside the algorithm itself. Among such methods, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm is well-known and commonly employed. In this method, the Hessian is corrected at each new iteration by rank-one matrices defined from several evaluations of the gradient only. The BFGS method has "super-linear convergence".

  For constrained problems, certain authors have developed so-called *Riemannian BFGS*, e.g. [152], that have the desirable convergence property in constrained problems. However, in this approach, the constraints are assumed to be known formally, by explicit expressions.

  In collaboration with ONERA-Meudon, we are exploring the possibility of representing constraints, in successive iterations, through local approximations of the constraint surfaces, splitting the design space locally into tangent and normal sub-spaces, and eliminating the normal coordinates through a linearization, or more generally a finite expansion, and applying the BFGS method through dependencies on the coordinates in the tangent subspace only. Preliminary experiments on the difficult Rosenbrock test-case, although in low dimensions, demonstrate the feasibility of this approach. On-going research is on theorizing this method, and testing cases of higher dimensions.

- **Multi-objective optimization for nanotechnologies.** Our team takes part in a larger collaboration with CEA/LETI (Grenoble), initiated by the Inria Project-Team Nachos, and related to the Maxwell equations. Our component in this activity relates to the optimization of nanophotonic devices, in particular with respect to the control of thermal loads. We have first identified a gradation of representative test-cases of increasing complexity:
- infrared micro-source;
- micro-photoacoustic cell;
- nanophotonic device.

These cases involve from a few geometric parameters to be optimized to a functional minimization subject to a finite-element solution involving a large number of dof's. CEA disposes of such codes, but considering the computational cost of the objective functions in the complex cases, the first part of our study is focused on the construction and validation of meta-models, typically of RBF-type. Multi-objective optimization will be carried out subsequently by MGDA, and possibly Nash games.

<div align="center">

AROMATH Project-Team

</div>

4. Application Domains

4.1. Geometric modeling for Design and Manufacturing.

The main domain of applications that we consider for the methods we develop is Computer Aided Design and Manufacturing.

Computer-Aided Design (CAD) involves creating digital models defined by mathematical constructions, from geometric, functional or aesthetic considerations. Computer-aided manufacturing (CAM) uses the geometrical design data to control the tools and processes, which lead to the production of real objects from their numerical descriptions.

CAD-CAM systems provide tools for visualizing, understanding, manipulating, and editing virtual shapes. They are extensively used in many applications, including automotive, shipbuilding, aerospace industries, industrial and architectural design, prosthetics, and many more. They are also widely used to produce computer animation for special effects in movies, advertising and technical manuals, or for digital content creation. Their economic importance is enormous. Their importance in education is also growing, as they are more and more used in schools and educational purposes.

CAD-CAM has been a major driving force for research developments in geometric modeling, which leads to very large software, produced and sold by big companies, capable of assisting engineers in all the steps from design to manufacturing.

Nevertheless, many challenges still need to be addressed. Many problems remain open, related to the use of efficient shape representations, of geometric models specific to some application domains, such as in architecture, naval engineering, mechanical constructions, manufacturing ...Important questions on the robustness and the certification of geometric computation are not yet answered. The complexity of the models which are used nowadays also appeals for the development of new approaches. The manufacturing environment is also increasingly complex, with new type of machine tools including: turning, 5-axes machining and wire EDM (Electrical Discharge Machining), 3D printer. It cannot be properly used without computer assistance, which raises methodological and algorithmic questions. There is an increasing need to combine design and simulation, for analyzing the physical behavior of a model and for optimal design.

The field has deeply changed over the last decades, with the emergence of new geometric modeling tools built on dedicated packages, which are mixing different scientific areas to address specific applications. It is providing new opportunities to apply new geometric modeling methods, output from research activities.

4.2. Geometric modeling for Numerical Simulation and Optimization

A major bottleneck in the CAD-CAM developments is the lack of interoperability of modeling systems and simulation systems. This is strongly influenced by their development history, as they have been following different paths.

The geometric tools have evolved from supporting a limited number of tasks at separate stages in product development and manufacturing, to being essential in all phases from initial design through manufacturing.

Current Finite Element Analysis (FEA) technology was already well established 40 years ago, when CAD-systems just started to appear, and its success stems from using approximations of both the geometry and the analysis model with low order finite elements (most often of degree ≤ 2).

There has been no requirement between CAD and numerical simulation, based on Finite Element Analysis, leading to incompatible mathematical representations in CAD and FEA. This incompatibility makes interoperability of CAD/CAM and FEA very challenging. In the general case today this challenge is addressed by expensive and time-consuming human intervention and software developments.

Improving this interaction by using adequate geometric and functional descriptions should boost the interaction between numerical analysis and geometric modeling, with important implications in shape optimization. In particular, it could provide a better feedback of numerical simulations on the geometric model in a design optimization loop, which incorporates iterative analysis steps.

The situation is evolving. In the past decade, a new paradigm has emerged to replace the traditional Finite Elements by B-Spline basis element of any polynomial degree, thus in principle enabling exact representation of all shapes that can be modeled in CAD. It has been demonstrated that the so-called isogeometric analysis approach can be far more accurate than traditional FEA.

It opens new perspectives for the interoperability between geometric modeling and numerical simulation. The development of numerical methods of high order using a precise description of the shapes raises questions on piecewise polynomial elements, on the description of computational domains and of their interfaces, on the construction of good function spaces to approximate physical solutions. All these problems involve geometric considerations and are closely related to the theory of splines and to the geometric methods we are investigating. We plan to apply our work to the development of new interactions between geometric modeling and numerical solvers.

ATHENA Project-Team

4. Application Domains

4.1. Applications of diffusion MRI

Clinical domain: Diagnosis of neurological disorder
Various examples of CNS diseases as Alzheimer's and Parkinson's diseases and others like multiple sclerosis, traumatic brain injury and schizophrenia have characteristic abnormalities in the micro-structure of brain tissues that are not apparent and cannot be revealed reliably by standard imaging techniques. Diffusion MRI can make visible these co-lateral damages to the fibers of the CNS white matter that connect different brain regions. This is why in our research, Diffusion MRI is the structural imaging modality that will be considered to recover the CNS connectivity.

4.2. Applications of M/EEG

Clinical domain: Diagnosis of neurological disorders
The dream of all M/EEG researchers is to alleviate the need for invasive recordings (electrocorticograms or intracerebral electrodes), which are often necessary prior to brain surgery, in order to precisely locate both pathological and vital functional areas. We are involved in this quest, particularly through our collaborations with the La Timone hospital in Marseille.
Subtopics include:

- Diagnosis of neurological disorders such as epilepsy, schizophrenia, tinnitus, ...
- Presurgical planning of brain surgery.
- Collaboration with the *Institut de Neurosciences des Systèmes* on these topics http://ins.univ-amu.fr/fr/.

Cognitive research

- Aims at better understanding the brain spatio-temporal organisation.
- Collaboration with the *Laboratory for Neurobiology of Cognition* in order to develop methods that suit their needs for sophisticated data analysis.

Brain Computer Interfaces (BCI) aim to allow direct control of external devices using brain signals such as measured through EEG. In our project, BCI can be seen as an application of EEG processing techniques, but also as an object of fundamental and applied research as they open the way for more dynamical and active brain cognitive protocols.

We develop a research collaboration with the eemagine/ANT-Neuro company. We collaborate with Nice University Hospital on the usage of BCI-based communication for ALS[0] patients.

---

[0]Nice University Hospital hosts a regional reference center for patients suffering from Amyotrophic Lateral Sclerosis

BIOCORE Project-Team

4. Application Domains

4.1. Bioenergy

Finding sources of renewable energy is a key challenge for our society. We contribute to this topic through two main domains for which a strong and acknowledged expertise has been acquired over the years. First, we consider anaerobic digesters, the field of expertise of the members of the team at the Laboratory of Enviromental Biotechnology (LBE), for the production of methane and/or biohydrogen from organic wastes. The main difficulty is to make these processes more reliable and exploit more efficiently the produced biogas by regulating both its quality and quantity despite high variability in the influent wastes. One of the specific applications that needs to be tackled is the production of biogas in a plant when the incoming organic waste results from the mixing of a finite number of substrates. The development of control laws that optimize the input mix of the substrates as a function of the actual state of the system is a key challenge for the viability of this industry.

The second topic consists in growing microalgae, the field of expertise of the members of the team at the Oceanographic Laboratory of Villefranche-sur-Mer (LOV), to produce biofuel. These microorganisms can synthesize lipids with a much higher productivity than terrestrial oleaginous species. The difficulty is to better understand the involved processes, which are mainly transient, to stimulate and optimize them on the basis of modeling and control strategies. Predicting and optimizing the productivity reached by these promising systems in conditions where light received by each cell is strongly related to hydrodynamics, is a crucial challenge.

Finally, for the energy balance of the process, it is important to couple microalgae and anaerobic digestion to optimize the solar energy that can be recovered from microalgae, as was explored within the ANR Symbiose project (2009-2012) [3].

4.2. CO$_2$ fixation and fluxes

Phytoplanktonic species, which assimilate CO_2 during photosynthesis, have received a lot of attention in the last years. Microalgal based processes have been developed in order to mitigate industrial CO_2. As for biofuel productions, many problems arise when dealing with microalgae which are more complex than bacteria or yeasts. Several models have been developed within our team to predict the CO_2 uptake in conditions of variable light and nitrogen availability. The first modeling challenge in that context consists in taking temperature effects and light gradient into account.

The second challenge consists in exploiting the microalgal bioreactors which have been developed in the framework of the quantification of carbon fluxes between ocean and atmospheres. The SEMPO platform (simulator of variable environment computer controlled), developed within the LOV team, has been designed to reproduce natural conditions that can take place in the sea and to accurately measure the cells behavior. This platform, for which our team has developed models and control methods over the years, is an original and unique tool to develop relevant models which stay valid in dynamic conditions. It is worth noting that a better knowledge of the photosynthetic mechanisms and improved photosynthesis models will benefit both thematics: CO_2 mitigation and carbon fluxes predictions in the sea.

4.3. Biological control for plants and micro-plants production systems

This research concentrates on the protection of cultures of photosynthetic organisms against their pests or their competitors. The cultures we study are crop and micro-algae productions. In both cases, the devices are more or less open to the outside, depending on the application (greenhouse/field, photobioreactor/raceway), so that they may give access to harmful pathogens and invading species. We opt for protecting the culture through the use of biocontrol in a broad sense.

In crop production, biocontrol is indeed a very promising alternative to reduce pesticide use: it helps protecting the environment, as well as the health of consumers and producers; it limits the development of resistance (in comparison to chemicals). The use of biocontrol agents, which are, generically, natural enemies (predators, parasitoids or pathogens) of crop pests [89], is however not widespread yet because it often lacks efficiency in real-life crop production systems (while its efficiency in the laboratory is much higher) and can fail to be economically competitive. Resistant crops are also used instead of pesticides to control pests and pathogens, but the latter eventually more or less rapidly overcome the resistance, so these crops need to be replaced by new resistant crops. As resistant genes are a potentially limited resource, a challenge is to ensure the durability of crop resistance. Our objective is to propose models that would help to explain which factors are locks that prevent the smooth transition from the laboratory to the agricultural crop, as well as develop new methods for the optimal deployment of the pests natural enemies and of crop resistance.

Microalgae production is faced with exactly the same problems since predators of the produced microalgae (e.g. zooplankton) or simply other species of microalgae can invade the photobioreactors and outcompete or eradicate the one that we wish to produce. Methods need therefore to be proposed for fighting the invading species; this could be done by introducing predators of the pest and so keeping it under control, or by controling the conditions of culture in order to reduce the possibility of invasion; the design of such methods could greatly take advantage of our knowledge developed in crop protection since the problems and models are related.

4.4. Biological depollution

These works will be carried out with the LBE , mainly on anaerobic treatment plants. This process, despite its strong advantages (methane production and reduced sludge production) can have several locally stable equilibria. In this sense, proposing reliable strategies to stabilize and optimise this process is a key issue. Because of the recent (re)development of anaerobic digestion, it is crucial to propose validated supervision algorithms for this technology. A problem of growing importance is to take benefit of various waste sources in order to adapt the substrate quality to the bacterial biomass activity and finally optimize the process. This generates new research topics for designing strategies to manage the fluxes of the various substrate sources meeting at the same time the depollution norms and providing a biogas of constant quality. In the past years, we have developed models of increasing complexity. However there is a key step that must be considered in the future: how to integrate the knowledge of the metabolisms in such models which represent the evolution of several hundreds bacterial species? How to improve the models integrating this two dimensional levels of complexity? With this perspective, we wish to better represent the competition between the bacterial species, and drive this competition in order to maintain, in the process, the species with the highest depollution capability. This approach, initiated in [92] must be extended from a theoretical point of view and validated experimentally.

BIOVISION Project-Team

4. Application Domains

4.1. Applications of virtual/augmented reality for low-vision

- **Rehabilitation**: Serious games are games designed for a primary purpose which is not pure entertainment. In our context, we think about serious games as a way to help low-vision patients in performing rehabilitation exercises. Virtual/augmented reality technology is a promising platform to develop such rehabilitation exercises targeted to specific pathologies. For example, with Age Macular Degeneration (AMD), our objective is to propose solutions allowing rehabilitation of visuo-perceptual-motor functions to optimally use residual portions of the peripheral retina and obtain efficient eccentric viewing.

- **Vision aid-systems**: A variety of aids for low-vision people are already on the market using different kinds of virtual/augmented reality platforms (dedicated or large public ones). They offer different functionalities (magnification, image enhancement, text to speech, face and object recognition). Our goal is to design new solutions allowing autonomous interaction in mixed reality environments, and take advantage of the improvement of functions obtained via rehabilitation protocols.

- **Cognitive research**: Virtual/augmented reality technology represents a new opportunity to conduct cognitive and behavioural research using virtual environments where all parameters can be psychophysically controlled. Our objective is to re-assess common theories by allowing patients to freely explore their environment in more ecological conditions.

4.2. Applications of vision modeling studies

- **Neuroscience research.** Making in-silico experiments is a way to reduce the experimental costs, to test hypotheses and design models, and to test algorithms. Our goal is to develop a large-scale simulations platform of impaired retinas, called Macular, allowing to mimic specific degeneracies or pharmacologically induced impairments, as well as to emulate electric stimulation by prostheses. In addition, the platform provides a realistic entry to models or simulators of the thalamus or the visual cortex, in contrast to the entries usually considered in modelling studies.

- **Education.** Macular is also targeted as a useful tool for educational purposes, illustrating for students how the retina works and responds to visual stimuli.

CAMIN Project-Team (section vide)

CASTOR Project-Team (section vide)

CHORALE Team

4. Application Domains

4.1. Transportation of people and goods

The researches developed in CHORALE can be applied in different applications fields. We are particularly interested in *transportation of people and goods*. CHORALE contribute to the development of Autonomous Connected Vehicles (e.g. Learning, Mapping, Localization, Navigation) and the associated services (e.g. towing, platooning, taxi).

4.2. Assistance and services robots

We are also interested in *assistance and service robotics*. By adding the social level in the representation of the environment, CHORALE develops social and proactive navigation.

4.3. Survey and monitoring of environments

CHORALE studies multi-robots systems to explore, survey and monitor different kind of environments (e.g. agricultural space, forest space, destroyed space), which are of particular importance in the field of *Exploration and monitoring of poorly structured environments*.

COATI Project-Team

4. Application Domains

4.1. Telecommunication Networks

COATI is mostly interested in telecommunications networks but also in the network structure appearing in social, molecular and transportation networks.

We focus on the design and management of heterogeneous physical and logical networks. The project has kept working on the design of backbone networks (optical networks, radio networks, IP networks). However, the fields of Software Defined Networks and Network Function Virtualization are growing in importance in our studies. In all these networks, we study routing algorithms and the evolution of the routing in case of any kind of topological modifications (maintenance operations, failures, capacity variations, etc.).

4.2. Other Domains

Our combinatorial tools may be well applied to solve many other problems in various areas (transport, biology, resource allocation, chemistry, smart-grids, speleology, etc.) and we collaborate with experts of some of these domains.

For instance, we collaborate with project-team ABS (Algorithms Biology Structure) from Sophia Antipolis on problems from Structural Biology (co-supervision of a PhD student). In the area of transportation networks, we collaborate with SMEs Benomad and Instant-System on dynamic car-pooling combined with multi-modal transportation systems in the context of ANR project Multimod started in January 2018. Last, we have started a collaboration with GREDEG (Groupe de Recherche en Droit, Economie et Gestion, Univ. Nice Sophia Antipolis) on the analysis of collaboration networks.

COFFEE Project-Team

4. Application Domains

4.1. Multiphase porous media flows and multi-physics coupling

Our research focuses on the numerical modeling of multiphase porous media flows accounting for complex geology and for nonlinear and multi-physics couplings. It is applied to various problems in the field of energy such as the simulation of geothermal systems in collaboration with BRGM, of nuclear waste repositories in collaboration with Andra, and of oil and gas recovery in collaboration with Total. Our research directions include the development of advanced numerical schemes adapted to polyhedral meshes and highly heterogeneous media in order to represent more accurately complex geologies. A special focus is made on the modeling of multiphase flows in network of faults or fractures represented as interfaces of co-dimension one coupled to the surrounding matrix. We also investigate nonlinear solvers adapted to the nonlinear couplings between gravity, capillary and viscous forces in highly heterogeneous porous media. In the same line, we study new domain decomposition algorithms to couple non-isothermal compositional liquid gas flows in a porous medium with free gas flows occurring at the interface between the ventilation gallery and the nuclear waste repository or between a geothermal reservoir and the atmosphere. We have begun exploring the coupling between the multiphase flow in the porous matrix and the solid mechanics involved in opening fractures.

4.2. Particulate and mixture flows

We investigate fluid mechanics models referred to as "multi–fluids" flows. A large part of our activity is more specifically concerned with the case where a disperse phase interacts with a dense phase. Such flows arise in numerous applications, like for pollutant transport and dispersion, the combustion of fuel particles in air, the modelling of fluidized beds, the dynamic of sprays and in particular biosprays with medical applications, engine fine particles emission... There are many possible modelings of such flows: microscopic models where the two phases occupy distinct domains and where the coupling arises through intricate interface conditions; macroscopic models which are of hydrodynamic (multiphase) type, involving non standard state laws, possibly with non conservative terms, and the so–called mesoscopic models. The latter are based on Eulerian–Lagrangian description where the disperse phase is described by a particle distribution function in phase space. Following this path we are led to a Vlasov-like equation coupled to a system describing the evolution of the dense phase that is either the Euler or the Navier-Stokes equations. It turns out that the leading effect in such models is the drag force. However, the role of other terms, of more or less phenomenological nature, deserves to be discussed (close packing terms, lift term, Basset force...). Of course the fluid/kinetic model is interesting in itself and needs further analysis and dedicated numerical schemes. In particular, in collaboration with the Atomic Energy Commission (CEA), we have proposed a semi-Lagrangian scheme for the simulation of particulate flows, extending the framework established in plasma physics to such flows.

We also think it is worthwhile to identify hydrodynamic regimes: it leads to discuss hierarchies of coupled hydrodynamic systems, the nature of which could be quite intriguing and original, while they share some common features of the porous media problems. We are particularly interested in revisiting the modeling of mixture flows through the viewpoint of kinetic models and hydrodynamic regimes. We propose to revisit the derivation of new mixture models, generalizing Kazhikov-Smagulov equations, through hydrodynamic asymptotics. The model is of "hybrid" type in the sense that the constraint reduces to the standard incompressibility condition when the disperse phase is absent, while it involves derivatives of the particle volume fraction when the disperse phase is present.

4.3. Fungal network growth

Members of the team have started an original research program devoted to fungal network growth. We started working on this subject through a collaboration with biologists and physicists at LIED (Université Paris Diderot) and probabilists in CMAP (Ecole Polytechnique) and Université Paris Sud, involving Rémi Catellier and Yves D'Angelo in LJAD in Nice. The motivation is to understand branching networks as an efficient space exploration strategy, with fungus *Podospora Anserina* being the biological model considered. This research is submitted as an ANR-project and has been supported by various local fundings.

DATASHAPE Project-Team (section vide)

DIANA Project-Team (section vide)

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

4. Application Domains

4.1. Algorithmic Differentiation

Algorithmic Differentiation of programs gives sensitivities or gradients, useful for instance for :

- optimum shape design under constraints, multidisciplinary optimization, and more generally any algorithm based on local linearization,
- inverse problems, such as parameter estimation and in particular 4Dvar data assimilation in climate sciences (meteorology, oceanography),
- first-order linearization of complex systems, or higher-order simulations, yielding reduced models for simulation of complex systems around a given state,
- adaption of parameters for classification tools such as Machine Learning systems, in which Adjoint Differentiation is also known as *backpropagation*.
- mesh adaptation and mesh optimization with gradients or adjoints,
- equation solving with the Newton method,
- sensitivity analysis, propagation of truncation errors.

4.2. Multidisciplinary optimization

A CFD program computes the flow around a shape, starting from a number of inputs that define the shape and other parameters. On this flow one can define optimization criteria e.g. the lift of an aircraft. To optimize a criterion by a gradient descent, one needs the gradient of the criterion with respect to all inputs, and possibly additional gradients when there are constraints. Adjoint AD is the most efficient way to compute these gradients.

4.3. Inverse problems and Data Assimilation

Inverse problems aim at estimating the value of hidden parameters from other measurable values, that depend on the hidden parameters through a system of equations. For example, the hidden parameter might be the shape of the ocean floor, and the measurable values of the altitude and velocities of the surface. Figure 1 shows an example of an inverse problem using the glaciology code ALIF (a pure C version of ISSM [30]) and its AD-adjoint produced by Tapenade.

One particular case of inverse problems is *data assimilation* [31] in weather forecasting or in oceanography. The quality of the initial state of the simulation conditions the quality of the prediction. But this initial state is not well known. Only some measurements at arbitrary places and times are available. A good initial state is found by solving a least squares problem between the measurements and a guessed initial state which itself must verify the equations of meteorology. This boils down to solving an adjoint problem, which can be done though AD [33]. The special case of *4Dvar* data assimilation is particularly challenging. The 4th dimension in "4D" is time, as available measurements are distributed over a given assimilation period. Therefore the least squares mechanism must be applied to a simulation over time that follows the time evolution model. This process gives a much better estimation of the initial state, because both position and time of measurements are taken into account. On the other hand, the adjoint problem involved is more complex, because it must run (backwards) over many time steps. This demanding application of AD justifies our efforts in reducing the runtime and memory costs of AD adjoint codes.
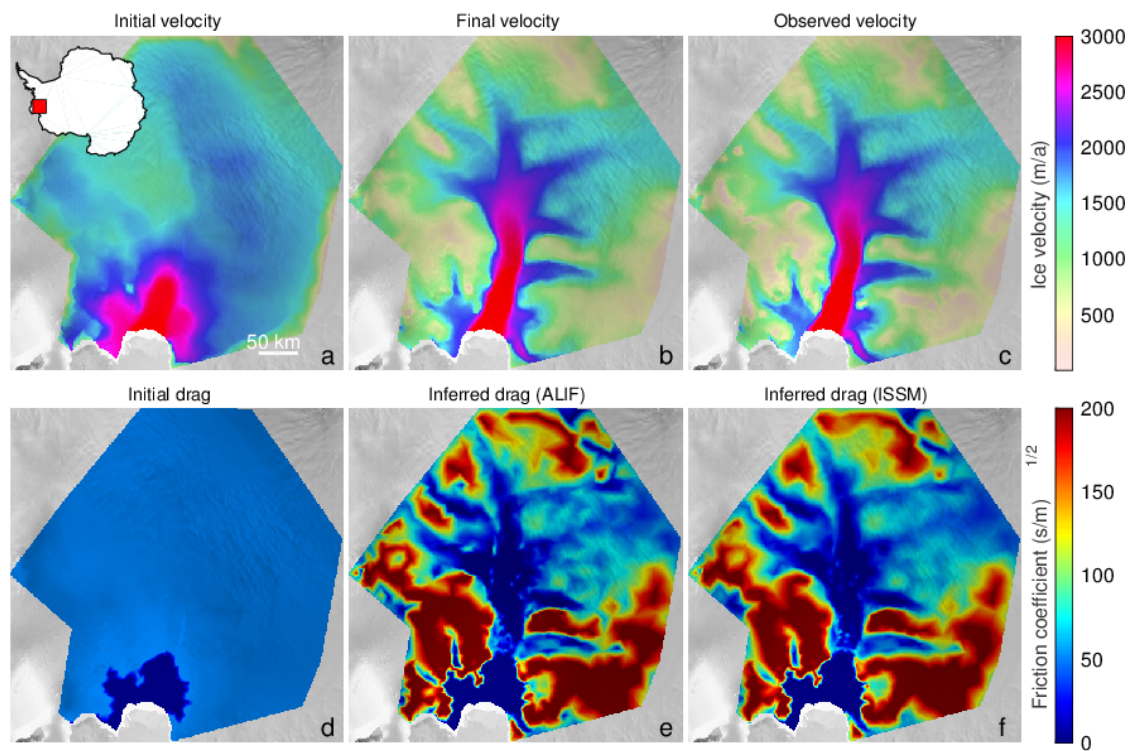
Figure 1. Assimilation of the basal friction under Pine Island glacier, West Antarctica. The final simulated surface velocity (b) is made to match the observed surface velocity (c), by estimation of the basal friction (e). A reference basal friction (f) is obtained by another data assimilation using the hand=written adjoint of ISSM

4.4. Linearization

Simulating a complex system often requires solving a system of Partial Differential Equations. This can be too expensive, in particular for real-time simulations. When one wants to simulate the reaction of this complex system to small perturbations around a fixed set of parameters, there is an efficient approximation: just suppose that the system is linear in a small neighborhood of the current set of parameters. The reaction of the system is thus approximated by a simple product of the variation of the parameters with the Jacobian matrix of the system. This Jacobian matrix can be obtained by AD. This is especially cheap when the Jacobian matrix is sparse. The simulation can be improved further by introducing higher-order derivatives, such as Taylor expansions, which can also be computed through AD. The result is often called a *reduced model*.

4.5. Mesh adaptation

Some approximation errors can be expressed by an adjoint state. Mesh adaptation can benefit from this. The classical optimization step can give an optimization direction not only for the control parameters, but also for the approximation parameters, and in particular the mesh geometry. The ultimate goal is to obtain optimal control parameters up to a precision prescribed in advance.

EPIONE Project-Team (section vide)

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

4. Application Domains

4.1. Introduction

Application domains are naturally linked to the problems described in Sections 3.2.1 and 3.2.2 . By and large, they split into a systems-and-circuits part and an inverse-source-and-boundary-problems part, united under a common umbrella of function-theoretic techniques as described in Section 3.3 .

4.2. Inverse magnetization problems

Participants: Laurent Baratchart, Sylvain Chevillard, Juliette Leblond, Konstantinos Mavreas.

Generally speaking, inverse potential problems, similar to the one appearing in Section 4.3 , occur naturally in connection with systems governed by Maxwell's equation in the quasi-static approximation regime. In particular, they arise in magnetic reconstruction issues. A specific application is to geophysics, which led us to form the Inria Associate Team IMPINGE (Inverse Magnetization Problems IN GEosciences) together with MIT and Vanderbilt University. Though this Associate Team reached the end of its term in 2018, the collaborations it has generated are still active. A joint work with Cerege (CNRS, Aix-en-Provence), in the framework of the ANR-project MagLune, completes this picture, see Sections 6.1.2 , 8.2.1 .

To set up the context, recall that the Earth's geomagnetic field is generated by convection of the liquid metallic core (geodynamo) and that rocks become magnetized by the ambient field as they are formed or after subsequent alteration. Their remanent magnetization provides records of past variations of the geodynamo, which is used to study important processes in Earth sciences like motion of tectonic plates and geomagnetic reversals. Rocks from Mars, the Moon, and asteroids also contain remanent magnetization which indicates the past presence of core dynamos. Magnetization in meteorites may even record fields produced by the young sun and the protoplanetary disk which may have played a key role in solar system formation.

For a long time, paleomagnetic techniques were only capable of analyzing bulk samples and compute their net magnetic moment. The development of SQUID microscopes has recently extended the spatial resolution to sub-millimeter scales, raising new physical and algorithmic challenges. The associate team IMPINGE aims at tackling them, experimenting with the SQUID microscope set up in the Paleomagnetism Laboratory of the department of Earth, Atmospheric and Planetary Sciences at MIT. Typically, pieces of rock are sanded down to a thin slab, and the magnetization has to be recovered from the field measured on a planar region at small distance from the slab.

Mathematically speaking, both inverse source problems for EEG from Section 4.3 and inverse magnetization problems described presently amount to recover the (3-D valued) quantity m (primary current density in case of the brain or magnetization in case of a thin slab of rock) from measurements of the potential:

$$V(x) = \int_\Omega \frac{\operatorname{div} m(x')\, dx'}{|x - x'|}\,, \tag{1}$$

outside the volume Ω of the object. Depending on the geometry of models, the magnetization distribution m may lie in a volume or spread out on a surface. This results in quite different identifiability properties, see [37] and Section 6.1.1 , but the two situations share a substantial mathematical common core.

Another timely instance of inverse magnetization problems lies with geomagnetism. Satellites orbiting around the Earth measure the magnetic field at many points, and nowadays it is a challenge to extract global information from those measurements. In collaboration with C. Gerhards (Geomathematics and Geoinformatics Group, Technische Universität Bergakademie Freiberg, Germany), we started to work on the problem of separating the magnetic field due to the magnetization of the globe's crust from the magnetic field due to convection in the liquid metallic core. The techniques involved are variants, in a spherical context, from those developed within the IMPINGE associate team for paleomagnetism, see Section 6.1.1 .

4.3. Inverse source problems in EEG

Participants: Paul Asensio, Laurent Baratchart, Juliette Leblond, Jean-Paul Marmorat, Masimba Nemaire.

Solving overdetermined Cauchy problems for the Laplace equation on a spherical layer (in 3-D) in order to extrapolate incomplete data (see Section 3.2.1) is a necessary ingredient of the team's approach to inverse source problems, in particular for applications to EEG, see [9]. Indeed, the latter involves propagating the initial conditions through several layers of different conductivities, from the boundary shell down to the center of the domain where the singularities (*i.e.* the sources) lie. Once propagated to the innermost sphere, it turns out that traces of the boundary data on 2-D cross sections coincide with analytic functions with branched singularities in the slicing plane [8], [42]. The singularities are related to the actual location of the sources, namely their moduli reach in turn a maximum when the plane contains one of the sources. Hence we are back to the 2-D framework of Section 3.3.3 , and recovering these singularities can be performed *via* best rational approximation. The goal is to produce a fast and sufficiently accurate initial guess on the number and location of the sources in order to run heavier descent algorithms on the direct problem, which are more precise but computationally costly and often fail to converge if not properly initialized. Our belief is that such a localization process can add a geometric, valuable piece of information to the standard temporal analysis of EEG signal records.

Numerical experiments obtained with our software FindSources3D give very good results on simulated data and we are now engaged in the process of handling real experimental data, simultaneously recorded by EEG and MEG devices, in collaboration with our partners at INS, hospital la Timone, Marseille (see Section 6.1.3).

Furthermore, another approach is being studied for EEG, that consists in regularizing the inverse source problem by a total variation constraint on the source term (a measure), added to the quadratic data approximation criterion. It is similar to the path that is taken for inverse magnetization problems (see Sections 4.2 and 6.1.1), and it presently focuses on surface-distributed models.

4.4. Identification and design of microwave devices

Participants: Laurent Baratchart, Sylvain Chevillard, Jean-Paul Marmorat, Martine Olivi, Fabien Seyfert.

This is joint work with Stéphane Bila (Xlim, Limoges).

One of the best training grounds for function-theoretic applications by the team is the identification and design of physical systems whose performance is assessed frequency-wise. This is the case of electromagnetic resonant systems which are of common use in telecommunications.

In space telecommunications (satellite transmissions), constraints specific to on-board technology lead to the use of filters with resonant cavities in the microwave range. These filters serve multiplexing purposes (before or after amplification), and consist of a sequence of cylindrical hollow bodies, magnetically coupled by irises (orthogonal double slits). The electromagnetic wave that traverses the cavities satisfies the Maxwell equations, forcing the tangent electrical field along the body of the cavity to be zero. A deeper study of the Helmholtz equation states that an essentially discrete set of wave vectors is selected. In the considered range of frequency, the electrical field in each cavity can be decomposed along two orthogonal modes, perpendicular to the axis of the cavity (other modes are far off in the frequency domain, and their influence can be neglected).

Each cavity (see Figure 1) has three screws, horizontal, vertical and midway (horizontal and vertical are two arbitrary directions, the third direction makes an angle of 45 or 135 degrees, the easy case is when all cavities show the same orientation, and when the directions of the irises are the same, as well as the input and output slits). Since screws are conductors, they behave as capacitors; besides, the electrical field on the surface has to be zero, which modifies the boundary conditions of one of the two modes (for the other mode, the electrical field is zero hence it is not influenced by the screw), the third screw acts as a coupling between the two modes. The effect of an iris is opposite to that of a screw: no condition is imposed on a hole, which results in a coupling between two horizontal (or two vertical) modes of adjacent cavities (in fact the iris is the union of two rectangles, the important parameter being their width). The design of a filter consists in finding the size of
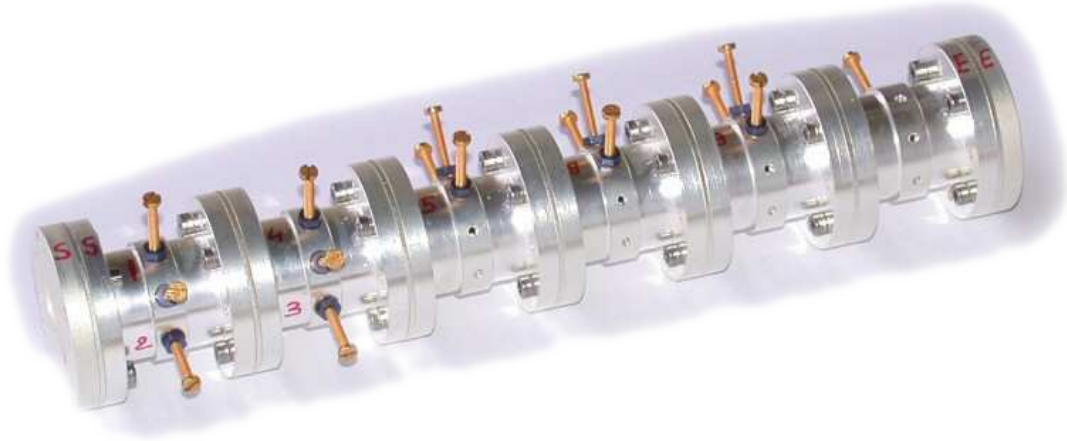
Figure 1. Picture of a 6-cavities dual mode filter. Each cavity (except the last one) has 3 screws to couple the modes within the cavity, so that 16 quantities must be optimized. Quantities such as the diameter and length of the cavities, or the width of the 11 slits are fixed during the design phase.

each cavity, and the width of each iris. Subsequently, the filter can be constructed and tuned by adjusting the screws. Finally, the screws are glued once a satisfactory response has been obtained. In what follows, we shall consider a typical example, a filter designed by the CNES in Toulouse, with four cavities near 11 GHz.

Near the resonance frequency, a good approximation to the Helmholtz equations is given by a second order differential equation. Thus, one obtains an electrical model of the filter as a sequence of electrically-coupled resonant circuits, each circuit being modeled by two resonators, one per mode, the resonance frequency of which represents the frequency of a mode, and whose resistance accounts for electric losses (surface currents) in the cavities.

This way, the filter can be seen as a quadripole, with two ports, when plugged onto a resistor at one end and fed with some potential at the other end. One is interested in the power which is transmitted and reflected. This leads one to define a scattering matrix S, which may be considered as the transfer function of a stable causal linear dynamical system, with two inputs and two outputs. Its diagonal terms $S_{1,1}$, $S_{2,2}$ correspond to reflections at each port, while $S_{1,2}$, $S_{2,1}$ correspond to transmission. These functions can be measured at certain frequencies (on the imaginary axis). The matrix S is approximately rational of order 4 times the number of cavities (that is 16 in the example on Figure 2), and the key step consists in expressing the components of the equivalent electrical circuit as functions of the S_{ij} (since there are no formulas expressing the lengths of the screws in terms of parameters of this electrical model). This representation is also useful to analyze the numerical simulations of the Maxwell equations, and to check the quality of a design, in particular the absence of higher resonant modes.

In fact, resonance is not studied via the electrical model, but via a low-pass equivalent circuit obtained upon linearizing near the central frequency, which is no longer conjugate symmetric (*i.e.* the underlying system may no longer have real coefficients) but whose degree is divided by 2 (8 in the example).

In short, the strategy for identification is as follows:

- measuring the scattering matrix of the filter near the optimal frequency over twice the pass band (which is 80MHz in the example).

- Solving bounded extremal problems for the transmission and the reflection (the modulus of he response being respectively close to 0 and 1 outside the interval measurement, cf. Section 3.3.1) in order to get a models for the scattering matrix as an analytic matrix-valued function. This provides us with a scattering matrix known to be close to a rational matrix of order roughly 1/4 of the number of data points.

- Approximating this scattering matrix by a true rational transfer-function of appropriate degree (8 in this example) via the Endymion or RARL2 software (cf. Section 3.3.2.2).

- A state space realization of S, viewed as a transfer function, can then be obtained, where additional symmetry constraints coming from the reciprocity law and possibly other physical features of the device have to be imposed.

- Finally one builds a realization of the approximant and looks for a change of variables that eliminates non-physical couplings. This is obtained by using algebraic-solvers and continuation algorithms on the group of orthogonal complex matrices (symmetry forces this type of transformation).
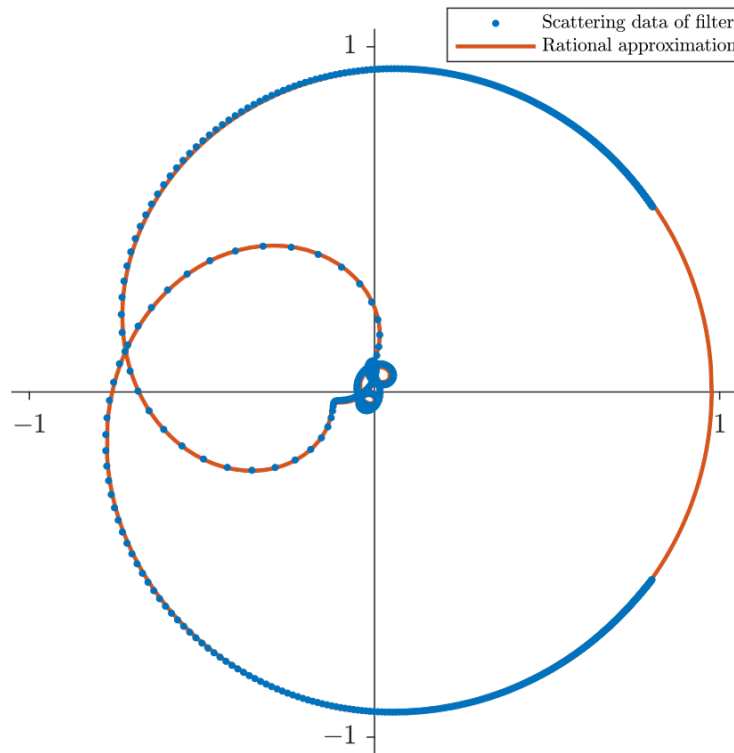

Figure 2. Nyquist Diagram of S_{22}. The rational approximation is of degree 8.

The final approximation is of high quality. This can be interpreted as a confirmation of the linearity assumption on the system: the relative L^2 error is less than 10^{-3}. This is illustrated by a reflection diagram (Figure 2).

The above considerations are valid for a large class of filters. These developments have also been used for the design of non-symmetric filters, which are useful for the synthesis of repeating devices.

The team further investigates problems relative to the design of optimal responses for microwave devices. The resolution of a quasi-convex Zolotarev problems was proposed, in order to derive guaranteed optimal multi-band filter responses subject to modulus constraints [12]. This generalizes the classical single band

design techniques based on Chebyshev polynomials and elliptic functions. The approach relies on the fact that the modulus of the scattering parameter $|S_{1,2}|$ admits a simple expression in terms of the filtering function $D = |S_{1,1}|/|S_{1,2}|$, namely

$$|S_{1,2}|^2 = \frac{1}{1 + D^2}.$$

The filtering function appears to be the ratio of two polynomials p_1/p_2, the numerator of the reflection and transmission scattering factors, that may be chosen freely. The denominator q is then obtained as the unique stable unitary polynomial solving the classical Feldtkeller spectral equation:

$$qq^* = p_1p_1^* + p_2p_2^*.$$

The relative simplicity of the derivation of a filter's response, under modulus constraints, owes much to the possibility of forgetting about Feldtkeller's equation and express all design constraints in terms of the filtering function. This no longer the case when considering the synthesis N-port devices for $N > 3$, like multiplexers, routers and power dividers, or when considering the synthesis of filters under matching conditions. The efficient derivation of multiplexers responses is the subject of recent investigation by Factas, using techniques based on constrained Nevanlinna-Pick interpolation (see Section 6.2).

Through contacts with CNES (Toulouse) and UPV (Bilbao), Apics got additionally involved in the design of amplifiers which, unlike filters, are active devices. A prominent issue here is stability. A twenty years back, it was not possible to simulate unstable responses, and only after building a device could one detect instability. The advent of so-called *harmonic balance* techniques, which compute steady state responses of linear elements in the frequency domain and look for a periodic state in the time domain of a network connecting these linear elements *via* static non-linearities made it possible to compute the harmonic response of a (possibly nonlinear and unstable) device [79]. This has had tremendous impact on design, and there is a growing demand for software analyzers. The team is also becoming active in this area.

In this connection, there are two types of stability involved. The first is stability of a fixed point around which the linearized transfer function accounts for small signal amplification. The second is stability of a limit cycle which is reached when the input signal is no longer small and truly nonlinear amplification is attained (*e.g.* because of saturation). Applications by the team so far have been concerned with the first type of stability, and emphasis is put on defining and extracting the "unstable part" of the response, see Section 6.3 . The stability check for limit cycles has made important theoretical advances, and numerical algorithms are now under investigation.

<div align="center">

FOCUS Project-Team

</div>

4. Application Domains

4.1. Ubiquitous Systems

The main application domain for Focus are ubiquitous systems, broadly systems whose distinctive features are: mobility, high dynamicity, heterogeneity, variable availability (the availability of services offered by the constituent parts of a system may fluctuate, and similarly the guarantees offered by single components may not be the same all the time), open-endedness, complexity (the systems are made by a large number of components, with sophisticated architectural structures). In Focus we are particularly interested in the following aspects.

- *Linguistic primitives* for programming dialogues among components.
- *Contracts* expressing the functionalities offered by components.
- *Adaptability and evolvability* of the behaviour of components.
- *Verification* of properties of component systems.
- Bounds on component *resource consumption* (e.g., time and space consumed).

4.2. Service Oriented Computing and Cloud Computing

Today the component-based methodology often refers to Service Oriented Computing. This is a specialized form of component-based approach. According to W3C, a service-oriented architecture is "a set of components which can be invoked, and whose interface descriptions can be published and discovered". In the early days of Service Oriented Computing, the term services was strictly related to that of Web Services. Nowadays, it has a much broader meaning as exemplified by the XaaS (everything as a service) paradigm: based on modern virtualization technologies, Cloud computing offers the possibility to build sophisticated service systems on virtualized infrastructures accessible from everywhere and from any kind of computing device. Such infrastructures are usually examples of sophisticated service oriented architectures that, differently from traditional service systems, should also be capable to elastically adapt on demand to the user requests.

GRAPHDECO Project-Team (section vide)

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

4. Application Domains

4.1. Agronomy

Agronomy is a strong expertise domain in the area of Montpellier. Some members of GraphIK are INRA researchers (computer scientists). We closely collaborate with the Montpellier research laboratory IATE, a join unit of INRA and other organisms. A major issue for INRA and more specifically IATE applications is modeling agrifood chains (i.e., the chain of all processes leading from the plants to the final products, including waste treatment). This modeling has several objectives. It provides better understanding of the processes from begin to end, which aids in decision making, with the aim of improving the quality of the products and decreasing the environmental impact. It also facilitates knowledge sharing between researchers, as well as the capitalization of expert knowledge and "know how". This last point is particularly important in areas strongly related to local know how (like in cheese or wine making), where knowledge is transmitted by experience, with the risk of non-sustainability of the specific skills. An agrifood chain analysis is a highly complex procedure since it relies on numerous criteria of various types: environmental, economical, functional, sanitary, etc. Quality objectives involve different stakeholders, technicians, managers, professional organizations, end-users, public organizations, etc. Since the goals of the implied stakeholders may be divergent dedicated knowledge and representation techniques are to be employed.

4.2. Data Journalism

One of today's major issues in data science is to design techniques and algorithms that allow analysts to efficiently infer useful information and knowledge by inspecting heterogeneous information sources, from structured data to unstructured content. We take data journalism as an emblematic use-case, which stands at the crossroad of multiple research fields: content analysis, data management, knowledge representation and reasoning, visualization and human-machine interaction. We are particularly interested in issues raised by the design of data and knowledge management systems that will support data journalism. These systems include an ontology (which typically expresses domain knowledge), heterogeneous data sources (provided with their own vocabulary and querying capabilities), and mappings that relate these data sources to the ontological vocabulary. Ontologies play a central role as they act both as a mediation layer that glue together pieces of knowledge extracted from data sources, and as an inference layer that allow to draw new knowledge.

Besides pure knowledge representation and reasoning issues, querying such systems raise issues at the crossroad of data and knowledge management. In particular, although mappings have been widely investigated in databases, they need to be revisited in the light of the reasoning capabilities enabled by the ontology. More generally, the consistency and the efficiency of the system cannot be ensured by considering the components of the system in isolation (i.e., the ontology, data sources and mappings), but require to study the interactions between these components and to consider the system as a whole.

HEPHAISTOS Project-Team (section vide)

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

4. Application Domains

4.1. Web

The Web is the natural application domain of the team. We are designing and implementing multitier languages for helping the development of Web applications. We are creating static and dynamic analyses for Web security. We are conducting empirical studies about privacy preservation on the Web.

4.2. Internet of Things

More recently, we have started focusing on *Internet of Things* (IoT) applications. They share many similarities with Web applications so most of the methodologies and expertises we have developed for the Web apply to IoT but the restricted hardware resources made available by many IoT devices demand new developments and new research explorations.

Kairos Project-Team

4. Application Domains

4.1. Cyber-Physical and Embedded Systems

We have historical contacts with industrial and academic partners in the domains of avionics and embedded electronics (Airbus, Thales, Safran). We have new collaborations in the fields of satellites (Thales Alenia Space) and connected cars (Renault Software Labs). These provide for use case and new issues in CPS co-modeling and co-design (Digital Twins) further described in New Results section.

4.2. Connected Objects in the Internet Of Things

Due to increasing collaborations with local partners, we have recently considered Smart Contracts (as popularized in Blockchain frameworks), as a way to formally established specification of behavioral system traces, applied to connected objects in a IoT environment. The new ANR project SIM is based on this.

LEMON Project-Team (section vide)

MATHNEURO Project-Team (section vide)

<div align="center">

MCTAO Project-Team

</div>

4. Application Domains

4.1. Aerospace Engineering

Participants: Bernard Bonnard, Jean-Baptiste Caillau, Thierry d'Argent, Lamberto Dell'Elce, Jean-Baptiste Pomet, Jérémy Rouot.

Space engineering is very demanding in terms of safe and high-performance control laws. It is therefore prone to fruitful industrial collaborations. McTAO now has an established expertise in space and celestial mechanics. Our collaborations with industry are mostly on orbit transfer problems with low-thrust propulsion. It can be orbit transfer to put a commercial satellite on station, in which case the dynamics are a Newtonian force field plus perturbations and the small control. There is also, currently, a renewed interest in low-thrust missions such as Lisa Pathfinder (ESA mission towards a Lagrange point of the Sun-Earth system) or BepiColombo (joint ESA-JAXA mission towards Mercury). Such missions look more like a controlled multibody system. In all cases the problem involves long orbit transfers, typically with many revolutions around the primary celestial body. When minimizing time, averaging techniques provide a good approximation. Another important criterion in practice is fuel consumption minimization (crucial because only a finite amount of fuel is onboard a satellite for all its "life"), which amounts to L^1-minimization. Both topics are studied by the team. We have a steady relationships with CNES and Thales Alenia Space (Cannes), that have financed or co-financed 3 PhDs and 2 post-docs in the Sophia location of the team in the decade and are a source of inspiration even at the methodological level. Team members also have close connections with Airbus-Safran (Les Mureaux) on launchers. Some of the authoritative papers in the field were written by team members, with an emphasis on the geometric analysis and on algorithms (coupling of shooting and continuation methods). There are also connections with peers more on the applied side, like D. Scheeres (Colorado Center for Astrodynamics Research at Boulder), the group of F. Bernelli (Politecnico Milano), and colleagues from U. Barcelona (A. Farrès, A. Jorba).

4.2. Magnetic resonance imaging (MRI)

Participants: Bernard Bonnard, Alice Nolot [Professeur Classes Préparatoires Troyes], Jérémy Rouot, Joseph Gergaud, Olivier Cots [ENSEEIHT, Toulouse], Stephen Glaser [TU München, Germany], Dominique Sugny [Univ. de Bourgogne].

The starting point of our interest in optimal control for quantum systems was a collaboration with physicist from ICB, University of Burgundy (Dominique Sugny), motivated by an ANR project where we worked on the control of molecular orientation in a dissipative environment using a laser field, and developed optimal control tools, combined with numerical simulations, to analyze the problem for Qubits. This was related to quantum computing rather than MRI. Using this expertise and under the impulse of Prof. S. Glaser and his group (Chemistry, TU München), we investigated Nuclear Magnetic resonance (NMR) for medical imaging (MRI), where the model is the Bloch equation describing the evolution of the Magnetization vector controlled by a magnetic field, but in fine is a specific Qubit model without decoherence. We worked on, and brought strong contributions to, the contrast problem: typically, given two chemical substances that have an importance in medicine, like oxygenated and de-oxygenated blood, find the (time-dependent) magnetic field that will produce the highest difference in brightness between these two species on the image resulting from Nuclear Magnetic Resonance. This has immediate and important industrial applications in medical imaging. Our contacts are with the above mentioned physics academic labs, who are themselves in contact with major companies. The team has produced and is producing important work on this problem. One may find a good overview in [50], a reference book has been published on the topic [54], a very complete numerical study comparing different optimization techniques was performed in [49]. We conduct this project in parallel with S. Glaser team, which validated experimentally the pertinence of the methods, the main achievement being the in vivo experiments

realized at the Creatis team of Insa Lyon showing the interest to use optimal control methods implemented in modern softwares in MRI in order to produce a better image in a shorter time. A goal is to arrive to a cartography of the optimal contrast with respect to the relaxation parameters using LMI techniques and numerical simulations with the Hamapth and Bocop code; note that the theoretical study is connected to the problem of understanding the behavior of the extremal solutions of a controlled pair of Bloch equations, and this is an ambitious task. Also, one of the difficulties to go from the obtained results, checkable on experiments, to practical control laws for production is to deal with magnetic field space inhomogeneities.

4.3. Swimming at low-Reynolds number

Participants: Bernard Bonnard, Yacine El Alaoui-Faris, Laetitia Giraldi, Clément Moreau, Alice Nolot, Jean-Baptiste Pomet, Jérémy Rouot.

Following the historical reference for low Reynolds number locomotion [72], the study of the swimming strategies of micro-organisms is attracting increasing attention in the recent literature. This is both because of the intrinsic biological interest, and for the possible implications these studies may have on the design of bio-inspired artificial replicas reproducing the functionalities of biological systems. In the case of micro-swimmers, the surrounding fluid is dominated by the viscosity effects of the water and becomes reversible. In this regime, it turns out that the infinite dimensional dynamics of the fluid do not have to be retained as state variables, so that the dynamics of a micro-swimmer can be expressed by ordinary differential equations if its shape has a finite number of degrees of freedom. Assuming this finite dimension, and if the control is the rate of deformation, one obtains a control system that is linear (affine without drift) with respect to the controls, *i.e.* the optimal control problem with a quadratic cost defines a sub-Riemannian structure (see section 3.2). This is the case where the shape is "fully actuated", *i.e.* if all the variables describing the shape are angles, there is an actuator on each of these angles. For artificial micro-swimmers, this is usually unrealistic, hence (artificial) magneto-elastic micro-swimmers, that are magnetized in order to be deformed by an external magnetic field. In this case, the control functions are the external magnetic field. In both cases, questions are controllability (straightforward in the fully actuated case), optimal control, possibly path planning. We collaborate with teams that have physical experiments for both.

- In collaboration with D. Takagi and M. Chyba (Univ of Hawaii), this approach is currently at the experimental level for copepod-like swimmer at the university of Hawaii: on the one hand, this zooplankton and its locomotion can be observed, and a robot micro swimmer mimicking a copepod has been constructed, but in fact large enough for direct actuation to be possible, and the low Reynolds number is achieved by using a more viscous fluid. This gives possibilities, through an inverse optimization problem, to determine what cost can be optimised by these crustaceans, see [40], [78], and to validate models on the robot.

- For magneto-elastic micro-robots, Y. El-Alaoui's PhD is co-advised with Stéphane Régnier from the robotics lab ISIR, Univ. Paris 6. Magneto-elastic micro-robots and their magnetic actuation are actually built at ISIR and the aim of the collaboration is to validate models and improve the existing control laws both in performance and in energy; of course, the micro scale does make things difficult.

The questions about optimality of periodic controls raised in section 3.2 are related to these applications for periodic deformations, or strokes, playan important role in locomotion.

4.4. Stability of high frequency amplifiers

Participants: Sébastien Fueyo, Gilles Lebeau, Jean-Baptiste Pomet, Laurent Baratchart [FACTAS project-team].

Nonlinear hyper-frequency amplifiers are ubiquitous in cell phone relays and many other devices. They must be as compact as possible, yielding a more complicated design. Computer Assisted Design tools are extensively used; for a given amplifier design, they provide frequency responses but fail to provide information of the stability of the response for each frequency. This stability is crucial for an unstable response will not be observed in practice; the actual device should not be built before stability is asserted. Predicting stability/instability from "simulations" in the Computer Assisted Design tool is of utmost importance (simulation between quotation marks because these simulations are in fact computations in the frequency domain). Potential transfer is important.

Some techniques do exist, see [77], based on creating some virtual perturbations and treating them as the input of a (linearized) control system to be "simulated" using the same tools. In an ongoing collaboration between McTAO and the project-team FACTAS, we work on the mathematical ground of these methods and in particular of the relation between stability and the property of the identified time-varying infinite dimensional systems. See recent developments in Section 7.14 .

4.5. Optimal control of microbial cells

Participants: Jean-Baptiste Caillau, Walid Djema [BIOCORE project-team], Laetitia Giraldi, Jean-Luc Gouzé [BIOCORE project-team], Sofya Maslovskaya, Jean-Baptiste Pomet, Agustín Yabo.

The growth of microorganisms is fundamentally an optimization problem which consists in dynamically allocating resources to cellular functions so as to maximize growth rate or another fitness criterion. Simple ordinary differential equation models, called self-replicators, have been used to formulate this problem in the framework of optimal and feedback control theory, allowing observations in microbial physiology to be explained. The resulting control problems are very challenging due to the nonlinearity of the models, parameter uncertainty, the coexistence of different time-scales, a dynamically changing environment, and various other physical and chemical constraints. In the framework of the ANR Maximic (PI Hidde de Jong, Inria Grenoble Rhône-Alpes) we aim at developing novel theoretical approaches for addressing these challenges in order to (i) study natural resource allocation strategies in microorganisms and (ii) propose new synthetic control strategies for biotechnological applications. In order to address (i), we develop extended self-replicator models accounting for the cost of regulation and energy metabolism in bacterial cells. We study these models by a combination of analytical and numerical approaches to derive optimal control solutions and a control synthesis, dealing with the bang-bang-singular structure of the solutions. Moreover, we define quasi-optimal feedback control strategies inspired by known regulatory mechanisms in the cell. To test whether bacteria follow the predicted optimal strategies, we quantify dynamic resource allocation in the bacterium Escherichia coli by monitoring, by means of time-lapse fluorescent microscopy, the expression of selected genes in single cells growing in a microfluidics device. In order to address (ii), we build self-replicator models that include a pathway for the production of a metabolite of interest. We also add a mechanism to turn off microbial growth by means of an external input signal, at the profit of the production of the metabolite. We formulate the maximization of the amount of metabolite produced as an optimal control problem, and derive optimal solutions and a control synthesis, as well as quasi-optimal feedback strategies satisfying chemical and physical design constraints. The proposed synthetic control strategies are being tested experimentally by growing E. coli strains capable of producing glycerol from glucose in a mini-bioreactor system. We aim at quantifying the amount of glucose consumed and glycerol produced, in the case of a predefined input signal (open-loop control) and the adaptive regulation of the input signal based on on-line measurements of the growth rate and the expression of fluorescent reporters of selected genes (closed-loop control). Currently, one PhD (A. Yabo) and one postdoc (S. Maslovskaya) are involved in these tasks and jointly supervised by colleagues from McTAO and Biocore teams at Sophia. Preliminary results concern the definition on extended (higher dimensional) models for the bacteria dynamics, check of second order optimality conditions on the resulting optimal control problem, and study of the turnpike phenomenon for these optimization problems.

MORPHEME Project-Team (section vide)

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

4. Application Domains

4.1. Electromagnetic wave propagation

Electromagnetic devices are ubiquitous in present day technology. Indeed, electromagnetism has found and continues to find applications in a wide array of areas, encompassing both industrial and societal purposes. Applications of current interest include (among others) those related to communications (e.g transmission through optical fiber lines), to biomedical devices (e.g microwave imaging, micro-antenna design for telemedecine, etc.), to circuit or magnetic storage design (electromagnetic compatibility, hard disc operation), to geophysical prospecting, and to non-destructive evaluation (e.g crack detection), to name but just a few. Equally notable and motivating are applications in defence which include the design of military hardware with decreased signatures, automatic target recognition (e.g bunkers, mines and buried ordnance, etc.) propagation effects on communication and radar systems, etc. Although the principles of electromagnetics are well understood, their application to practical configurations of current interest, such as those that arise in connection with the examples above, is significantly complicated and far beyond manual calculation in all but the simplest cases. These complications typically arise from the geometrical characteristics of the propagation medium (irregular shapes, geometrical singularities), the physical characteristics of the propagation medium (heterogeneity, physical dispersion and dissipation) and the characteristics of the sources (wires, etc.).

Although many of the above-mentioned application contexts can potentially benefit from numerical modeling studies, the team currently concentrates its efforts on two physical situations.

4.1.1. *Microwave interaction with biological tissues*

Two main reasons motivate our commitment to consider this type of problem for the application of the numerical methodologies developed in the NACHOS project-team:

- First, from the numerical modeling point of view, the interaction between electromagnetic waves and biological tissues exhibit the three sources of complexity identified previously and are thus particularly challenging for pushing one step forward the state-of-the art of numerical methods for computational electromagnetics. The propagation media is strongly heterogeneous and the electromagnetic characteristics of the tissues are frequency dependent. Interfaces between tissues have rather complicated shapes that cannot be accurately discretized using cartesian meshes. Finally, the source of the signal often takes the form of a complicated device (e.g a mobile phone or an antenna array).

- Second, the study of the interaction between electromagnetic waves and living tissues is of interest to several applications of societal relevance such as the assessment of potential adverse effects of electromagnetic fields or the utilization of electromagnetic waves for therapeutic or diagnostic purposes. It is widely recognized nowadays that numerical modeling and computer simulation of electromagnetic wave propagation in biological tissues is a mandatory path for improving the scientific knowledge of the complex physical mechanisms that characterize these applications.

Despite the high complexity both in terms of heterogeneity and geometrical features of tissues, the great majority of numerical studies so far have been conducted using variants of the widely known FDTD method due to Yee [61]. In this method, the whole computational domain is discretized using a structured (cartesian) grid. Due to the possible straightforward implementation of the algorithm and the availability of computational power, FDTD is currently the leading method for numerical assessment of human exposure to electromagnetic waves. However, limitations are still seen, due to the rather difficult departure from the commonly used rectilinear grid and cell size limitations regarding very detailed structures of human tissues. In this context, the general objective of the contributions of the NACHOS project-team is to demonstrate the benefits of high order unstructured mesh based Maxwell solvers for a realistic numerical modeling of the interaction of electromagnetic waves and biological tissues with emphasis on applications related to numerical dosimetry.

Since the creation of the team, our works on this topic have mainly been focussed on the study of the exposure of humans to radiations from mobile phones or wireless communication systems (see Fig. 1). This activity has been conducted in close collaboration with the team of Joe Wiart at Orange Labs/Whist Laboratory (http://whist.institut-telecom.fr/en/index.html) (formerly, France Telecom Research & Development) in Issy-les-Moulineaux [8].
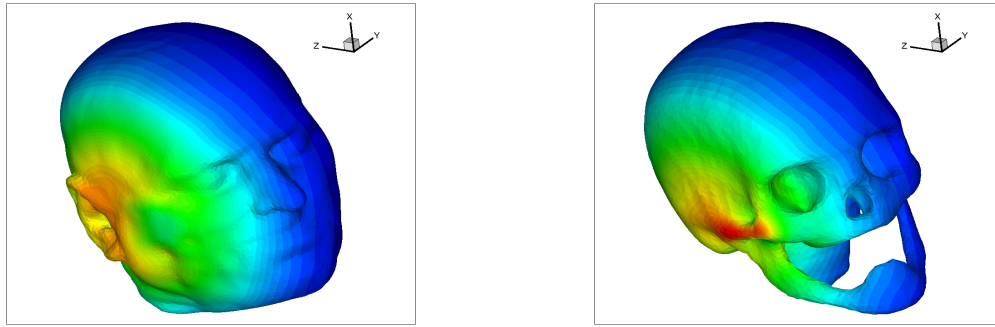

Figure 1. Exposure of head tissues to an electromagnetic wave emitted by a localized source. Top figures: surface triangulations of the skin and the skull. Bottom figures: contour lines of the amplitude of the electric field.

4.1.2. Light-matter interaction on the nanoscale

Nanostructuring of materials has opened up a number of new possibilities for manipulating and enhancing light-matter interactions, thereby improving fundamental device properties. Low-dimensional semiconductors, like quantum dots, enable one to catch the electrons and control the electronic properties of a material, while photonic crystal structures allow to synthesize the electromagnetic properties. These technologies may, e.g., be employed to make smaller and better lasers, sources that generate only one photon at a time, for applications in quantum information technology, or miniature sensors with high sensitivity. The incorporation of metallic structures into the medium add further possibilities for manipulating the propagation of electromagnetic waves. In particular, this allows subwavelength localisation of the electromagnetic field and, by subwavelength structuring of the material, novel effects like negative refraction, e.g. enabling super lenses, may be realized. Nanophotonics is the recently emerged, but already well defined, field of science and technology aimed at establishing and using the peculiar properties of light and light-matter interaction in various nanostructures. Nanophotonics includes all the phenomena that are used in optical sciences for the development of optical devices. Therefore, nanophotonics finds numerous applications such as in optical microscopy, the design of optical switches and electromagnetic chips circuits, transistor filaments, etc. Because of its numerous scientific and technological applications (e.g. in relation to telecommunication, energy production and biomedicine), nanophotonics represents an active field of research increasingly relying on numerical modeling beside experimental studies.

Plasmonics is a related field to nanophotonics. Metallic nanostructures whose optical scattering is dominated by the response of the conduction electrons are considered as plasmomic media. If the structure presents an interface with e.g. a dielectric with a positive permittivity, collective oscillations of surface electrons create surface-plasmons-polaritons (SPPs) that propagate along the interface. SPPs are guided along metal-dielectric interfaces much in the same way light can be guided by an optical fiber, with the unique characteristic of subwavelength-scale confinement perpendicular to the interface. Nanofabricated systems that exploit SPPs offer fascinating opportunities for crafting and controlling the propagation of light in matter. In particular, SPPs can be used to channel light efficiently into nanometer-scale volumes, leading to direct modification of mode dispersion properties (substantially shrinking the wavelength of light and the speed of light pulses for example), as well as huge field enhancements suitable for enabling strong interactions with non-linear

materials. The resulting enhanced sensitivity of light to external parameters (for example, an applied electric field or the dielectric constant of an adsorbed molecular layer) shows great promise for applications in sensing and switching. In particular, very promising applications are foreseen in the medical domain [53]- [62].

Numerical modeling of electromagnetic wave propagation in interaction with metallic nanostructures at optical frequencies requires to solve the system of Maxwell equations coupled to appropriate models of physical dispersion in the metal, such as the Drude and Drude-Lorentz models. Here again, the FDTD method is a widely used approach for solving the resulting system of PDEs [58]. However, for nanophotonic applications, the space and time scales, in addition to the geometrical characteristics of the considered nanostructures (or structured layouts of the latter), are particularly challenging for an accurate and efficient application of the FDTD method. Recently, unstructured mesh based methods have been developed and have demonstrated their potentialities for being considered as viable alternatives to the FDTD method [56]- [57]- [51]. Since the end of 2012, nanophotonics/plasmonics is increasingly becoming a focused application domain in the research activities of the team in close collaboration with physicists from CNRS laboratories, and also with researchers from international institutions.
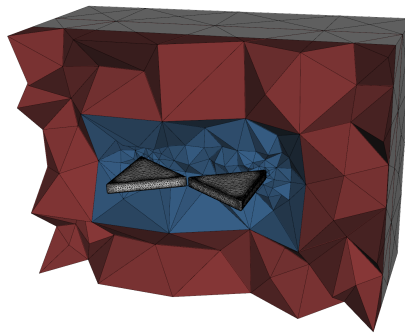

Figure 2. Simulation of the field enhancement at the tip of a gold bowtie nanoantenna (PhD thesis of Jonathan Viquerat).

4.2. Elastodynamic wave propagation

Elastic wave propagation in interaction with solids are encountered in a lot of scientific and engineering contexts. One typical example is geoseismic wave propagation for earthquake dynamics or resource prospection.

4.2.1. *Earthquake dynamics*

To understand the basic science of earthquakes and to help engineers better prepare for such an event, scientists want to identify which regions are likely to experience the most intense shaking, particularly in populated sediment-filled basins. This understanding can be used to improve buildings in high hazard areas and to help engineers design safer structures, potentially saving lives and property. In the absence of deterministic earthquake prediction, forecasting of earthquake ground motion based on simulation of scenarios is one of the most promising tools to mitigate earthquake related hazard. This requires intense modeling that meets the spatial and temporal resolution scales of the continuously increasing density and resolution of the seismic instrumentation, which record dynamic shaking at the surface, as well as of the basin models. Another important issue is to improve the physical understanding of the earthquake rupture processes and seismic wave propagation. Large-scale simulations of earthquake rupture dynamics and wave propagation are currently the only means to investigate these multiscale physics together with data assimilation and inversion. High resolution models are also required to develop and assess fast operational analysis tools for real time seismology and early warning systems.

Numerical methods for the propagation of seismic waves have been studied for many years. Most of existing numerical software rely on finite difference type methods. Among the most popular schemes, one can cite the staggered grid finite difference scheme proposed by Virieux [59] and based on the first order velocity-stress hyperbolic system of elastic waves equations, which is an extension of the scheme derived by Yee [61] for the solution of the Maxwell equations. Many improvements of this method have been proposed, in particular, higher order schemes in space or rotated staggered-grids allowing strong fluctuations of the elastic parameters. Despite these improvements, the use of cartesian grids is a limitation for such numerical methods especially when it is necessary to incorporate surface topography or curved interface. Moreover, in presence of a non planar topography, the free surface condition needs very fine grids (about 60 points by minimal Rayleigh wavelength) to be approximated. In this context, our objective is to develop high order unstructured mesh based methods for the numerical solution of the system of elastodynamic equations for elastic media in a first step, and then to extend these methods to a more accurate treatment of the heterogeneities of the medium or to more complex propagation materials such as viscoelastic media which take into account the intrinsic attenuation. Initially, the team has considered in detail the necessary methodological developments for the large-scale simulation of earthquake dynamics [1]. More recently, the team has collaborated with CETE Méditerranée which is a regional technical and engineering centre whose activities are concerned with seismic hazard assessment studies, and IFSTTAR (https://www.ifsttar.fr/en/welcome/) which is the French institute of science and technology for transport, development and networks, conducting research studies on control over aging, risks and nuisances.
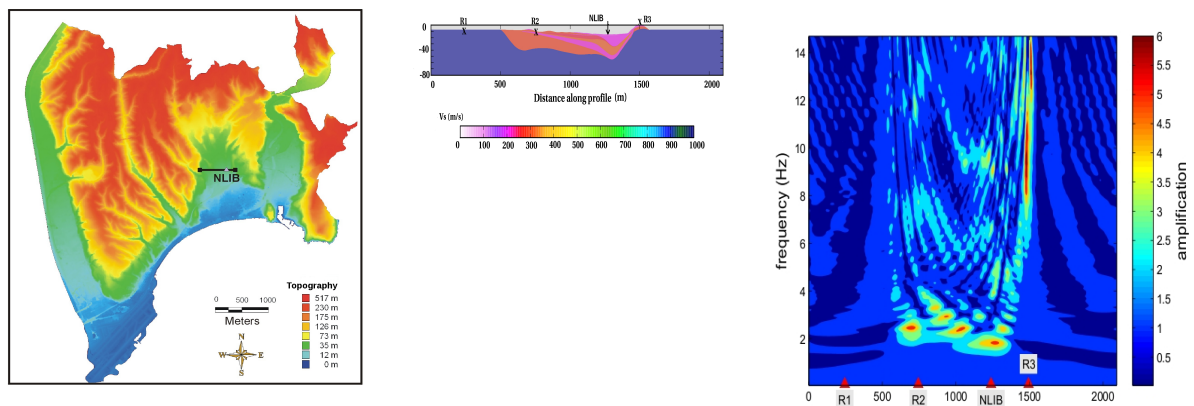

Figure 3. Propagation of a plane wave in a heterogeneous model of Nice area (provided by CETE Méditerranée). Left figure: topography of Nice and location of the cross-section used for numerical simulations (black line). Middle figure: S-wave velocity distribution along the cross-section in the Nice basin. Right figure: transfer functions (amplification) for a vertically incident plane wave ; receivers every 5 m at the surface. This numerical simulation was performed using a numerical method for the solution of the elastodynamics equations coupled to a Generalized Maxwell Body (GMB) model of viscoelasticity (PhD thesis of Fabien Peyrusse).

4.2.2. Seismic exploration

This application topic is considered in close collaboration with the MAGIQUE-3D project-team at Inria Bordeaux - Sud-Ouest which is coordinating the Depth Imaging Partnership (DIP -http://dip.inria.fr) between Inria and TOTAL. The research program of DIP includes different aspects of the modeling and numerical simulation of sesimic wave propagation that must be considered to construct an efficient software suites for producing accurate images of the subsurface. Our common objective with the MAGIQUE-3D project-team is to design high order unstructured mesh based methods for the numerical solution of the system

of elastodynamic equations in the time-domain and in the frequency-domain, that will be used as forward modelers in appropriate inversion procedures.

4. Application Domains

4.1. Network Science

Network Science is a multidisciplinary body of knowledge, principally concerned with the emergence of global properties in a network of individual agents, from the "local" properties of this network, namely, the way agents interact with each other. The central model of "networks" is the graph (of Graph Theory/Operations Research), with nodes representing the different entities managing information and taking decisions, and the links representing the fact that entities interact, or not. Links are usually equipped with a "weight" that measures the intensity of interaction. Adding evolution rules to this quite elementary representation leads to dynamic network models, the properties of which Network Science tries to analyze.

A classical example of properties sought in networks is the famous "six degrees of separation" (or "small world") property: how and why does it happen so frequently? Another ubiquitous property of real-life networks is the Zipf or "scale-free" distribution for degrees. Some of these properties, when properly exploited, lead to successful business opportunities: just consider the PageRank algorithm of Google, which miraculously connects the relevance of some Web information with the relevance of the other information that points to it.

4.2. Network Engineering

In its primary acceptation, Network Science involves little or no engineering: phenomena are assumed to be "natural" and emerge without intervention. However, the idea comes fast to intervene in order to modify the outcome of the phenomenon. This is where NEO is positioned. Beyond the mostly descriptive approach of Network Science, we aim at using the techniques of Operations Research so as to engineer complex networks.

To quote just two examples: controlling the spread of diseases through a "network" of people is of primarily interest for mankind. Similarly, controlling the spread of information or reputation through a social network is of great interest in the Internet. Precisely: given the impact of web visibility on business income, it is tempting (and quite common) to manipulate the graph of the web by adding links so as to drive the PageRank algorithm to a desired outcome.

Another interesting example is the engineering of community structures. Recently, thousands of papers have been written on the topic of community *detection* problem. In most of the works, the researchers propose methods, most of the time, heuristics, for detecting communities or dense subgraphs inside a large network. Much less effort has been put in the understanding of community formation process and even much less effort has been dedicated to the question of how one can influence the process of community formation, e.g. in order to increase overlap among communities and reverse the fragmentation of the society.

Our ambition for the medium term is to reach an understanding of the behavior of complex networks that will make us capable of influencing or producing a certain property in said network. For this purpose, we will develop families of models to capture the essential structure, dynamics, and uncertainty of complex networks. The "solution" of these models will provide the correspondence between metrics of interest and model parameters, thus opening the way to the synthesis of effective control techniques.

In the process of tackling real, very large size networks, we increasingly deal with large graph data analysis and the development of decision techniques with low algorithmic complexity, apt at providing answers from large datasets in reasonable time.

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

3. Application Domains

3.1. Domain 1: Privacy in smart environments

Privacy in smart environments. One illustrative example is our latest work on privacy-preserving smart-metering [2]. Several countries throughout the world are planning to deploy smart meters in house-holds in the very near future. Traditional electrical meters only measure total consumption on a given period of time (i.e., one month or one year). As such, they do not provide accurate information of when the energy was consumed. Smart meters, instead, monitor and report consumption in intervals of few minutes. They allow the utility provider to monitor, almost in real-time, consumption and possibly adjust generation and prices according to the demand. Billing customers by how much is consumed and at what time of day will probably change consumption habits to help matching energy consumption with production. In the longer term, with the advent of smart appliances, it is expected that the smart grid will remotely control selected appliances to reduce demand. Although smart metering might help improving energy management, it creates many new privacy problems. Smart-meters provide very accurate consumption data to electricity providers. As the interval of data collected by smart meters decreases, the ability to disaggregate low-resolution data increases. Analysing high-resolution consumption data, Non-intrusive Appliance Load Monitoring (NALM) can be used to identify a remarkable number of electric appliances (e.g., water heaters, well pumps, furnace blowers, refrigerators, and air conditioners) employing exhaustive appliance signature libraries. We developed DREAM, DiffeRentially privatE smArt Metering, a scheme that is private under the differential privacy model and therefore provides strong and provable guarantees. With our scheme, an (electricity) supplier can periodically collect data from smart-meters and derive aggregated statistics while learning only limited information about the activities of individual households. For example, a supplier cannot tell from a user's trace when he watched TV or turned on heating.

3.2. Domain 2: Big Data and Privacy

We believe that another important problem will be related to privacy issues in big data. Public datasets are used in a variety of applications spanning from genome and web usage analysis to location-based and recommendation systems. Publishing such datasets is important since they can help us analyzing and understanding interesting patterns. For example, mobility trajectories have become widely collected in recent years and have opened the possibility to improve our understanding of large-scale social networks by investigating how people exchange information, interact, and develop social interactions. With billion of handsets in use worldwide, the quantity of mobility data is gigantic. When aggregated, they can help understand complex processes, such as the spread of viruses, and build better transportation systems. While the benefits provided by these datasets are indisputable, they unfortunately pose a considerable threat to individual privacy. In fact, mobility trajectories might be used by a malicious attacker to discover potential sensitive information about a user, such as his habits, religion or relationships. Because privacy is so important to people, companies and researchers are reluctant to publish datasets by fear of being held responsible for potential privacy breaches. As a result, only very few of them are actually released and available. This limits our ability to analyze such data to derive information that could benefit the general public. It is now an urgent need to develop Privacy-Preserving Data Analytics (PPDA) systems that collect and transform raw data into a version that is immunized against privacy attacks but that still preserves useful information for data analysis. This is one of the objectives of Privatics. There exists two classes of PPDA according to whether the entity that is collecting and anonymizing the data is trusted or not. In the trusted model, that we refer to as Privacy-Preserving Data Publishing (PPDP), individuals trust the publisher to which they disclose their data. In the untrusted model, that we refer to as Privacy-Preserving Data Collection (PPDC), individuals do not trust the data publisher. They may add some noise to their data to protect sensitive information from the data publisher.

Privacy-Preserving Data Publishing: In the trusted model, individuals trust the data publisher and disclose all their data to it. For example, in a medical scenario, patients give their true information to hospitals to receive proper treatment. It is then the responsibility of the data publisher to protect privacy of the individuals' personal data. To prevent potential data leakage, datasets must be sanitized before possible release. Several proposals have been recently proposed to release private data under the Differential Privacy model [25, 56, 26, 57, 50]. However most of these schemes release a "snapshot" of the datasets at a given period of time. This release often consists of histograms. They can, for example, show the distributions of some pathologies (such as cancer, flu, HIV, hepatitis, etc.) in a given population. For many analytics applications, "snapshots" of data are not enough, and sequential data are required. Furthermore, current work focusses on rather simple data structures, such as numerical data. Release of more complex data, such as graphs, are often also very useful. For example, recommendation systems need the sequences of visited websites or bought items. They also need to analyse people connection graphs to identify the best products to recommend. Network trace analytics also rely on sequences of events to detect anomalies or intrusions. Similarly, traffic analytics applications typically need sequences of visited places of each user. In fact, it is often essential for these applications to know that user A moved from position 1 to position 2, or at least to learn the probability of a move from position 1 to position 2. Histograms would typically represent the number of users in position 1 and position 2, but would not provide the number of users that moved from position 1 to position 2. Due to the inherent sequentiality and high-dimensionality of sequential data, one major challenge of applying current data sanitization solutions on sequential data comes from the uniqueness of sequences (e.g., very few sequences are identical). This fact makes existing techniques result in poor utility. Schemes to privately release data with complex data structures, such as sequential, relational and graph data, are required. This is one the goals of Privatics. In our current work, we address this challenge by employing a variable-length n-gram model, which extracts the essential information of a sequential database in terms of a set of variable-length n - grams [15]. We then intend to extend this approach to more complex data structures.

Privacy-Preserving Data Collection: In the untrusted model, individuals do not trust their data publisher. For example, websites commonly use third party web analytics services, such as Google Analytics to obtain aggregate traffic statistics such as most visited pages, visitors' countries, etc. Similarly, other applications, such as smart metering or targeted advertising applications, are also tracking users in order to derive aggregated information about a particular class of users. Unfortunately, to obtain this aggregate information, services need to track users, resulting in a violation of user privacy. One of our goals is to develop Privacy-Preserving Data Collection solutions. We propose to study whether it is possible to provide efficient collection/aggregation solutions without tracking users, i.e. without getting or learning individual contributions.

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

4. Application Domains

4.1. Mathematical Components

The Mathematical Components is the main by-product of an effort started almost two decades ago to provide a formally verified proof for a major theorem in group theory. Because this major theorem had a proof published in books of several hundreds of pages, with elements coming from character theory, other coming from algebra, and some coming from real analysis, it was an exercice in building a large library, with results in many domains, and in establishing clear guidelines for further increase and data search.

This library has proved to be a useful repository of mathematical facts for a wide area of applications, so that it has a growing community of users in many contries (Denmark, France, Germany, Japan, Singapore, Spain, Sweden, UK, USA, at the time of writing these lines in 2019) and for a wide variety of topics (transcendental number theory, elliptic curve cryptography, articulated robot kinematics, recently block chain foundations).

Interesting questions on this library range around the importance of decidability and proof irrelevance, the way to structure knowledge to automatically inherit theorems from one topic to another, the way to generate infrastructure to make this automation efficient and predictable. In particular, we want to concentrate on adding a new mathematical topic to this library: real analysis and then complex analysis (Mathematical Components Analysis).

On the front of automation, we are convinced that a higher level language is required to describe similarities between theories, to generate theorems that are immediate consequences of structures, etc, and for this reason, we invest in the development of a new language on top of the proof assistant (ELPI).

4.2. Proofs in cryptography

When we work on cryptography, we are interested in the formal verification of proofs showing that some cryptographic primitives provide good guarantees against unwanted access to information. Over the years we have developed a technique for this kind of reasoning that relies on a programing logic (close to Hoare logic) with probabilistic apsects and the capability to establish relations between several implementations of a problem. The resulting programming logic is called *probabilistic relational Hoare logic*. In more recent work, we have also started to study questions of *side-channel* attacks, where we wish to guarantee that opponents cannot gain access to protected knowledge, even if they observe specific features of execution, like execution time (to which the answer lies in *constant-time* execution) or partial access to memory bits (to which the answer lies in *masking*).

For this domain of application, we choose to work with a specific proof tool (EasyCrypt), which combines powerful first-order reasoning and uses of automatic tools, with a specific support for probabilistic relational Hoare Logic. The development of this EasyCrypt proof tool is one of the objectives of our team.

When it comes to formal proofs of resistance to side-channel attack, we contend that it is necessary to verify formally that the compiler used in the production of actually running code respects the resistance properties that were established in formally verified proofs. One of our objectives is to describe such a compiler (Jasmin) and show its strength on a variety of applications.

4.3. Proofs for robotics

Robots are man-made artifacts where numerous design decisions can be argued based on logical or mathematical principles. For this reason, we wish to use this domain of application as a focus for our investigations. The questions for which we are close to providing answers involve precision issues in numeric computation, obstacle avoidance and motion planning (including questions of graph theory), articulated limb cinematics and dynamics, and balance and active control.

From the mathematical perspective, these topics require that we improve our library to cover real algebraic geometry, computational geometry, real analysis, graph theory, and refinement relations between abstract algorithms and executable programs.

In the long run, we hope to exhibit robots where pieces of software and part of the design has been subject to formal verification.

Stars Project-Team (section vide)

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

4. Application Domains

4.1. Domain 1

In addition to tackling enduring scientific challenges, our research on geometric modeling and processingis motivated by applications to computational engineering, reverse engineering, digital mapping and urbanplanning. The main deliverable of our research will be algorithms with theoretical foundations. Ultimatelywe wish to contribute making geometry modeling and processing routine for practitioners who deal with real-world data. Our contributions may also be used as a sound basis for future software and technology developments.

Our first ambition for technology transfer is to consolidate the components of our research experiments in the form of new software components for the CGAL (Computational Geometry Algorithms Library) library. Consolidation being best achieved with the help of an engineer, we will search for additional funding. Through CGAL, we wish to contribute to the "standard geometric toolbox", so as to provide a generic answer to application needs instead of fragmenting our contributions. We already cooperate with the Inria spin-offcompany Geometry Factory, which commercializes CGAL, maintains it and provide technical support.

Our second ambition is to increase the research momentum of companies through advising Cifre Ph.D. thesesand postdoctoral fellows on topics that match our research program.

TOSCA Team

4. Application Domains

4.1. Application domains

TOSCA is interested in developing stochastic models and probabilistic numerical methods. Our present motivations come from models with singular coefficients, with applications in Geophysics, Molecular Dynamics and Neurosciences; Lagrangian modeling in Fluid Dynamics and Meteorology; Population Dynamics, Evolution and Genetics; Neurosciences; and Financial Mathematics.

4.1.1. *Stochastic models with singular coefficients: Analysis and simulation*

Stochastic differential equations with discontinuous coefficients arise in Geophysics, Chemistry, Molecular Dynamics, Neurosciences, Oceanography, etc. In particular, they model changes of diffusion of fluids, or diffractions of particles, along interfaces.

For practiners in these fields, Monte Carlo methods are popular as they are easy to interpret — one follows particles — and are in general easy to set up. However, dealing with discontinuities presents many numerical and theoretical challenges. Despite its important applications, ranging from brain imaging to reservoir simulation, very few teams in mathematics worldwide are currently working in this area. The Tosca project-team has tackled related problems for several years providing rigorous approach. Based on stochastic analysis as well as interacting with researchers in other fields, we developed new theoretical and numerical approaches for extreme cases such as Markov processes whose generators are of divergence form with discontinuous diffusion coefficient.

The numerical approximation of singular stochastic processes can be combined with backward stochastic differential equations (BSDEs) or branching diffusions to obtain Monte Carlo methods for quasi-linear PDEs with discontinuous coefficients. The theory of BSDEs has been extensively developed since the 1980s, but the general assumptions for their existence can be quite restrictive. Although the probabilistic interpretation of quasi-linear PDEs with branching diffusions has been known for a long time, there have been only a few works on the related numerical methods.

Another motivation to consider stochastic dynamics in a discontinuous setting came to us from time evolution of fragmentation and coagulation phenomena, with the objective to elaborate stochastic models for the avalanche formation of soils, snow, granular materials or other geomaterials. Most of the models and numerical methods for avalanches are deterministic and involve a wide variety of physical parameters such as the density of the snow, the yield, the friction coefficient, the pressure, the basal topography, etc. One of these methods consists in studying the safety factor (or limit load) problem, related to the shallow flow of a visco-plastic fluid/solid with heterogeneous thickness over complex basal topography. The resulting nonlinear partial differential equation of this last theory involves many singularities, which motivates us to develop an alternative stochastic approach based on our past works on coagulation and fragmentation. Our approach consists in studying the evolution of the size of a typical particle in a particle system which fragments in time.

4.1.2. *Stochastic Lagrangian modeling in Computational Fluid Dynamics*

Stochastic Lagrangian models were introduced in the eighties to simulate complex turbulent flows, particularly two-phase flows. In Computational Fluid Dynamics (CFD), they are intensively used in the so-called Probability Density Functions (PDF) methods in order to model and compute the reaction-phase terms in the fundamental equations of fluid motions. The PDF methods are currently developed in various laboratories by specialists in scientific computation and physicists. However, to our knowledge, we are innovating in two ways:

- our theoretical studies are the pioneering mathematical analysis of Lagrangian stochastic models in CFD;
- our work on the Stochastic Downscaling Method (SDM) for wind simulation is the first attempt to solve the fundamental equations themselves by a fully 3D stochastic particle method.

We emphasize that our numerical analysis is essential to the SDM development which takes benefits from our deep expertise on numerical schemes for McKean-Vlasov-non-linear SDEs.

4.1.3. Population Dynamics, Evolution and Genetics

The activity of the team on stochastic modeling in population dynamics and genetics mainly concerns application in adaptive dynamics, a branch of evolutionary biology studying the interplay between ecology and evolution, ecological modeling, population genetics in growing populations, and stochastic control of population dynamics, with applications to cancer growth modeling. Stochastic modeling in these areas mainly considers individual-based models, where the birth and death of each individual is described. This class of model is well-developed in Biology, but their mathematical analysis is still fragmentary. Another important topic in population dynamics is the study of populations conditioned to non-extinction, and of the corresponding stationary distributions, called quasi-stationary distributions (QSD). This domain has been the object of a lot of studies since the 1960's, but we made recently significant progresses on the questions of existence, convergence and numerical approximation of QSDs using probabilistic tools rather than the usual spectral tools.

Our activity in population dynamics also involves a fully new research project on cancer modeling at the cellular level by means of branching processes. In 2010 the International Society for Protons Dynamics in Cancer was launched in order to create a critical mass of scientists engaged in research activities on Proton Dynamics in Cancer, leading to the facilitation of international collaboration and translation of research to clinical development. Actually, a new branch of research on cancer evolution is developing intensively; it aims in particular to understand the role of proteins acting on cancerous cells' acidity, their effects on glycolysis and hypoxia, and the benefits one can expect from controlling pH regulators in view of proposing new therapies.

4.1.4. Stochastic modeling in Neuroscience

It is generally accepted that many different neural processes that take place in the brain involve noise. Indeed, one typically observes experimentally underlying variability in the spiking times of an individual neuron in response to an unchanging stimulus, while a predictable overall picture emerges if one instead looks at the average spiking time over a whole group of neurons. Sources of noise that are of interest include ionic currents crossing the neural membrane, synaptic noise, and the global effect of the external environment (such as other parts of the brain).

It is likely that these stochastic components play an important role in the function of both the neurons and the networks they form. The characterization of the noise in the brain, its consequences at a functional level and its role at both a microscopic (individual neuron) level and macroscopic level (network of thousands of neurons) is therefore an important step towards understanding the nervous system.

To this end, a large amount of current research in the neuroscientific literature has involved the addition of noise to classical purely deterministic equations resulting in new phenomena being observed. The aim of the project is thus to rigorously study these new equations in order to be able to shed more light on the systems they describe.

4.1.5. Stochastic modeling in Financial Mathematics

4.1.5.1. Technical Analysis

In the financial industry, there are three main approaches to investment: the fundamental approach, where strategies are based on fundamental economic principles; the technical analysis approach, where strategies are based on past price behavior; and the mathematical approach where strategies are based on mathematical models and studies. The main advantage of technical analysis is that it avoids model specification, and thus calibration problems, misspecification risks, etc. On the other hand, technical analysis techniques have limited theoretical justifications, and therefore no one can assert that they are risk-less, or even efficient.

4.1.5.2. Financial Risks Estimation and Hedging

Popular models in financial mathematics usually assume that markets are perfectly liquid. In particular, each trader can buy or sell the amount of assets he/she wants at the same price (the "market price"). They moreover assume that the decision taken by the trader does not affect the price of the asset (the small investor assumption). In practice, the assumption of perfect liquidity is never satisfied but the error due to liquidity is generally negligible with respect to other sources of error such as model error or calibration error, etc.

Derivatives of interest rates are singular for at least two reasons: firstly the underlying (interest rate) is not directly exchangeable, and secondly the liquidity costs usually used to hedge interest rate derivatives have large variation in times.

Due to recurrent crises, the problem of risk estimation is now a crucial issue in finance. Regulations have been enforced (Basel Committee II). Most asset management software products on the markets merely provide basic measures (VaR, Tracking error, volatility) and basic risk explanation features (e.g., "top contributors" to risk, sector analysis, etc).

4.1.5.3. Energy and Carbon Markets

With the rise of renewable energy generation (from solar, wind, waves...), engineers face new challenges which heavily rely on stochastic and statistical problems.

Besides, in the context of the beginning of the second phase (the Kyoto phase) in 2008 of the European carbon market, together with the fact that French carbon tax was scheduled to come into law on Jan. 1, 2010, the year 2009 was a key year for the carbon price modeling. Our research approach adopts the point of view of the legislator and energy producers. We used both financial mathematical tools and a game theory approach. Today, with the third phase of the EU-ETS, that didn't yet start, and the report form the Cour des Comptes (October 2013) that pointed out (among many others point) the lack of mathematical modeling on such carbon market design, we continue our research in this direction.

4.1.5.4. Optimal Stopping Problems

The theory of optimal stopping is concerned with the problem of taking a decision at the best time, in order to maximise an expected reward (or minimise an expected cost). We work on the general problem of optimal stopping with random discounting and additional cost of observation.

4.1.5.5. First hitting times distributions

Diffusion hitting times are of great interest in finance (a typical example is the study of barrier options) and also in Geophysics and Neurosciences. On the one hand, analytic expressions for hitting time densities are well known and studied only in some very particular situations (essentially in Brownian contexts). On the other hand, the study of the approximation of the hitting times for stochastic differential equtions is an active area of research since very few results still are available in the literature.

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

4. Application Domains

4.1. Social Semantic Web

A number of evolutions have changed the face of information systems in the past decade but the advent of the Web is unquestionably a major one and it is here to stay. From an initial wide-spread perception of a public documentary system, the Web as an object turned into a social virtual space and, as a technology, grew as an application design paradigm (services, data formats, query languages, scripting, interfaces, reasoning, etc.). The universal deployment and support of its standards led the Web to take over nearly all of our information systems. As the Web continues to evolve, our information systems are evolving with it.

Today in organizations, not only almost every internal information system is a Web application, but these applications more and more often interact with external Web applications. The complexity and coupling of these Web-based information systems call for specification methods and engineering tools. From capturing the needs of users to deploying a usable solution, there are many steps involving computer science specialists and non-specialists.

We defend the idea of relying on Semantic Web formalisms to capture and reason on the models of these information systems supporting the design, evolution, interoperability and reuse of the models and their data as well as the workflows and the processing.

4.2. Linked Data on the Web and on Intranets

With billions of triples online (see Linked Open Data initiative), the Semantic Web is providing and linking open data at a growing pace and publishing and interlinking the semantics of their schemas. Information systems can now tap into and contribute to this Web of data, pulling and integrating data on demand. Many organisations also started to use this approach on their intranets leading to what is called linked enterprise data.

A first application domain for us is the publication and linking of data and their schemas through Web architectures. Our results provide software platforms to publish and query data and their schemas, to enrich these data in particular by reasoning on their schemas, to control their access and licenses, to assist the workflows that exploit them, to support the use of distributed datasets, to assist the browsing and visualization of data, etc.

Examples of collaboration and applied projects include: SMILK Joint Laboratory, Corese, DBpedia.fr.

4.3. Assisting Web-based Epistemic Communities

In parallel with linked open data on the Web, social Web applications also spread virally (e.g. Facebook growing toward 1.5 billion users) first giving the Web back its status of a social read-write media and then putting it back on track to its full potential of a virtual place where to act, react and interact. In addition, many organizations are now considering deploying social Web applications internally to foster community building, expert cartography, business intelligence, technological watch and knowledge sharing in general.

By reasoning on the Linked Data and the semantics of the schemas used to represent social structures and Web resources, we provide applications supporting communities of practice and interest and fostering their interactions in many different contexts (e-learning, business intelligence, technical watch, etc.).

We use typed graphs to capture and mix: social networks with the kinds of relationships and the descriptions of the persons; compositions of Web services with types of inputs and outputs; links between documents with their genre and topics; hierarchies of classes, thesauri, ontologies and folksonomies; recorded traces and suggested navigation courses; submitted queries and detected frequent patterns; timelines and workflows; etc.

Our results assist epistemic communities in their daily activities such as biologists exchanging results, business intelligence and technological watch networks informing companies, engineers interacting on a project, conference attendees, students following the same course, tourists visiting a region, mobile experts on the field, etc. Examples of collaboration and applied projects: EduMICS, OCKTOPUS, Vigiglobe, Educlever, Gayatech.

4.4. Linked Data for a Web of Diversity

We intend to build on our results on explanations (provenance, traceability, justifications) and to continue our work on opinions and arguments mining toward the global analysis of controversies and online debates. One result would be to provide new search results encompassing the diversity of viewpoints and providing indicators supporting opinion and decision making and ultimately a Web of trust. Trust indicators may require collaborations with teams specialized in data certification, cryptography, signature, security services and protocols, etc. This will raise the specific problem of interaction design for security and privacy. In addition, from the point of view of the content, this requires to foster the publication and coexistence of heterogeneous data with different points of views and conceptualizations of the world. We intend to pursue the extension of formalisms to allow different representations of the world to co-exist and be linked and we will pay special attention to the cultural domain and the digital humanities. Examples of collaboration and applied projects: Zoomathia, Seempad, SMILK,

4.5. Artificial Web Intelligence

We intend to build on our experience in artificial intelligence (knowledge representation, reasoning) and distributed artificial intelligence (multi-agent systems - MAS) to enrich formalisms and propose alternative types of reasoning (graph-based operations, reasoning with uncertainty, inductive reasoning, non-monotonic, etc.) and alternative architectures for linked data with adequate changes and extensions required by the open nature of the Web. There is a clear renewed interest in AI for the Web in general and for Web intelligence in particular. Moreover distributed AI and MAS provide both new architectures and new simulation platforms for the Web. At the macro level, the evolution accelerated with HTML5 toward Web pages as full applications and direct Page2Page communication between browser clearly is a new area for MAS and P2P architectures. Interesting scenarios include the support of a strong decentralization of the Web and its resilience to degraded technical conditions (downscaling the Web), allowing pages to connect in a decentralized way, forming a neutral space, and possibly going offline and online again in erratic ways. At the micro level, one can imagine the place RDF and SPARQL could take as data model and programming model in the virtual machines of these new Web pages and, of course, in the Web servers. RDF is also used to serialize and encapsulate other languages and becomes a pivot language in linking very different applications and aspects of applications. Example of collaboration and applied projects: MoreWAIS, Corese, Vigiglobe collaboration.

4.6. Human-Data Interaction (HDI) on the Web

We need more interaction design tools and methods for linked data access and contribution. We intend to extend our work on exploratory search coupling it with visual analytics to assist sense making. It could be a continuation of the Gephi extension that we built targeting more support for non experts to access and analyze data on a topic or an issue of their choice. More generally speaking SPARQL is inappropriate for common users and we need to support a larger variety of interaction means with linked data. We also believe linked data and natural language processing (NLP) have to be strongly integrated to support natural language based interactions. Linked Open Data (LOD) for NLP, NLP for LOD and Natural Dialog Processing for querying, extracting and asserting data on the Web is a priority to democratize its use. Micro accesses and micro contributions are important to ensure public participation and also call for customized interfaces and thus for methods and tools to generate these interfaces. In addition, the user profiles are being enriched now with new data about the user such as her current mental and physical state, the emotion she just expressed or her cognitive performances. Taking into account this information to improve the interactions, change the behavior of the system and adapt the interface is a promising direction. And these human-data interaction

means should also be available for "small data", helping the user to manage her personal information and to link it to public or collective one, maintaining her personal and private perspective as a personal Web of data. Finally, the continuous knowledge extractions, updates and flows add the additional problem of representing, storing, querying and interacting with dynamic data. Examples of collaboration and applied projects: QAKIS, Sychonext collaboration, ALOOF, DiscoveryHub, WASABI, MoreWAIS.

Web-augmented interactions with the world: The Web continues to augment our perception and interaction with reality. In particular, Linked Open Data enable new augmented reality applications by providing data sources on almost any topic. The current enthusiasm for the Web of Things, where every object has a corresponding Web resource, requires evolutions of our vision and use of the Web architecture. This vision requires new techniques as the ones mentioned above to support local search and contextual access to local resources but also new methods and tools to design Web-based human devices interactions, accessibility, etc. These new usages are placing new requirements on the Web Architecture in general and on the semantic Web models and algorithms in particular to handle new types of linked data. They should support implicit requests considering the user context as a permanent query. They should also simplify our interactions with devices around us jointly using our personal preferences and public common knowledge to focus the interaction on the vital minimum that cannot be derived in another way. For instance the access to the Web of data for a robot can completely change the quality of the interactions it can offer. Again, these interactions and the data they require raise problems of security and privacy. Examples of collaboration and applied projects: ALOOF, AZKAR, MoreWAIS.

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

4. Application Domains

4.1. Data-intensive Scientific Applications

The application domains covered by Zenith are very wide and diverse, as they concern data-intensive scientific applications, i.e., most scientific applications. Since the interaction with scientists is crucial to identify and tackle data management problems, we are dealing primarily with application domains for which Montpellier has an excellent track record, i.e., agronomy, environmental science, life science, with scientific partners like INRA, IRD and CIRAD. However, we are also addressing other scientific domains (e.g. astronomy, oil extraction, music processing) through our international collaborations.

Let us briefly illustrate some representative examples of scientific applications on which we have been working on.

- **Management of astronomical catalogs**. An example of data-intensive scientific applications is the management of astronomical catalogs generated by the Dark Energy Survey (DES) project on which we are collaborating with researchers from Brazil. In this project, huge tables with billions of tuples and hundreds of attributes (corresponding to dimensions, mainly double precision real numbers) store the collected sky data. Data are appended to the catalog database as new observations are performed and the resulting database size has hundres of TB. Scientists around the globe can query the database with queries that may contain a considerable number of attributes. The volume of data that this application holds poses important challenges for data management. In particular, efficient solutions are needed to partition and distribute the data in several servers. An efficient partitioning scheme should try to minimize the number of fragments accessed in the execution of a query, thus reducing the overhead associated to handle the distributed execution.

- **Personal health data analysis and privacy** Today, it is possible to acquire data on many domains related to personal data. For instance, one can collect data on her daily activities, habits or health. It is also possible to measure performance in sports. This can be done thanks to sensors, communicating devices or even connected glasses. Such data, once acquired, can lead to valuable knowledge for these domains. For people having a specific disease, it might be important to know if they belong to a specific category that needs particular care. For an individual, it can be interesting to find a category that corresponds to her performances in a specific sport and then adapt her training with an adequate program. Meanwhile, for privacy reasons, people will be reluctant to share their personal data and make them public. Therefore, it is important to provide them with solutions that can extract such knowledge from everybody's data, while guaranteeing that their private data won't be disclosed to anyone.

- **Botanical data sharing**. Botanical data is highly decentralized and heterogeneous. Each actor has its own expertise domain, hosts its own data, and describes them in a specific format. Furthermore, botanical data is complex. A single plant's observation might include many structured and unstructured tags, several images of different organs, some empirical measurements and a few other contextual data (time, location, author, etc.). A noticeable consequence is that simply identifying plant species is often a very difficult task; even for the botanists themselves (the so-called taxonomic gap). Botanical data sharing should thus speed up the integration of raw observation data, while providing users an easy and efficient access to integrated data. This requires to deal with social-based data integration and sharing, massive data analysis and scalable content-based information retrieval. We address this application in the context of the French initiative Pl@ntNet, with CIRAD and IRD.

- **Biological data integration and analysis**.

Biology and its applications, from medicine to agronomy and ecology, are now producing massive data, which is revolutionizing the way life scientists work. For instance, using plant phenotyping platforms such as PhenoDyn and PhenoArch at INRA Montpellier, quantitative genetic methods allow to identify genes involved in phenotypic variation in response to environmental conditions. These methods produce large amounts of data at different time intervals (minutes to months), at different sites and at different scales ranging from small tissue samples to the entire plant until whole plant population. Analyzing such big data creates new challenges for data management and data integration.

- **Audio heritage preservation**.

    Since the end of the 19$^{\text{th}}$ century, France has commissioned ethnologists to record the world's immaterial audio heritage. This results in datasets of dozens of thousands of audio recordings from all countries and more than 1200 ethnies. Today, this data is gathered under the name of 'Archives du CNRS — Musée de l'Homme' and is handled by the CREM (Centre de Recherche en Ethno-Musicologie). Profesional scientists in digital humanities are accessing this data daily for their investigations, and several important challenges arise to ease their work. The KAMoulox project, lead by A. Liutkus, targets at offering online processing tools for the scientists to automatically restore this old material on demand.

These application examples illustrate the diversity of requirements and issues which we are addressing with our scientific application partners. To further validate our solutions and extend the scope of our results, we also want to foster industrial collaborations, even in non scientific applications, provided that they exhibit similar challenges.