



RESEARCH CENTER

FIELD

**Applied Mathematics, Computation  
and Simulation**

Activity Report 2012

**Section Scientific Foundations**

Edition: 2013-04-24



## COMPUTATIONAL MODELS AND SIMULATION

1. BACCHUS Team	5
2. CAD Team	9
3. CAGIRE Team	11
4. CALVI Project-Team	14
5. CASTOR Team	20
6. COFFEE Project-Team	22
7. CONCHA Project-Team	23
8. DEFI Project-Team	26
9. GAMMA3 Project-Team (section vide)	29
10. IPSO Project-Team	30
11. MC2 Project-Team	36
12. MICMAC Project-Team	42
13. NACHOS Project-Team	44
14. NANO-D Team	48
15. OPALE Project-Team	52
16. POEMS Project-Team	54
17. SCIPORT Team	56
18. SIMPAF Project-Team	60

## MODELING, OPTIMIZATION, AND CONTROL OF DYNAMIC SYSTEMS

19. APICS Project-Team	64
20. BIPOP Project-Team	73
21. COMMANDS Project-Team	75
22. CORIDA Project-Team	77
23. DISCO Project-Team	81
24. GECO Team	83
25. MAXPLUS Project-Team	85
26. MCTAO Team	92
27. NECS Project-Team	96
28. NON-A Project-Team	100

## OPTIMIZATION, LEARNING AND STATISTICAL METHODS

29. CLASSIC Project-Team	107
30. DOLPHIN Project-Team	109
31. GEOSTAT Project-Team	115
32. MISTIS Project-Team	120
33. MODAL Project-Team	124
34. REALOPT Project-Team	125
35. SELECT Project-Team	128
36. SEQUEL Project-Team	129
37. SIERRA Project-Team	136
38. TAO Project-Team	137

## STOCHASTIC METHODS AND MODELS

39. ALEA Project-Team .....	143
40. ASPI Project-Team .....	144
41. CQFD Project-Team .....	150
42. I4S Team .....	153
43. MATHRISK Team (section vide) .....	160
44. REGULARITY Project-Team .....	161
45. TOSCA Project-Team .....	170

## BACCHUS Team

### 3. Scientific Foundations

#### 3.1. Numerical schemes for fluid mechanics

**Participants:** Rémi Abgrall, Mario Ricchiuto, Pietro Marco Congedo, Cécile Dobrzynski, Héloïse Beaugendre, Pierre-Henri Maire, Luc Mieussens.

A large number of engineering problems involve fluid mechanics. They may involve the coupling of one or more physical models. An example is provided by aeroelastic problems, which have been studied in details by other Inria teams. Another example is given by flows in pipelines where the fluid (a mixture of air–water–gas) does not have well-known physical properties, and there are even more exotic situations that will be discussed later. Another application is the influence of fluid flow on noise production. Problems in aeroacoustics are indeed becoming more and more important in everyday life. In some occasions, one needs specific numerical tools to take into account *e.g.* a fluids' exotic equation of state, or because the amount of required computational resources becomes huge, as in unsteady flows. Another situation where specific tools are needed is when one is interested in very specific physical quantities, such as *e.g.* the lift and drag of an airfoil, a situation where commercial tools can only provide a very crude answer.

It is a fact that there are many commercial codes. They allow users to simulate a lot of different flow types. The quality of the results is however far from optimal in many cases. Moreover, the numerical technology implemented in these codes is often not the most recent. To give a few examples, consider the noise generated by wake vortices in supersonic flows (external aerodynamics/aeroacoustics), or the direct simulation of a 3D compressible mixing layer in a complex geometry (as in combustion chambers). Up to our knowledge, due to the very different temporal and physical scales that need to be captured, a direct simulation of these phenomena is not in the reach of the most recent technologies because the numerical resources required are currently unavailable. *We need to invent specific algorithms for this purpose.*

In order to simulate efficiently these complex physical problems, we are working on some fundamental aspects of the numerical analysis of non linear hyperbolic problems. *Our goal is to develop more accurate and more efficient schemes that can adapt to modern computer architectures.*

More precisely, *we are working on a class of numerical schemes*, known in literature as Residual Distribution schemes, *specifically tailored to unstructured and hybrid meshes*. They have the most possible compact stencil that is compatible with the expected order of accuracy. *This accuracy is at least of second order, and it can go up to any order of accuracy, even though fourth order is considered for practical applications.* Since the stencil is compact, the implementation on parallel machines becomes simple. These schemes are very flexible in nature, which is so far one of the most important advantage over other techniques. This feature has allowed us to adapt the schemes to the requirements of different physical situations (*e.g.* different formulations allow either an efficient explicit time advancement for problems involving small time-scales, or a fully implicit space-time variant which is unconditionally stable and allows to handle stiff problems where only the large time scales are relevant). This flexibility has also enabled to devise a variant using the same data structure of the popular Discontinuous Galerkin schemes, which are also part of our scientific focus.

The compactness of the second order version of the schemes enables us to use efficiently the high performance parallel linear algebra tools developed by the team. However, the high order versions of these schemes, which are under development, require modifications to these tools taking into account the nature of the data structure used to reach higher orders of accuracy. This leads to new scientific problems at the border between numerical analysis and computer science. In parallel to these fundamental aspects, we also work on adapting more classical numerical tools to complex physical problems such as those encountered in interface flows, turbulent or multiphase flows, geophysical flows, and material science. An effort for developing a more predictive tool for multiphase compressible flows is also underway. Within this project, several advancements have been performed, *i.e.* considering a more complete systems of equations including viscosity, working on the

thermodynamic modeling of complex fluids, and developing stochastic methods for uncertainty quantification in compressible flows.

We expect within a few years to be able to demonstrate the potential of our developments on applications ranging from the reproduction of the complex multidimensional interactions between tidal waves and estuaries, unsteady aerodynamics and aeroacoustics associated to both external and internal compressible flows, compressible ideal and non-ideal MHD (in relation with the ITER project), and the behavior of complex materials. This will be achieved by means of a multi-disciplinary effort involving our research on residual discretizations schemes, the parallel advances in algebraic solvers and partitioners, and the strong interactions with specialists in computer science, scientific computing, physics, mechanics, and mathematical modeling.

Our research in numerical algorithms has led to the development of the `Realfluids` platform which is described in section 5.3. New software developments are under way in the field of free surface flows and complex materials modeling. These developments are performed in the code `SLOWS` (Shallow-water FLOWS) for free surface flows, and in the solver `COCA` (CodeOxydationCompositesAutocicatrissants) for the simulation of the self-healing process in composite materials. These developments will be described in sections 5.10 and 5.2.

This work is supported by the EU-Strep IDIHOM, various research contracts and in part by the ANEMOS project and the ANR-Emergence `RealFluids` grant. A large part of the team also beneficiates of the ADDECCO ERC grant.

## 3.2. Uncertainty quantification

**Participants:** Rémi Abgrall, Mario Ricchiuto, Pietro Marco Congedo.

Another topic of interest is the quantification of uncertainties in non linear problems. In many applications, the physical model is not known accurately. The typical example is that of turbulence models in aeronautics. These models all depend on a number of parameters which can radically change the output of the simulation. Being impossible to lump the large number of temporal and spatial scales of a turbulent flow in a few model parameters, these values are often calibrated to quantitatively reproduce a certain range of effects observed experimentally. A similar situation is encountered in many applications such as real gas or multiphase flows, where the equation of state form suffer from uncertainties, and free surface flows with sediment transport, where often both the hydrodynamic model and the sediment transport model depend on several parameters, and may have more than one formal expression.

This type of uncertainty, called *epistemic*, is associated with a lack of knowledge and could be reduced by further experiments and investigation. Instead, another type of uncertainty, called *aleatory*, is related to the intrinsic aleatory quality of a physical measure and can not be reduced. The dependency of the numerical simulation from these uncertainties can be studied by propagation of chaos techniques such as those developed during the recent years via polynomial chaos techniques. Different implementations exist, depending whether the method is intrusive or not. The accuracy of these methods is still a matter of research, as well how they can handle an as large as possible number of uncertainties or their versatility with respect to the structure of the random variable pdfs. Our objective is to develop some non-intrusive or semi-intrusive methods, trying to define a unified framework for obtaining a reliable and accurate numerical solution at a moderate computational cost. Dealing with high dimensional representation of stochastic inputs in design optimization is computationally prohibitive. In fact, for a robust design, statistics of the fitness functions are also important, then uncertainty quantification (UQ) becomes the predominant issue to handle if a large number of uncertainties is taken into account. Several methods are proposed in literature to consider high dimension stochastic problem but their accuracy on realistic problems where highly non-linear effects could exist is not proven at all. We developed several efficient global strategies for robust optimization: the first class of method is based on the extension of simplex stochastic collocation to the optimization space, the second one consists in hybrid strategies using ANOVA decomposition.

This part of our activities is supported by the ERC grant ADDECCO, the ANR-MN project UFO and the associated team AQUARIUS.

### 3.3. Meshes and scalable discrete data structures

**Participants:** Cécile Dobrzynski, Sébastien Fourestier, Algiane Froehly, Cédric Lachat, François Pellegrini.

#### 3.3.1. Adaptive dynamic mesh partitioning

Many simulations which model the evolution of a given phenomenon along with time (turbulence and unsteady flows, for instance) need to re-mesh some portions of the problem graph in order to capture more accurately the properties of the phenomenon in areas of interest. This re-meshing is performed according to criteria which are closely linked to the undergoing computation and can involve large mesh modifications: while elements are created in critical areas, some may be merged in areas where the phenomenon is no longer critical.

Performing such re-meshing in parallel creates additional problems. In particular, splitting an element which is located on the frontier between several processors is not an easy task, because deciding when splitting some element, and defining the direction along which to split it so as to preserve numerical stability most, require shared knowledge which is not available in distributed memory architectures. Ad-hoc data structures and algorithms have to be devised so as to achieve these goals without resorting to extra communication and synchronization which would impact the running speed of the simulation.

Most of the works on parallel mesh adaptation attempt to parallelize in some way all the mesh operations: edge swap, edge split, point insertion, etc. It implies deep modifications in the (re)mesher and often leads to bad performance in term of CPU time. An other work [51] proposes to base the parallel re-meshing on existing mesher and load balancing to be able to modify the elements located on the frontier between several processors.

In addition, the preservation of load balance in the re-meshed simulation requires dynamic redistribution of mesh data across processing elements. Several dynamic repartitioning methods have been proposed in the literature [52], [50], which rely on diffusion-like algorithms and the solving of flow problems to minimize the amount of data to be exchanged between processors. However, integrating such algorithms into a global framework for handling adaptive meshes in parallel has yet to be done.

The path that we are following bases on the decomposition of the areas to remesh into boules that can be processed concurrently, each by a sequential remesher. It requires to devise scalable algorithms for building such boules, scheduling them on the available processors, reconstructing the remeshed mesh and redistributing its data. This research started within the context of the PhD of Cédric Lachat, funded by a CORDI grant of EPI PUMAS and is continued thanks to a funding by ADT grant E1 Gaucho.

#### 3.3.2. Graph partitioning and static mapping

Unlike their predecessors of two decades ago, today's very large parallel architectures can no longer implement a uniform memory model. They are based on a hierarchical structure, in which cores are assembled into chips, chips are assembled into boards, boards are assembled into cabinets and cabinets are interconnected through high speed, low latency communication networks. On these systems, communication is non uniform: communicating with cores located on the same chip is cheaper than with cores on other boards, and much cheaper than with cores located in other cabinets. The advent of these massively parallel, non uniform machines impacts the design of the software to be executed on them, both for applications and for service tools. It is in particular the case for the software whose task is to balance workload across the cores of these architectures.

A common method for task allocation is to use graph partitioning tools. The elementary computations to perform are represented by vertices and their dependencies by edges linking two vertices that need to share some piece of data. Finding good solutions to the workload distribution problem amounts to computing partitions with small vertex or edge cuts and that balance evenly the weights of the graph parts. Yet, computing efficient partitions for non uniform architectures requires to take into account the topology of the target architecture. When processes are assumed to coexist simultaneously for all the duration of the program, this generalized optimization problem is called mapping. In this problem, the communication cost function to minimize incorporates architecture-dependent, locality improving terms, such as the dilation of each edge (that

is, by how much it is “stretched” across the graph representing the target architecture), which is sometimes also expressed as some “hop metric”. A mapping is called static if it is computed prior to the execution of the program and is never modified at run-time.

The sequential *Scotch* tool being developed within the BACCHUS team (see Section 5.9) was able to perform static mapping since its first version, in 1994, but this feature was not widely known nor used by the community. With the increasing need to map very large problem graphs onto very large and strongly non uniform parallel machines, there is an increasing demand for parallel static mapping tools. Since, in the context of dynamic repartitioning, parallel mapping software will have to run on their target architectures, parallel mapping and remapping algorithms suitable for efficient execution on such heterogeneous architectures have to be investigated. This leads to solve three interwoven challenges:

- scalability: such algorithms must be able to map graphs of more than a billion vertices onto target architectures comprising millions of cores;
- heterogeneity: not only do these algorithms must take into account the topology of the target architecture they map graphs onto, but they also have themselves to run efficiently on these very architectures;
- asynchronicity: most parallel partitioning algorithms use collective communication primitives, that is, some form of heavy synchronization. With the advent of machines having several millions of cores, and in spite of the continuous improvement of communication subsystems, the demand for more asynchronicity in parallel algorithms is likely to increase.

This research takes place within the context of the PhD of Sébastien Fourestier.



## CAD Team

### 3. Scientific Foundations

#### 3.1. Geometry continuity and $\varepsilon$ Geometry Continuity

The mathematical background of parametric surfaces is Differential Geometry. In differential geometry, Riemann (1826-1866), Shiing-Shen Chern (1911-2004), continuities play a very important kernel role. In 1980s, more and more engineering design using geometry modeling softwares found the problems of the parametric continuities. And the order of the parametric continuity depends on how the curve is parameterized. To day, engineers and scientists try to find a kind of continuities, which are the intuitive intrinsic properties of curves and surfaces, and the orders of the continuities are independent of the parameterization.

$G$ -Continuity could be defined as the smoothness properties of a curve or a surface that are more than its order of differentiability. This problem is complex and progress in this domain is very slow. We proposed new ways to make through the bottleneck. Furthermore, we also wanted to fill the gap between the traditional mathematics and modern computer science. Hence, we developed the theories of epsilon-geometry continuities to accommodate the representation and the rounding errors of float-point arithmetic, and design new geometric modeling operators under the constraints of epsilon-geometry continuities.

#### 3.2. Two main challenges in Computer Aided Design

##### 3.2.1. Robustness tolerance, error control

Based on this theoretical contribution, we also proposed several elegant solutions to the most important challenges in Computer Aided Design (see Lees A Piegl. "Ten challenges in Computer-Aided-Design". *Jal of CAD* 2005. 37 (4): 461-470): robustness, tolerances, error control, geometric arrangement, beautification and modelling of complex shapes. During CAD processes one uses a myriad of tolerances, many of which are not directly related to the actual manufacturing process. Some interesting questions here include: What are the most relevant machining tolerances? How to set the army of computational tolerances, e.g. those of systems of equations, to guarantee machining within the required accuracy? How tolerances in different spaces, e.g. in model space and in parameter space, are related. Numerical instabilities also account for the majority of computational errors in commercial CAD systems.

The problems related to robustness haunt every programmer who has ever worked on commercial systems. Fixing numerical bugs can be very frustrating, and often times results in patching up the code simply because no solution exists to remedy the problem.

##### 3.2.2. Geometry beautification, Geometry operators and Shape generation

Although geometric uncertainties are related to robustness and tolerance, there are a number of extra issues well worth deeper investigations. Geometric arrangements are full of special cases. The most notable ones are: cases of touch, overlapping, containment, etc.; cases of parallelism, perpendicularity, coincidence, etc.; axes of symmetrical data, data clustering, dense or sparse data, etc.; cases of degeneracy, discontinuity, inconsistencies, etc.; problems with cracks, excess material, lack of detail, etc. In just about any code that deals with geometry, the number of special cases is significantly larger than the general ones. Data explosion is the result of careless selection of the methods, e.g. parameter space-based sampling, and improper implementation, e.g. recursive algorithms. Some of the relevant issues are: sampling: over sampling, sampling in incorrect places, etc; procedural definitions, e.g. lofting a large set of curves or merging surfaces may result in an explosion of control points.

Furthermore, although CAD processes are supposed to produce valid and "made to order" models, the reality is that most (if not all) models are rough and require post-processing, i.e. beautification. Some of the most frequently needed tasks are: removing unwanted edges, corners, cracks, etc.; removing bumps, oscillations, curvature extremes, etc.; healing incorrect models, e.g. removing holes in triangulations; smoothing, fairing, re-shaping, etc.

### **3.3. Computer Graphics**

In Computer Graphics, objectives were to prove the capability of the team to address some topics as Computational Photography, Rendering and Computer Animation. Work in Progress in these topics are described in the following chapter.

## CAGIRE Team

### 3. Scientific Foundations

#### 3.1. Computational fluid mechanics: resolving versus modelling small scales of turbulence

A typical continuous solution of the Navier Stokes equations is governed by a spectrum of time and space scales. The broadness of that spectrum is directly controlled by the Reynolds number defined as the ratio between the inertial forces and the viscous forces. This number is quite helpful to determine if the flow is turbulent or not. In the former case, it indicates the range of scales of fluctuations that are present in the flow under study. Typically, for instance for the velocity field, the ratio between the largest scale (the integral length scale) to the smallest one (Kolmogorov scale) scales as  $Re^{3/4}$ . The smallest scales may have a certain effect on the largest ones which implies that an accurate framework for the computation of flows must take into account all these scales. This can be achieved either by solving directly the Navier-Stokes equations (Direct numerical simulations or DNS) or by first applying a time filtering (Reynolds Average Navier-Stokes or RANS) or a spatial filtering operator to the Navier-Stokes equations (large-eddy simulations or LES). The new terms brought about by the filtering operator have to be modelled. From a computational point of view, the RANS approach is the less demanding, which explains why historically it has been the workhorse in both the academic and the industrial sectors. Although it has permitted quite a substantive progress in the understanding of various phenomena such as turbulent combustion or heat transfer, its inability to provide a time-dependent information has led to promote in the last decade the recourse to either LES or DNS. By simulating the large scale structures while modelling the smallest ones supposed to be more isotropic, LES proved to be quite a step through that permits to fully take advantage of the increasing power of computers to study complex flow configurations. In the same time, DNS was progressively applied to geometries of increasing complexity (channel flows, jets, turbulent premixed flames), and proved to be a formidable tool that permits (i) to improve our knowledge of turbulent flows and (ii) to test (i.e. validate or invalidate) and improve the numerous modelling hypotheses inherently associated to the RANS and LES approaches. From a numerical point of view, if the steady nature of the RANS equations allows to perform iterative convergence on finer and finer meshes, this is no longer possible for LES or DNS which are time-dependent. It is therefore necessary to develop high accuracy schemes in such frameworks. Considering that the Reynolds number in an engine combustion chamber is significantly larger than 10000, a direct numerical simulation of the whole flow domain is not conceivable on a routine basis but the simulation of generic flows which feature some of the phenomena present in a combustion chamber is accessible considering the recent progresses in High Performance Computing (HPC). Along these lines, our objective is to develop a DNS tool to simulate a jet in crossflow configuration which is the generic flow of an aeronautical combustion chamber as far as its effusion cooling is concerned.

#### 3.2. Computational fluid mechanics: numerical methods

All the methods we describe are mesh-based methods: the computational domain is divided into *cells*, that have an elementary shape: triangle and quadrangle in two dimensions, and tetrahedra, hexahedra, pyramids, and prism in three dimensions. If the cells are only regular hexahedra, the mesh is said to be *structured*. Otherwise, it is said to be unstructured. If the mesh is composed of more than one sort of elementary shape, the mesh is said to be *hybrid*.

The basic numerical model for the computation of internal flows is based on the Navier-Stokes equations. For fifty years, many sorts of numerical approximation have been tried for this sort of system: finite differences, finite volumes, and finite elements.

The finite differences have met a great success for some equations, but for the approximation of fluid mechanics, they suffer from two drawbacks. First, structured meshes must be used. This drawback can be very limiting in the context of internal aerodynamics, in which the geometries can be very complex. The other problem is that finite difference schemes do not include any upwinding process, which is essential for convection dominated flows.

The finite volumes methods have imposed themselves in the last thirty years in the context of aerodynamic. They intrinsically contain an upwinding mechanism, so that they are naturally stable for linear as much as for nonlinear convective flows. The extension to diffusive flows has been done in [11]. Whereas the extension to second order with the MUSCL method is widely spread, the extension to higher order has always been a strong drawback of finite volumes methods. For such an extension, reconstruction methods have been developed (ENO, WENO). Nevertheless, these methods need to use a stencil that increases quickly with the order, which induces problems for the parallelisation and the efficiency of the implementation. Another natural extension of finite volume methods are the so-called discontinuous Galerkin methods. These methods are based on the Galerkin' idea of projecting the weak formulation of the equations on a finite dimensional space. But on the contrary to the conforming finite elements method, the approximation space is composed of functions that are continuous (typically: polynomials) inside each cell, but that are discontinuous on the sides. The discontinuous Galerkin methods are currently very popular, because they can be used with many sort of partial differential equations. Moreover, the fact that the approximation is discontinuous allows to use modern mesh adaptation (hanging nodes, which appear in non conforming mesh adaptation), and adaptive order, in which the high order is used only where the solution is smooth.

Discontinuous Galerkin methods were introduced by Reed and Hill [31] and first studied by Lesaint and Raviart [24]. The extension to the Euler system with explicit time integration was mainly led by Shu, Cockburn and their collaborators. The steps of time integration and slope limiting were similar to high order ENO schemes, whereas specific constraints given by the finite elements nature of the scheme were progressively solved, for scalar conservation laws [15], [14], one dimensional systems [13], multidimensional scalar conservation laws [12], and multidimensional systems [16]. For the same system, we can also cite the work of [18], [22], which is slightly different: the stabilisation is made by adding a nonlinear stabilisation term, and the time integration is implicit. Then, the extension to the compressible Navier-Stokes system was made by Bassi and Rebay [10], first by a mixed type finite element method, and then simplified by means of lifting operators. The extension to the  $k - \omega$  RANS system was made in [9]. Another type of discontinuous Galerkin method for Navier Stokes is the so-called Symmetric Interior Penalty (SIP) method. It is used for example by Hartmann and Houston [20]. The symmetric nature of the discretization is particularly well suited with mesh adaptation by means of the adjoint equation resolution [21]. Last, we note that the discontinuous Galerkin method was already successfully tested in [17] at Direct Numerical Simulation scale for very moderate Reynolds, and also by Munz'team in Stuttgart [25], with local time stepping.

For concluding this section, there already exist numerical schemes based on the discontinuous Galerkin method which proved to be efficient for computing compressible viscous flows. Nevertheless, there remain things to be improved, which include for example: efficient shock capturing term methods for supersonic flows, high order discretization of curved boundaries, or low Mach behaviour of these schemes (this last point will be detailed in the next subsection). Another drawback of the discontinuous Galerkin methods is that they are very computationally costly, due to the accurate representation of the solution. A particular care must be taken on the implementation for being efficient.

### 3.3. Experimental aspects

A great deal of experiments has been devoted to the study of jet in crossflow configurations. They essentially differ one from each other by the hole shape (cylindrical or shaped), the hole axis inclination, the way by which the hole is fed, the characteristics of the crossflow and the jet (turbulent or not, isothermal or not), the number of holes considered and last but not least the techniques used to investigate the flow. A good starting point to assess the diversity of the studies carried out is given by [26]. For inclined cylindrical holes, the experimental

database produced by Gustafsson and Johansson <sup>2</sup> represents a sound reference base and for normal injection, the work by [32] served as reference for LES simulations [30]. For shaped holes, the studies are less numerous and are aimed at assessing the influence of the hole shape on various flow properties such as the heat transfer at the wall [23]. In 2007, A. Most developed at UPPA a test facility for studying jet in crossflow issued from shaped holes [27]. The hole shape was chosen as a 12.5 scale of the holes (i.e. at scale 1) drilled by laser in a combustion chamber. His preliminary 2-component PIV results have been used to test RANS simulations [28] and LES [29]. This test facility will be used in the framework of the present project to investigate a 1-hole plane i.e. an isolated jet in crossflow. PIV and LDV metrology will be used.

---

<sup>2</sup>[http://www.tfd.chalmers.se/~gujo/WS11\\_2005/Slanted\\_jet/INDEX.HTM](http://www.tfd.chalmers.se/~gujo/WS11_2005/Slanted_jet/INDEX.HTM)

## CALVI Project-Team

### 3. Scientific Foundations

#### 3.1. Kinetic models for plasma and beam physics

plasma physics, beam physics, kinetic models, reduced models, Vlasov equation, modeling, mathematical analysis, asymptotic analysis, existence, uniqueness

Plasmas and particle beams can be described by a hierarchy of models including  $N$ -body interaction, kinetic models and fluid models. Kinetic models in particular are posed in phase-space and involve specific difficulties. We perform a mathematical analysis of such models and try to find and justify approximate models using asymptotic analysis.

##### 3.1.1. Models for plasma and beam physics

The **plasma state** can be considered as the **fourth state of matter**, obtained for example by bringing a gas to a very high temperature ( $10^4 K$  or more). The thermal energy of the molecules and atoms constituting the gas is then sufficient to start ionization when particles collide. A globally neutral gas of neutral and charged particles, called **plasma**, is then obtained. Intense charged particle beams, called nonneutral plasmas by some authors, obey similar physical laws.

The hierarchy of models describing the evolution of charged particles within a plasma or a particle beam includes  $N$ -body models where each particle interacts directly with all the others, kinetic models based on a statistical description of the particles and fluid models valid when the particles are at a thermodynamical equilibrium.

In a so-called *kinetic model*, each particle species  $s$  in a plasma or a particle beam is described by a distribution function  $f_s(\mathbf{x}, \mathbf{v}, t)$  corresponding to the statistical average of the particle distribution in phase-space corresponding to many realisations of the physical system under investigation. The product  $f_s d\mathbf{x} d\mathbf{v}$  is the average number of particles of the considered species, the position and velocity of which are located in a bin of volume  $d\mathbf{x} d\mathbf{v}$  centered around  $(\mathbf{x}, \mathbf{v})$ . The distribution function contains a lot more information than what can be obtained from a fluid description, as it also includes information about the velocity distribution of the particles.

A kinetic description is necessary in collective plasmas where the distribution function is very different from the Maxwell-Boltzmann (or Maxwellian) distribution which corresponds to the thermodynamical equilibrium, otherwise a fluid description is generally sufficient. In the limit when collective effects are dominant with respect to binary collisions, the corresponding kinetic equation is the *Vlasov equation*

$$\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \frac{\partial f_s}{\partial \mathbf{x}} + \frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_s}{\partial \mathbf{v}} = 0,$$

which expresses that the distribution function  $f$  is conserved along the particle trajectories which are determined by their motion in their mean electromagnetic field. The Vlasov equation which involves a self-consistent electromagnetic field needs to be coupled to the Maxwell equations in order to compute this field

$$\begin{aligned} -\frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} + \nabla \times \mathbf{B} &= \mu_0 \mathbf{J}, \\ \frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} &= 0, \\ \operatorname{div} \mathbf{E} &= \frac{\rho}{\varepsilon_0}, \\ \operatorname{div} \mathbf{B} &= 0, \end{aligned}$$

which describes the evolution of the electromagnetic field generated by the charge density

$$\rho(\mathbf{x}, t) = \sum_s q_s \int f_s(\mathbf{x}, \mathbf{v}, t) d\mathbf{v},$$

and current density

$$\mathbf{J}(\mathbf{x}, t) = \sum_s q_s \int f_s(\mathbf{x}, \mathbf{v}, t) \mathbf{v} d\mathbf{v},$$

associated to the charged particles.

When binary particle-particle interactions are dominant with respect to the mean-field effects then the distribution function  $f$  obeys the Boltzmann equation

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} = Q(f, f),$$

where  $Q$  is the nonlinear Boltzmann collision operator. In some intermediate cases, a collision operator needs to be added to the Vlasov equation.

The numerical solution of the three-dimensional Vlasov-Maxwell system represents a considerable challenge due to the huge size of the problem. Indeed, the Vlasov-Maxwell system is nonlinear and posed in phase space. It thus depends on seven variables: three configuration space variables, three velocity space variables and time, for each species of particles. This feature makes it essential to use every possible option to find a reduced model wherever possible, in particular when there are geometrical symmetries or small terms which can be neglected.

### 3.1.2. *Mathematical and asymptotic analysis of kinetic models*

The mathematical analysis of the Vlasov equation is essential for a thorough understanding of the model as well for physical as for numerical purposes. It has attracted many researchers since the end of the 1970s. Among the most important results which have been obtained, we can cite the existence of strong and weak solutions of the Vlasov-Poisson system by Horst and Hunze [74], see also Bardos and Degond [51]. The existence of a weak solution for the Vlasov-Maxwell system has been proved by Di Perna and Lions [59]. An overview of the theory is presented in a book by Glassey [71].

Many questions concerning for example uniqueness or existence of strong solutions for the three-dimensional Vlasov-Maxwell system are still open. Moreover, there is a realm of approached models that need to be investigated. In particular, the Vlasov-Darwin model for which we could recently prove the existence of global solutions for small initial data [52].

On the other hand, the asymptotic study of the Vlasov equation in different physical situations is important in order to find or justify reduced models. One situation of major importance in tokamaks, used for magnetic fusion as well as in atmospheric plasmas, is the case of a large external magnetic field used for confining the particles. The magnetic field tends to incurve the particle trajectories which eventually, when the magnetic field is large, are confined along the magnetic field lines. Moreover, when an electric field is present, the particles drift in a direction perpendicular to the magnetic and to the electric field. The new time scale linked to the cyclotron frequency, which is the frequency of rotation around the magnetic field lines, comes in addition to the other time scales present in the system like the plasma frequencies of the different particle species. Thus, many different time scales as well as length scales linked in particular to the different Debye length are present in the system. Depending on the effects that need to be studied, asymptotic techniques allow to find reduced models. In this spirit, in the case of large magnetic fields, recent results have been obtained by Golse and Saint-Raymond [72], [80] as well as by Brenier [57]. Our group has also contributed to this problem using homogenization techniques to justify the guiding center model and the finite Larmor radius model which are used by physicist in this setting [67], [65], [66].

Another important asymptotic problem yielding reduced models for the Vlasov-Maxwell system is the fluid limit of collisionless plasmas. In some specific physical situations, the infinite system of velocity moments of the Vlasov equations can be closed after a few of those, thus yielding fluid models.

## 3.2. Development of simulation tools

Numerical methods, Vlasov equation, unstructured grids, adaptivity, numerical analysis, convergence, Semi-Lagrangian method The development of efficient numerical methods is essential for the simulation of plasmas and beams. Indeed, kinetic models are posed in phase space and thus the number of dimensions is doubled. Our main effort lies in developing methods using a phase-space grid as opposed to particle methods. In order to make such methods efficient, it is essential to consider means for optimizing the number of mesh points. This is done through different adaptive strategies. In order to understand the methods, it is also important to perform their mathematical analysis. Since a few years we are interested also with solvers that uses Particle In Cell method. This new issue allows us to enrich some parts of our research activities previously centered on the Semi-Lagrangian approach.

### 3.2.1. Introduction

The numerical integration of the Vlasov equation is one of the key challenges of computational plasma physics. Since the early days of this discipline, an intensive work on this subject has produced many different numerical schemes. One of those, namely the Particle-In-Cell (PIC) technique, has been by far the most widely used. Indeed it belongs to the class of Monte Carlo particle methods which are independent of dimension and thus become very efficient when dimension increases which is the case of the Vlasov equation posed in phase space. However these methods converge slowly when the number of particles increases, hence if the complexity of grid based methods can be decreased, they can be the better choice in some situations. This is the reason why one of the main challenges we address is the development and analysis of adaptive grid methods.

### 3.2.2. Convergence analysis of numerical schemes

Exploring grid based methods for the Vlasov equation, it becomes obvious that they have different stability and accuracy properties. In order to fully understand what are the important features of a given scheme and how to derive schemes with the desired properties, it is essential to perform a thorough mathematical analysis of this scheme, investigating in particular its stability and convergence towards the exact solution.

### 3.2.3. The semi-Lagrangian method

The semi-Lagrangian method consists in computing a numerical approximation of the solution of the Vlasov equation on a phase space grid by using the property of the equation that the distribution function  $f$  is conserved along characteristics. More precisely, for any times  $s$  and  $t$ , we have

$$f(\mathbf{x}, \mathbf{v}, t) = f(\mathbf{X}(s; \mathbf{x}, \mathbf{v}, t), \mathbf{V}(s; \mathbf{x}, \mathbf{v}, t), s),$$

where  $(\mathbf{X}(s; \mathbf{x}, \mathbf{v}, t), \mathbf{V}(s; \mathbf{x}, \mathbf{v}, t))$  are the characteristics of the Vlasov equation which are solution of the system of ordinary differential equations

$$\begin{aligned} \frac{d\mathbf{X}}{ds} &= \mathbf{V}, \\ \frac{d\mathbf{V}}{ds} &= \mathbf{E}(\mathbf{X}(s), s) + \mathbf{V}(s) \times \mathbf{B}(\mathbf{X}(s), s), \end{aligned} \tag{1}$$

with initial conditions  $\mathbf{X}(t) = \mathbf{x}$ ,  $\mathbf{V}(t) = \mathbf{v}$ .



From this property,  $f^n$  being known one can induce a numerical method for computing the distribution function  $f^{n+1}$  at the grid points  $(\mathbf{x}_i, \mathbf{v}_j)$  consisting in the following two steps:

1. For all  $i, j$ , compute the origin of the characteristic ending at  $\mathbf{x}_i, \mathbf{v}_j$ , i.e. an approximation of  $\mathbf{X}(t_n; \mathbf{x}_i, \mathbf{v}_j, t_{n+1}), \mathbf{V}(t_n; \mathbf{x}_i, \mathbf{v}_j, t_{n+1})$ .
2. As

$$f^{n+1}(\mathbf{x}_i, \mathbf{v}_j) = f^n(\mathbf{X}(t_n; \mathbf{x}_i, \mathbf{v}_j, t_{n+1}), \mathbf{V}(t_n; \mathbf{x}_i, \mathbf{v}_j, t_{n+1})),$$

$f^{n+1}$  can be computed by interpolating  $f^n$  which is known at the grid points at the points  $\mathbf{X}(t_n; \mathbf{x}_i, \mathbf{v}_j, t_{n+1}), \mathbf{V}(t_n; \mathbf{x}_i, \mathbf{v}_j, t_{n+1})$ .

This method can be simplified by performing a time-splitting separating the advection phases in physical space and velocity space, as in this case the characteristics can be solved explicitly.

### 3.2.4. Adaptive semi-Lagrangian methods

Uniform meshes are most of the time not efficient to solve a problem in plasma physics or beam physics as the distribution of particles is evolving a lot as well in space as in time during the simulation. In order to get optimal complexity, it is essential to use meshes that are fitted to the actual distribution of particles. If the global distribution is not uniform in space but remains locally mostly the same in time, one possible approach could be to use an unstructured mesh of phase space which allows to put the grid points as desired. Another idea, if the distribution evolves a lot in time is to use a different grid at each time step which is easily feasible with a semi-Lagrangian method. And finally, the most complex and powerful method is to use a fully adaptive mesh which evolves locally according to variations of the distribution function in time. The evolution can be based on a posteriori estimates or on multi-resolution techniques.

### 3.2.5. Particle-In-Cell codes

The Particle-In-Cell method [56] consists in solving the Vlasov equation using a particle method, i.e. advancing numerically the particle trajectories which are the characteristics of the Vlasov equation, using the equations of motion which are the ordinary differential equations defining the characteristics. The self-fields are computed using a standard method on a structured or unstructured grid of physical space. The coupling between the field solve and the particle advance is done on the one hand by depositing the particle data on the grid to get the charge and current densities for Maxwell's equations and, on the other hand, by interpolating the fields at the particle positions. This coupling is one of the difficult issues and needs to be handled carefully.

### 3.2.6. Maxwell's equations in singular geometry

The solutions to Maxwell's equations are *a priori* defined in a function space such that the curl and the divergence are square integrable and that satisfy the electric and magnetic boundary conditions. Those solutions are in fact smoother (all the derivatives are square integrable) when the boundary of the domain is smooth or convex. This is no longer true when the domain exhibits non-convex *geometrical singularities* (corners, vertices or edges).

Physically, the electromagnetic field tends to infinity in the neighbourhood of the re-entrant singularities, which is a challenge to the usual finite element methods. Nodal elements cannot converge towards the physical solution. Edge elements demand considerable mesh refinement in order to represent those infinities, which is not only time- and memory-consuming, but potentially catastrophic when solving time dependent equations: the CFL condition then imposes a very small time step. Moreover, the fields computed by edge elements are discontinuous, which can create considerable numerical noise when the Maxwell solver is embedded in a plasma (e.g. PIC) code.

In order to overcome this dilemma, a method consists in splitting the solution as the sum of a *regular* part, computed by nodal elements, and a *singular* part which we relate to singular solutions of the Laplace operator, thus allowing to calculate a local analytic representation. This makes it possible to compute the solution precisely without having to refine the mesh.

This *Singular Complement Method* (SCM) had been developed [49] and implemented [48] in plane geometry. An especially interesting case is axisymmetric geometry. This is still a 2D geometry, but more realistic than the plane case; despite its practical interest, it had been subject to much fewer theoretical studies [54]. The non-density result for regular fields was proven [58], the singularities of the electromagnetic field were related to that of modified Laplacians [45], and expressions of the singular fields were calculated [46]. Thus the SCM was extended to this geometry. It was then implemented by F. Assous (now at Bar-Ilan University, Israel) and S. Labrunie in a PIC–finite element Vlasov–Maxwell code [47].

As a byproduct, space-time regularity results were obtained for the solution to time-dependent Maxwell’s equation in presence of geometrical singularities in the plane and axisymmetric cases [70], [46].

### 3.3. Large size problems

Parallelism, domain decomposition, code transformation

#### 3.3.1. Introduction

The applications we consider lead to very large size computational problems for which we need to apply modern computing techniques enabling to use efficiently many computers including traditional high performance parallel computers and computational grids.

The full Vlasov-Maxwell system yields a very large computational problem mostly because the Vlasov equation is posed in six-dimensional phase-space. In order to tackle the most realistic possible physical problems, it is important to use all the modern computing power and techniques, in particular parallelism and grid computing.

#### 3.3.2. Parallelization of numerical methods

An important issue for the practical use of the methods we develop is their parallelization. We address the problem of tuning these methods to homogeneous or heterogeneous architectures with the aim of meeting increasing computing resources requirements.

Most of the considered numerical methods apply a series of operations identically to all elements of a geometric data structure: the mesh of phase space. Therefore these methods intrinsically can be viewed as a data-parallel algorithm. A major advantage of this data-parallel approach derives from its scalability. Because operations may be applied identically to many data items in parallel, the amount of parallelism is dictated by the problem size.

Parallelism, for such data-parallel PDE solvers, is achieved by partitioning the mesh and mapping the sub-meshes onto the processors of a parallel architecture. A good partition balances the workload while minimizing the communications overhead. Many interesting heuristics have been proposed to compute near-optimal partitions of a (regular or irregular) mesh. For instance, the heuristics based on space-filling curves [73] give very good results for a very low cost.

Adaptive methods include a mesh refinement step and can highly reduce memory usage and computation volume. As a result, they induce a load imbalance and require to dynamically distribute the adaptive mesh. A problem is then to combine distribution and resolution components of the adaptive methods with the aim of minimizing communications. Data locality expression is of major importance for solving such problems. We use our experience of data-parallelism and the underlying concepts for expressing data locality [81], optimizing the considered methods and specifying new data-parallel algorithms.

As a general rule, the complexity of adaptive methods requires to define software abstractions allowing to separate/integrate the various components of the considered numerical methods (see [79] as an example of such modular software infrastructure).

Another key point is the joint use of heterogeneous architectures and adaptive meshes. It requires to develop new algorithms which include new load balancing techniques. In that case, it may be interesting to combine several parallel programming paradigms, i.e. data-parallelism with other lower-level ones.

Moreover, exploiting heterogeneous architectures requires the use of a run time support associated with a programming interface that enables some low-level hardware characteristics to be unified. Such run time support is the basis for heterogeneous algorithms. Candidates for such a run time support may be specific implementations of MPI such as MPICH-G2 (a grid-enabled MPI implementation on top of the GLOBUS tool kit for grid computing [64]).

Our general approach for designing efficient parallel algorithms is to define code transformations at any level. These transformations can be used to incrementally tune codes to a target architecture and they warrant code reusability.

## CASTOR Team

### 3. Scientific Foundations

#### 3.1. Plasma Physics

**Participants:** Hervé Guillard, Boniface Nkonga, Afeintou Sangam, Richard Pasquetti, Audrey Bonnement, Marie Martin, Cédric Lachat, Laure Combe, Jacques Blum, Cédric Boulbe, Sebastian Minjeaud.

In order to fulfil the increasing demand, alternative energy sources have to be developed. Indeed, the current rate of fossil fuel usage and its serious adverse environmental impacts (pollution, greenhouse gas emissions, ...) lead to an energy crisis accompanied by potentially disastrous global climate changes.

Controlled fusion power is one of the most promising alternatives to the use of fossil resources, potentially with a unlimited source of fuel. France with the ITER (<http://www.iter.org/default.aspx>) and Laser Megajoule (<http://www-lmj.cea.fr/>) facilities is strongly involved in the development of these two parallel approaches to master fusion that are magnetic and inertial confinement. Although the principles of fusion reaction are well understood from nearly sixty years, (the design of tokamak dates back from studies done in the '50 by Igor Tamm and Andreï Sakharov in the former Soviet Union), the route to an industrial reactor is still long and the application of controlled fusion for energy production is beyond our present knowledge of related physical processes. In magnetic confinement, beside technological constraints involving for instance the design of plasma-facing component, one of the main difficulties in the building of a controlled fusion reactor is the poor confinement time reached so far. This confinement time is actually governed by turbulent transport that therefore determines the performance of fusion plasmas. The prediction of the level of turbulent transport in large machines such as ITER is therefore of paramount importance for the success of the researches on controlled magnetic fusion.

The other route for fusion plasma is inertial confinement. In this latter case, large scale hydrodynamical instabilities prevent a sufficient large energy deposit and lower the return of the target. Therefore, for both magnetic and inertial confinement technologies, the success of the projects is deeply linked to the theoretical understanding of plasma turbulence and flow instabilities as well as to mathematical and numerical improvements enabling the development of predictive simulation tools.

#### 3.2. Turbulence Modelling

**Participants:** Alain Dervieux, Boniface Nkonga, Richard Pasquetti.

Fluid turbulence has a paradoxical situation in science. The Navier-Stokes equations are an almost perfect model that can be applied to any flow. However, they cannot be solved for any flow of direct practical interest. Turbulent flows involve instability and strong dependence to parameters, chaotic succession of more or less organised phenomena, small and large scales interacting in a complex manner. It is generally necessary to find a compromise between neglecting a huge number of small events and predicting more or less accurately some larger events and trends.

In this direction, CASTOR wishes to contribute to the progress of methods for the prediction of fluid turbulence. Taking benefit of its experience in numerical methods for complex applications, CASTOR works out models for predicting flows around complex obstacles, that can be moved or deformed by the flow, and involving large turbulent structures. Taking into account our ambition to provide also short term methods for industrial problems, we consider methods applying to high Reynolds flows, and in particular, methods hybridizing Large Eddy Simulation (LES) with Reynolds Averaging.

Turbulence is the indirect cause of many other phenomena. Fluid-structure interaction is one of them, and can manifest itself for example in Vortex Induced Motion or Vibration. These phenomena can couple also with liquid-gas interfaces and bring new problems. Of particular interest is also the study of turbulence generated noise. In this field, though acoustic phenomena can also in principle be described by the Navier-Stokes equations, they are not generally numerically solved by flow solvers but rather by specialized linear and nonlinear acoustic solvers. An important question is the investigation of the best way to combine a LES simulation with the acoustic propagation of the waves it produces.

### 3.3. Astrophysical and Environmental flows

**Participants:** Hervé Guillard, Boniface Nkonga, Sebastian Minjeaud.

Although it seems inappropriate to address the modeling of experimental devices of the size of a tokamak and for instance, astrophysical systems with the same mathematical and numerical tools, it has long been recognized that the behaviour of these systems have a profound unity. This has for consequence for instance that any large conference on plasma physics includes sessions on astrophysical plasmas as well as sessions on laboratory plasmas. CASTOR does not intend to consider fluid models coming from Astrophysics or Environmental flows for themselves. However, the team is interested in the numerical approximation of some problems in this area as they provide interesting reduced models for more complex phenomena. To be more precise, let us give some concrete examples : The development of Rossby waves <sup>1</sup> a common problem in weather prediction has a counterpart in the development of magnetic shear induced instabilities in tokamaks and the understanding of this latter type of instabilities has been largely improved by the Rossby wave model. A second example is the water bag model of plasma physics that has a lot in common with multi-layer shallow water system.

To give a last example, we can stress that the development of the so-called well-balanced finite volume schemes used nowadays in many domains of mathematical physics or engineering was largely motivated by the desire to suppress some problems appearing in the approximation of the shallow water system.

Our goal is therefore to use astrophysical or geophysical models to investigate some numerical questions in contexts that, in contrast with plasma physics or fluid turbulence, do not require huge three dimensional computations but are still of interest for themselves and not only as toy models.

---

<sup>1</sup>Rossby waves are giant meanders in high altitude wind that have major influence on weather. Oceanic Rossby waves are also known to exist and to affect the world ocean circulation

## **COFFEE Project-Team**

### **3. Scientific Foundations**

#### **3.1. Scientific Foundations**

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

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

To be more specific:

- We consider evolution problems describing highly heterogeneous flows (with different phases or with high density ratio). In turn, we are led to deal with non linear systems of PDEs of convection and/or convection–diffusion type.
- The nature of the coupling between the equations can be two-fold, which leads to different difficulties, both in terms of analysis and conception of numerical methods. For instance, the system can couple several equations of different types (elliptic/parabolic, parabolic/hyperbolic, parabolic or elliptic with algebraic constraints, parabolic with degenerate coefficients....). Furthermore, the unknowns can depend on different sets of variables, a typical example being the fluid/kinetic models for particulate flows. In turn, the simulation cannot use a single numerical approach to treat all the equations. Instead, hybrid methods have to be designed which raise the question of fitting them in an appropriate way, both in terms of consistency of the discretization and in terms of stability of the whole computation. For the problems under consideration, the coupling can also arise through interface conditions. It naturally occurs when the physical conditions are highly different in subdomains of the physical domain in which the flows takes place. Hence interface conditions are intended to describe the exchange (of mass, energy...) between the domains. Again it gives rise to rather unexplored mathematical questions, and for numerics it yields the question of defining a suitable matching at the discrete level, that is requested to preserve the properties of the continuous model.
- By nature the problems we wish to consider involve many different scales (of time or length basically). It raises two families of mathematical questions. In terms of numerical schemes, the multiscale feature induces the presence of stiff terms within the equations, which naturally leads to stability issues. A clear understanding of scale separation helps in designing efficient methods, based on suitable splitting techniques for instance. On the other hand asymptotic arguments can be used to derive hierarchy of models and to identify physical regimes in which a reduced set of equations can be used.

## CONCHA Project-Team

### 3. Scientific Foundations

#### 3.1. Challenges related to numerical simulations of complex flows

First, we describe some typical difficulties in our fields of application which require the improvement of established and the development of new methods.

- **Coupling of equations and models**  
The general equations of fluid dynamics consist in a strongly coupled nonlinear system. Its mathematical nature depends on the precise model, but in general contains hyperbolic, parabolic, and elliptic parts. The spectrum of physical phenomena described by these equations is very large: convection, diffusion, waves... In addition, it is often necessary to couple different models in order to describe different parts of a mechanical system: chemistry, fluid-fluid-interaction, fluid-solid-interaction...
- **Robustness with respect to physical parameters**  
The values of physical parameters such as diffusion coefficients and constants describing different state equations and material laws lead to different behaviour characterized for example by the Reynolds, Mach, and Weissenberg numbers. Optimized numerical methods are available in many situations, but it remains a challenging problem in some fields of applications to develop robust discretizations and solution algorithms.
- **Multiscale phenomena**  
The inherent nonlinearities lead to an interplay of a wide range of physical modes, well-known for example from the study of turbulent flows. Since the resolution of all modes is often unreachable, it is a challenging task to develop numerical methods, which are still able to reproduce the essential features of the physical phenomenon under study.

#### 3.2. Stabilized and discontinuous finite element methods

The discontinuous Galerkin method [68], [66], [44], [43] has gained enormous success in CFD due to its flexibility, links with finite volume methods, and its local conservation properties. In particular, it seems to be the most widely used finite element method for the Euler equations [45]. On the other hand, the main drawback of this approach is the large number of unknowns as compared to standard finite element methods. The situation is even worse if one counts the population of the resulting system matrices. In order to find a more efficient approach, it seems therefore important to study the connections with other finite element methods.

In view of the ubiquitous problem of large Péclet numbers, stabilization techniques have been introduced since a long time. They are either based on upwinding or additional terms in the discrete variational formulation. The drawback of the first technique is a loss in consistency which generally leads to large numerical diffusion. The grand-father of the second technique is the SUPG/GLS method [57], [67]. Recently, new approaches have been developed, which try to avoid coupling of the different equations due to the residuals. In this context we cite LPS (local projection stabilization) [62], [55], [48][5] and CIP (continuous interior penalty) [58], [59].

#### 3.3. Finite element methods on quadrilateral and hexahedral meshes

The construction of finite element methods on quadrilateral, and particularly, hexahedral meshes can be a complicated task; especially the development of mixed and non-conforming methods is an active field of research. The difficulties arise not only from the fact that adequate degrees of freedom have to be found, but also from the non-constantness of the element Jacobians; an arbitrary hexahedron, which we define as the image of the unit cube under a tri-linear transformation, does in general not have plane faces, which implies for example, that the normal vector is not constant on a side.

In collaboration with Eric Dubach (Associate professor at LMAP) and Jean-Marie Thomas (Former professor at LMAP) we have built a new class of finite element functions (named pseudo-conforming) on quadrilateral and hexahedral meshes. The degrees of freedom are the same as those of classical iso-parametric finite elements but the basis functions are defined as polynomials on each element of the mesh. On general quadrilaterals and hexahedra, our method leads to a non-conforming method; in the particular case of parallelotopes, the new finite elements coincide with the classical ones [61], [60].

### 3.4. Finite element methods for interface problems

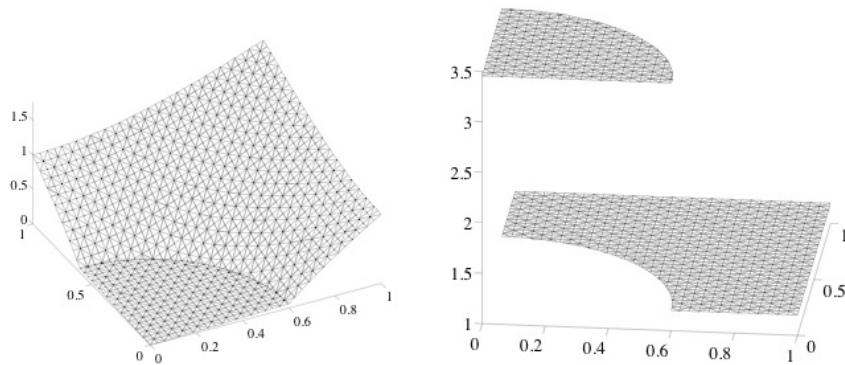


Figure 1. Incompressible elasticity with discontinuous material properties (left: modulus of velocities, right: pressure; from [46]).

The NXFEM (Nitsche eXtended finite element method) has been developed in [63] and [64]. It is based on a pure variational formulation with standard finite element spaces, which are locally enriched in such a way that the accurate capturing of an interface not aligned with the underlying mesh is possible, giving a rigorous formulation of the very popular XFEM. A typical computation for the Stokes problem with varying, piecewise constant viscosity is shown in Figure 1. This technology opens the door to many applications in the field of fluid mechanics, such as immiscible flows, free surface flows and so on.

### 3.5. Adaptivity

Adaptive finite element methods are becoming a standard tool in numerical simulations, and their application in CFD is one of the main topics of Concha. Such methods are based on a posteriori error estimates of the discretization error avoiding explicit knowledge of properties of the solution, in contrast to a priori error estimates. The estimator is used in an adaptive loop by means of a local mesh refinement algorithm. The mathematical theory of these algorithms has for a long time been bounded to the proof of upper and lower bounds, but has made important improvements in recent years. For illustration, a typical sequence of adaptively refined meshes on an  $L$ -shaped domain is shown in Figure 2.

The theoretical analysis of mesh-adaptive methods, even in the most standard case of the Poisson problem, is in its infancy. The first important results in this direction concern the convergence of the sequence of solution generated by the algorithm (the standard a priori error analysis does not apply since the global mesh-size does not necessarily go to zero). In order to prove convergence, an unavoidable data approximation term has to be treated in addition to the error estimator [69]. These results do not say anything about the convergence speed, that is the number of unknowns required to achieve a given accuracy. Such complexity estimates are the subject of active research, the first fundamental result in this direction is [54].



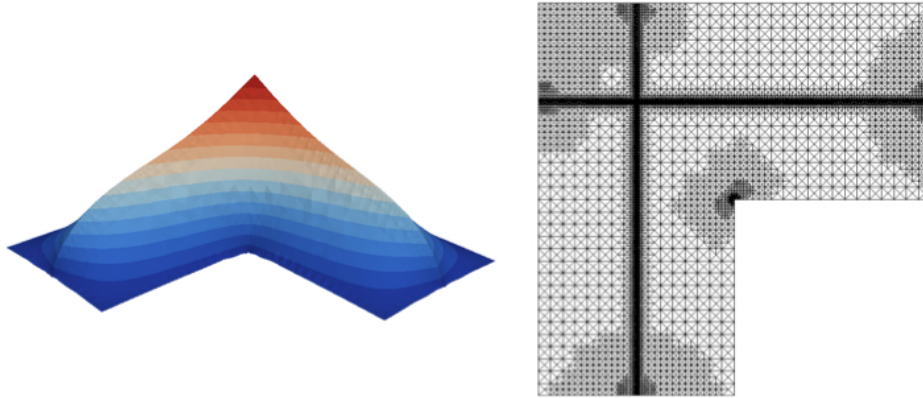


Figure 2. Solution with rough right-hand-side in a corner domain and adaptively refined mesh (from [50]).

Our first contribution [23] to this field has been the introduction of a new adaptive algorithm which makes use of an adaptive marking strategy, which refines according to the data oscillations only if they are by a certain factor larger than the estimator. This algorithm allowed us to prove geometric convergence and quasi-optimal complexity, avoiding additional iteration as used before [71]. We have extended our results to conforming FE without inner node refinement [51] and to mixed FE [50]. In this case, a major additional difficulty arises from the fact that, due to the saddle-point formulation, the orthogonality relation known from continuous FEM does not hold. In addition, we have considered the case of incomplete solution of the discrete systems. To this end, we have developed a simple adaptive stopping criterion based on comparison of the iteration error with the discretization error estimator, see also [49].

Goal-oriented error estimation has been introduced in [52]. It allows to error control and adaptivity directly oriented to the computation of physical quantities, such as the drag and lift coefficient, the Nusselt number, and other physical quantities.

## **DEFI Project-Team**

### **3. Scientific Foundations**

#### **3.1. Scientific Foundations**

The research activity of our team is dedicated to the design, analysis and implementation of efficient numerical methods to solve inverse and shape/topological optimization problems in connection with wave imaging, structural design, non-destructive testing and medical imaging modalities. We are particularly interested in the development of fast methods that are suited for real-time applications and/or large scale problems. These goals require to work on both the physical and the mathematical models involved and indeed a solid expertise in related numerical algorithms.

This section intends to give a general overview of our research interests and themes. We choose to present them through the specific academic example of inverse scattering problems (from inhomogeneities), which is representative of foreseen developments on both inversion and (topological) optimization methods. The practical problem would be to identify an inclusion from measurements of diffracted waves that result from the interaction of the sought inclusion with some (incident) waves sent into the probed medium. Typical applications include biomedical imaging where using micro-waves one would like to probe the presence of pathological cells, or imaging of urban infrastructures where using ground penetrating radars (GPR) one is interested in finding the location of buried facilities such as pipelines or waste deposits. This kind of applications requires in particular fast and reliable algorithms.

By “imaging” we shall refer to the inverse problem where the concern is only the location and the shape of the inclusion, while “identification” may also indicate getting informations on the inclusion physical parameters.

Both problems (imaging and identification) are non linear and ill-posed (lack of stability with respect to measurements errors if some careful constrains are not added). Moreover, the unique determination of the geometry or the coefficients is not guaranteed in general if sufficient measurements are not available. As an example, in the case of anisotropic inclusions, one can show that an appropriate set of data uniquely determine the geometry but not the material properties.

These theoretical considerations (uniqueness, stability) are not only important in understanding the mathematical properties of the inverse problem, but also guide the choice of appropriate numerical strategies (which information can be stably reconstructed) and also the design of appropriate regularization techniques. Moreover, uniqueness proofs are in general constructive proofs, i.e. they implicitly contain a numerical algorithm to solve the inverse problem, hence their importance for practical applications. The sampling methods introduced below are one example of such algorithms.

A large part of our research activity is dedicated to numerical methods applied to the first type of inverse problems, where only the geometrical information is sought. In its general setting the inverse problem is very challenging and no method can provide a universal satisfactory solution to it (regarding the balance cost-precision-stability). This is why in the majority of the practically employed algorithms, some simplification of the underlying mathematical model is used, according to the specific configuration of the imaging experiment. The most popular ones are geometric optics (the Kirchhoff approximation) for high frequencies and weak scattering (the Born approximation) for small contrasts or small obstacles. They actually give full satisfaction for a wide range of applications as attested by the large success of existing imaging devices (radar, sonar, echography, X-ray tomography, etc.), that rely on one of these approximations.

Generally speaking, the used simplifications result in a linearization of the inverse problem and therefore are usually valid only if the latter is weakly non-linear. The development of these simplified models and the improvement of their efficiency is still a very active research area. With that perspective we are particularly interested in deriving and studying higher order asymptotic models associated with small geometrical parameters such as: small obstacles, thin coatings, wires, periodic media, .... Higher order models usually introduce some non linearity in the inverse problem, but are in principle easier to handle from the numerical point of view than in the case of the exact model.

A larger part of our research activity is dedicated to algorithms that avoid the use of such approximations and that are efficient where classical approaches fail: i.e. roughly speaking when the non linearity of the inverse problem is sufficiently strong. This type of configuration is motivated by the applications mentioned below, and occurs as soon as the geometry of the unknown media generates non negligible multiple scattering effects (multiply-connected and closely spaced obstacles) or when the used frequency is in the so-called resonant region (wave-length comparable to the size of the sought medium). It is therefore much more difficult to deal with and requires new approaches. Our ideas to tackle this problem will be motivated and inspired by recent advances in shape and topological optimization methods and also the introduction of novel classes of imaging algorithms, so-called sampling methods.

The sampling methods are fast imaging solvers adapted to multi-static data (multiple receiver-transmitter pairs) at a fixed frequency. Even if they do not use any linearization of the forward model, they rely on computing the solutions to a set of linear problems of small size, that can be performed in a completely parallel procedure. Our team has already a solid expertise in these methods applied to electromagnetic 3-D problems. The success of such approaches was their ability to provide a relatively quick algorithm for solving 3-D problems without any need for a priori knowledge on the physical parameters of the targets. These algorithms solve only the imaging problem, in the sense that only the geometrical information is provided.

Despite the large efforts already spent in the development of this type of methods, either from the algorithmic point of view or the theoretical one, numerous questions are still open. These attractive new algorithms also suffer from the lack of experimental validations, due to their relatively recent introduction. We also would like to invest on this side by developing collaborations with engineering research groups that have experimental facilities. From the practical point of view, the most potential limitation of sampling methods would be the need of a large amount of data to achieve a reasonable accuracy. On the other hand, optimization methods do not suffer from this constraint but they require good initial guess to ensure convergence and reduce the number of iterations. Therefore it seems natural to try to combine the two class of methods in order to calibrate the balance between cost and precision.

Among various shape optimization methods, the Level Set method seems to be particularly suited for such a coupling. First, because it shares similar mechanism as sampling methods: the geometry is captured as a level set of an “indicator function” computed on a cartesian grid. Second, because the two methods do not require any a priori knowledge on the topology of the sought geometry. Beyond the choice of a particular method, the main question would be to define in which way the coupling can be achieved. Obvious strategies consist in using one method to pre-process (initialization) or post-process (find the level set) the other. But one can also think of more elaborate ones, where for instance a sampling method can be used to optimize the choice of the incident wave at each iteration step. The latter point is closely related to the design of so called “focusing incident waves” (which are for instance the basis of applications of the time-reversal principle). In the frequency regime, these incident waves can be constructed from the eigenvalue decomposition of the data operator used by sampling methods. The theoretical and numerical investigations of these aspects are still not completely understood for electromagnetic or elastodynamic problems.

Other topological optimization methods, like the homogenization method or the topological gradient method, can also be used, each one provides particular advantages in specific configurations. It is evident that the development of these methods is very suited to inverse problems and provide substantial advantage compared to classical shape optimization methods based on boundary variation. Their applications to inverse problems has not been fully investigated. The efficiency of these optimization methods can also be increased for adequate asymptotic configurations. For instance small amplitude homogenization method can be used as an efficient relaxation method for the inverse problem in the presence of small contrasts. On the other hand, the topological gradient method has shown to perform well in localizing small inclusions with only one iteration.

A broader perspective would be the extension of the above mentioned techniques to time-dependent cases. Taking into account data in time domain is important for many practical applications, such as imaging in cluttered media, the design of absorbing coatings or also crash worthiness in the case of structural design.

For the identification problem, one would like to also have information of the physical properties of the targets. Of course optimization methods is a tool of choice for these problems. However, in some applications only

a qualitative information is needed and obtaining it in a cheaper way can be performed using asymptotic theories combined with sampling methods. We also refer here to the use of so called transmission eigenvalues as qualitative indicators for non destructive testing of dielectrics.

We are also interested in parameter identification problems arising in diffusion-type problems. Our research here is mostly motivated by applications to the imaging of biological tissues with the technique of Diffusion Magnetic Resonance Imaging (DMRI). Roughly speaking DMRI gives a measure of the average distance travelled by water molecules in a certain medium and can give useful information on cellular structure and structural change when the medium is biological tissue. In particular, we would like to infer from DMRI measurements changes in the cellular volume fraction occurring upon various physiological or pathological conditions as well as the average cell size in the case of tumor imaging. The main challenges here are 1) correctly model measured signals using diffusive-type time-dependent PDEs 2) numerically handle the complexity of the tissues 3) use the first two to identify physically relevant parameters from measurements. For the last point we are particularly interested in constructing reduced models of the multiple-compartment Bloch-Torrey partial differential equation using homogenization methods.

**GAMMA3 Project-Team (section vide)**

## IPSO Project-Team

### 3. Scientific Foundations

#### 3.1. Structure-preserving numerical schemes for solving ordinary differential equations

**Participants:** François Castella, Philippe Chartier, Erwan Faou, Vilmart Gilles.

ordinary differential equation, numerical integrator, invariant, Hamiltonian system, reversible system, Lie-group system

In many physical situations, the time-evolution of certain quantities may be written as a Cauchy problem for a differential equation of the form

$$\begin{aligned} y'(t) &= f(y(t)), \\ y(0) &= y_0. \end{aligned} \tag{2}$$

For a given  $y_0$ , the solution  $y(t)$  at time  $t$  is denoted  $\varphi_t(y_0)$ . For fixed  $t$ ,  $\varphi_t$  becomes a function of  $y_0$  called the *flow* of (1). From this point of view, a numerical scheme with step size  $h$  for solving (1) may be regarded as an approximation  $\Phi_h$  of  $\varphi_h$ . One of the main questions of *geometric integration* is whether *intrinsic* properties of  $\varphi_t$  may be passed on to  $\Phi_h$ .

This question can be more specifically addressed in the following situations:

##### 3.1.1. Reversible ODEs

The system (1) is said to be  $\rho$ -reversible if there exists an involutive linear map  $\rho$  such that

$$\rho \circ \varphi_t = \varphi_t^{-1} \circ \rho = \varphi_{-t} \circ \rho. \tag{3}$$

It is then natural to require that  $\Phi_h$  satisfies the same relation. If this is so,  $\Phi_h$  is said to be *symmetric*. Symmetric methods for reversible systems of ODEs are just as much important as *symplectic* methods for Hamiltonian systems and offer an interesting alternative to symplectic methods.

##### 3.1.2. ODEs with an invariant manifold

The system (1) is said to have an invariant manifold  $g$  whenever

$$\mathcal{M} = \{y \in \mathbb{R}^n; g(y) = 0\} \tag{4}$$

is kept *globally* invariant by  $\varphi_t$ . In terms of derivatives and for sufficiently differentiable functions  $f$  and  $g$ , this means that

$$\forall y \in \mathcal{M}, g'(y)f(y) = 0.$$

As an example, we mention Lie-group equations, for which the manifold has an additional group structure. This could possibly be exploited for the space-discretisation. Numerical methods amenable to this sort of problems have been reviewed in a recent paper [56] and divided into two classes, according to whether they use  $g$  explicitly or through a projection step. In both cases, the numerical solution is forced to live on the manifold at the expense of some Newton's iterations.

### 3.1.3. Hamiltonian systems

Hamiltonian problems are ordinary differential equations of the form:

$$\begin{aligned}\dot{p}(t) &= -\nabla_q H(p(t), q(t)) \in \mathbb{R}^d \\ \dot{q}(t) &= \nabla_p H(p(t), q(t)) \in \mathbb{R}^d\end{aligned}\quad (5)$$

with some prescribed initial values  $(p(0), q(0)) = (p_0, q_0)$  and for some scalar function  $H$ , called the Hamiltonian. In this situation,  $H$  is an invariant of the problem. The evolution equation (4) can thus be regarded as a differential equation on the manifold

$$\mathcal{M} = \{(p, q) \in \mathbb{R}^d \times \mathbb{R}^d; H(p, q) = H(p_0, q_0)\}.$$

Besides the Hamiltonian function, there might exist other invariants for such systems: when there exist  $d$  invariants in involution, the system (4) is said to be *integrable*. Consider now the parallelogram  $P$  originating from the point  $(p, q) \in \mathbb{R}^{2d}$  and spanned by the two vectors  $\xi \in \mathbb{R}^{2d}$  and  $\eta \in \mathbb{R}^{2d}$ , and let  $\omega(\xi, \eta)$  be the sum of the *oriented* areas of the projections over the planes  $(p_i, q_i)$  of  $P$ ,

$$\omega(\xi, \eta) = \xi^T J \eta,$$

where  $J$  is the *canonical symplectic* matrix

$$J = \begin{bmatrix} 0 & I_d \\ -I_d & 0 \end{bmatrix}.$$

A continuously differentiable map  $g$  from  $\mathbb{R}^{2d}$  to itself is called *symplectic* if it preserves  $\omega$ , i.e. if

$$\omega(g'(p, q)\xi, g'(p, q)\eta) = \omega(\xi, \eta).$$

A fundamental property of Hamiltonian systems is that their exact flow is symplectic. Integrable Hamiltonian systems behave in a very remarkable way: as a matter of fact, their invariants persist under small perturbations, as shown in the celebrated theory of Kolmogorov, Arnold and Moser. This behavior motivates the introduction of *symplectic* numerical flows that share most of the properties of the exact flow. For practical simulations of Hamiltonian systems, symplectic methods possess an important advantage: the error-growth as a function of time is indeed linear, whereas it would typically be quadratic for non-symplectic methods.

### 3.1.4. Differential-algebraic equations

Whenever the number of differential equations is insufficient to determine the solution of the system, it may become necessary to solve the differential part and the constraint part altogether. Systems of this sort are called *differential-algebraic* systems. They can be classified according to their index, yet for the purpose of this expository section, it is enough to present the so-called index-2 systems

$$\begin{aligned}\dot{y}(t) &= f(y(t), z(t)), \\ 0 &= g(y(t)),\end{aligned}\quad (6)$$

where initial values  $(y(0), z(0)) = (y_0, z_0)$  are given and assumed to be consistent with the constraint manifold. By constraint manifold, we imply the intersection of the manifold

$$\mathcal{M}_1 = \{y \in \mathbb{R}^n, g(y) = 0\}$$

and of the so-called hidden manifold

$$\mathcal{M}_2 = \{(y, z) \in \mathbb{R}^n \times \mathbb{R}^m, \frac{\partial g}{\partial y}(y)f(y, z) = 0\}.$$

This manifold  $\mathcal{M} = \mathcal{M}_1 \cap \mathcal{M}_2$  is the manifold on which the exact solution  $(y(t), z(t))$  of (5) lives.

There exists a whole set of schemes which provide a numerical approximation lying on  $\mathcal{M}_1$ . Furthermore, this solution can be projected on the manifold  $\mathcal{M}$  by standard projection techniques. However, it is worth mentioning that a projection destroys the symmetry of the underlying scheme, so that the construction of a symmetric numerical scheme preserving  $\mathcal{M}$  requires a more sophisticated approach.

### 3.2. Highly-oscillatory systems

**Participants:** François Castella, Philippe Chartier, Nicolas Crouseilles, Erwan Faou, Florian Méhats, Mohammed Lemou, Gilles Vilmart.

second-order ODEs, oscillatory solutions, Schrödinger and wave equations, step size restrictions.

In applications to molecular dynamics or quantum dynamics for instance, the right-hand side of (1) involves *fast* forces (short-range interactions) and *slow* forces (long-range interactions). Since *fast* forces are much cheaper to evaluate than *slow* forces, it seems highly desirable to design numerical methods for which the number of evaluations of slow forces is not (at least not too much) affected by the presence of fast forces.

A typical model of highly-oscillatory systems is the second-order differential equations

$$\ddot{q} = -\nabla V(q) \tag{7}$$

where the potential  $V(q)$  is a sum of potentials  $V = W + U$  acting on different time-scales, with  $\nabla^2 W$  positive definite and  $\|\nabla^2 W\| \gg \|\nabla^2 U\|$ . In order to get a bounded error propagation in the linearized equations for an explicit numerical method, the step size must be restricted according to

$$h\omega < C,$$

where  $C$  is a constant depending on the numerical method and where  $\omega$  is the highest frequency of the problem, i.e. in this situation the square root of the largest eigenvalue of  $\nabla^2 W$ . In applications to molecular dynamics for instance, *fast* forces deriving from  $W$  (short-range interactions) are much cheaper to evaluate than *slow* forces deriving from  $U$  (long-range interactions). In this case, it thus seems highly desirable to design numerical methods for which the number of evaluations of slow forces is not (at least not too much) affected by the presence of fast forces.

Another prominent example of highly-oscillatory systems is encountered in quantum dynamics where the Schrödinger equation is the model to be used. Assuming that the Laplacian has been discretized in space, one indeed gets the *time*-dependent Schrödinger equation:

$$i\dot{\psi}(t) = \frac{1}{\varepsilon} H(t)\psi(t), \tag{8}$$



where  $H(t)$  is finite-dimensional matrix and where  $\varepsilon$  typically is the square-root of a mass-ratio (say electron/ion for instance) and is small ( $\varepsilon \approx 10^{-2}$  or smaller). Through the coupling with classical mechanics ( $H(t)$  is obtained by solving some equations from classical mechanics), we are faced once again with two different time-scales, 1 and  $\varepsilon$ . In this situation also, it is thus desirable to devise a numerical method able to advance the solution by a time-step  $h > \varepsilon$ .

### 3.3. Geometric schemes for the Schrödinger equation

**Participants:** François Castella, Philippe Chartier, Erwan Faou, Florian Méhats, Gilles Vilmart.

Schrödinger equation, variational splitting, energy conservation.

Given the Hamiltonian structure of the Schrödinger equation, we are led to consider the question of energy preservation for time-discretization schemes.

At a higher level, the Schrödinger equation is a partial differential equation which may exhibit Hamiltonian structures. This is the case of the time-dependent Schrödinger equation, which we may write as

$$i\varepsilon \frac{\partial \psi}{\partial t} = H\psi, \quad (9)$$

where  $\psi = \psi(x, t)$  is the wave function depending on the spatial variables  $x = (x_1, \dots, x_N)$  with  $x_k \in \mathbb{R}^d$  (e.g., with  $d = 1$  or  $3$  in the partition) and the time  $t \in \mathbb{R}$ . Here,  $\varepsilon$  is a (small) positive number representing the scaled Planck constant and  $i$  is the complex imaginary unit. The Hamiltonian operator  $H$  is written

$$H = T + V$$

with the kinetic and potential energy operators

$$T = - \sum_{k=1}^N \frac{\varepsilon^2}{2m_k} \Delta_{x_k} \quad \text{and} \quad V = V(x),$$

where  $m_k > 0$  is a particle mass and  $\Delta_{x_k}$  the Laplacian in the variable  $x_k \in \mathbb{R}^d$ , and where the real-valued potential  $V$  acts as a multiplication operator on  $\psi$ .

The multiplication by  $i$  in (8) plays the role of the multiplication by  $J$  in classical mechanics, and the energy  $\langle \psi | H | \psi \rangle$  is conserved along the solution of (8), using the physicists' notations  $\langle u | A | u \rangle = \langle u, Au \rangle$  where  $\langle \cdot, \cdot \rangle$  denotes the Hermitian  $L^2$ -product over the phase space. In quantum mechanics, the number  $N$  of particles is very large making the direct approximation of (8) very difficult.

The numerical approximation of (8) can be obtained using projections onto submanifolds of the phase space, leading to various PDEs or ODEs: see [60], [59] for reviews. However the long-time behavior of these approximated solutions is well understood only in this latter case, where the dynamics turns out to be finite dimensional. In the general case, it is very difficult to prove the preservation of qualitative properties of (8) such as energy conservation or growth in time of Sobolev norms. The reason for this is that backward error analysis is not directly applicable for PDEs. Overwhelming these difficulties is thus a very interesting challenge.

A particularly interesting case of study is given by symmetric splitting methods, such as the Strang splitting:

$$\psi_1 = \exp(-i(\delta t)V/2) \exp(i(\delta t)\Delta) \exp(-i(\delta t)V/2)\psi_0 \quad (10)$$

where  $\delta t$  is the time increment (we have set all the parameters to 1 in the equation). As the Laplace operator is unbounded, we cannot apply the standard methods used in ODEs to derive long-time properties of these schemes. However, its projection onto finite dimensional submanifolds (such as Gaussian wave packets space or FEM finite dimensional space of functions in  $x$ ) may exhibit Hamiltonian or Poisson structure, whose long-time properties turn out to be more tractable.

### 3.4. High-frequency limit of the Helmholtz equation

**Participant:** François Castella.

waves, Helmholtz equation, high oscillations.

The Helmholtz equation models the propagation of waves in a medium with variable refraction index. It is a simplified version of the Maxwell system for electro-magnetic waves.

The high-frequency regime is characterized by the fact that the typical wavelength of the signals under consideration is much smaller than the typical distance of observation of those signals. Hence, in the high-frequency regime, the Helmholtz equation at once involves highly oscillatory phenomena that are to be described in some asymptotic way. Quantitatively, the Helmholtz equation reads

$$i\alpha_\varepsilon u_\varepsilon(x) + \varepsilon^2 \Delta_x u_\varepsilon + n^2(x)u_\varepsilon = f_\varepsilon(x). \quad (11)$$

Here,  $\varepsilon$  is the small adimensional parameter that measures the typical wavelength of the signal,  $n(x)$  is the space-dependent refraction index, and  $f_\varepsilon(x)$  is a given (possibly dependent on  $\varepsilon$ ) source term. The unknown is  $u_\varepsilon(x)$ . One may think of an antenna emitting waves in the whole space (this is the  $f_\varepsilon(x)$ ), thus creating at any point  $x$  the signal  $u_\varepsilon(x)$  along the propagation. The small  $\alpha_\varepsilon > 0$  term takes into account damping of the waves as they propagate.

One important scientific objective typically is to describe the high-frequency regime in terms of *rays* propagating in the medium, that are possibly refracted at interfaces, or bounce on boundaries, etc. Ultimately, one would like to replace the true numerical resolution of the Helmholtz equation by that of a simpler, asymptotic model, formulated in terms of rays.

In some sense, and in comparison with, say, the wave equation, the specificity of the Helmholtz equation is the following. While the wave equation typically describes the evolution of waves between some initial time and some given observation time, the Helmholtz equation takes into account at once the propagation of waves over *infinitely long* time intervals. Qualitatively, in order to have a good understanding of the signal observed in some bounded region of space, one readily needs to be able to describe the propagative phenomena in the whole space, up to infinity. In other words, the “rays” we refer to above need to be understood from the initial time up to infinity. This is a central difficulty in the analysis of the high-frequency behaviour of the Helmholtz equation.

### 3.5. From the Schrödinger equation to Boltzmann-like equations

**Participant:** François Castella.

Schrödinger equation, asymptotic model, Boltzmann equation.

The Schrödinger equation is the appropriate way to describe transport phenomena at the scale of electrons. However, for real devices, it is important to derive models valid at a larger scale.

In semi-conductors, the Schrödinger equation is the ultimate model that allows to obtain quantitative information about electronic transport in crystals. It reads, in convenient adimensional units,

$$i\partial_t \psi(t, x) = -\frac{1}{2} \Delta_x \psi + V(x)\psi, \quad (12)$$

where  $V(x)$  is the potential and  $\psi(t, x)$  is the time- and space-dependent wave function. However, the size of real devices makes it important to derive simplified models that are valid at a larger scale. Typically, one wishes to have kinetic transport equations. As is well-known, this requirement needs one to be able to describe “collisions” between electrons in these devices, a concept that makes sense at the macroscopic level, while it does not at the microscopic (electronic) level. Quantitatively, the question is the following: can one obtain the Boltzmann equation (an equation that describes collisional phenomena) as an asymptotic model for the Schrödinger equation, along the physically relevant micro-macro asymptotics? From the point of view of modelling, one wishes here to understand what are the “good objects”, or, in more technical words, what are the relevant “cross-sections”, that describe the elementary collisional phenomena. Quantitatively, the Boltzmann equation reads, in a simplified, linearized, form :

$$\partial_t f(t, x, v) = \int_{\mathbf{R}^3} \sigma(v, v') [f(t, x, v') - f(t, x, v)] dv'. \quad (13)$$

Here, the unknown is  $f(x, v, t)$ , the probability that a particle sits at position  $x$ , with a velocity  $v$ , at time  $t$ . Also,  $\sigma(v, v')$  is called the cross-section, and it describes the probability that a particle “jumps” from velocity  $v$  to velocity  $v'$  (or the converse) after a collision process.

## MC2 Project-Team

### 3. Scientific Foundations

#### 3.1. Introduction

We are mainly concerned with complex fluid mechanics problems. The complexity consists of the rheological nature of the fluids (non newtonian fluids), of the coupling phenomena (in shape optimization problems), of the geometry (micro-channels) or of multi-scale phenomena arising in turbulence or in tumor growth modeling. Our goal is to understand these phenomena and to simulate and/or to control them. The subject is wide and we will restrict ourselves to three directions: the first one consists in studying low Reynolds number interface problems in multi-fluid flows with applications to complex fluids, microfluidics and biology - the second one deals with numerical simulation of Newtonian fluid flows with emphasis on the coupling of methods to obtain fast solvers.

Even if we deal with several kinds of applications, there is a strong scientific core at each level of our project. Concerning the model, we are mainly concerned with incompressible flows and we work with the classical description of incompressible fluid dynamics. For the numerical methods, we use the penalization method to describe the obstacles or the boundary conditions for high Reynolds flows, for shape optimization, for interface problems in biology or in microfluidics. This allows us to use only cartesian meshes. Moreover, we use the level-set method for interface problems, for shape optimization and for fluid structure interaction. Finally, for the implementation, strong interaction exists between the members of the team and the modules of the numerical codes are used by all the team and we want to build the platform **eLYSe** to systematize this approach.

#### 3.2. Multi-fluid flows and application for complex fluids, microfluidics

**Participants:** Angelo Iollo, Charles-Henri Bruneau, Thierry Colin, Mathieu Colin, Kévin Santugini.

Multi-fluid flows, microfluidics

By a complex fluid, we mean a fluid containing some mesoscopic objects, *i.e.* structures whose size is intermediate between the microscopic size and the macroscopic size of the experiment. The aim is to study complex fluids containing surfactants in large quantities. It modifies the viscosity properties of the fluids and surface-tension phenomena can become predominant.

Microfluidics is the study of fluids in very small quantities, in micro-channels (a micro-channel is typically 1 cm long with a section of  $50\mu\text{m} \times 50\mu\text{m}$ ). They are many advantages of using such channels. First, one needs only a small quantity of liquid to analyze the phenomena. Furthermore, very stable flows and quite unusual regimes may be observed, which enables to perform more accurate measurements. The idea is to couple numerical simulations with experiments to understand the phenomena, to predict the flows and compute some quantities like viscosity coefficients for example. Flows in micro-channels are often at low Reynolds numbers. The hydrodynamical part is therefore stable. However, the main problem is to produce real 3D simulations covering a large range of situations. For example we want to describe diphasic flows with surface tension and sometimes surface viscosity. Surface tension enforces the stability of the flow. The size of the channel implies that one can observe some very stable phenomena. For example, using a "T" junction, a very stable interface between two fluids can be observed. In a cross junction, one can also have formation of droplets that travel along the channel. Some numerical difficulties arise from the surface tension term. With an explicit discretization of this term, a restrictive stability condition appears for very slow flows [71]. Our partner is the LOF, a Rhodia-Bordeaux 1-CNRS laboratory.

One of the main points is the wetting phenomena at the boundary. Note that the boundary conditions are fundamental for the description of the flow since the channels are very shallow. The wetting properties cannot be neglected at all. Indeed, for the case of a two non-miscible fluids system, if one considers no-slip boundary conditions, then since the interface is driven by the velocity of the fluids, it shall not move on the boundary. The experiments shows that this is not true: the interface is moving and in fact all the dynamics start from the boundary and then propagate in the whole volume of fluids. Even with low Reynolds numbers, the wetting effects can induce instabilities and are responsible of hardly predictable flows. Moreover, the fluids that are used are often visco-elastic and exhibit "unusual" slip length. Therefore, we cannot use standard numerical codes and have to adapt the usual numerical methods to our case to take into account the specificities of our situations. Moreover, we want to obtain reliable models and simulations that can be as simple as possible and that can be used by our collaborators. As a summary, the main specific points of the physics are: the multi-fluid simulations at low Reynolds number, the wetting problems and the surface tension that are crucial, the 3D characteristic of the flows, the boundary conditions that are fundamental due to the size of the channels. We need to handle complex fluids. Our collaborators in this lab are J.-B. Salmon, P. Guillot, A. Colin.

The evolution of non-newtonian flows in webs of micro-channels are therefore useful to understand the mixing of oil, water and polymer for enhanced oil recovery for example. Complex fluids arising in cosmetics are also of interest. We also need to handle mixing processes.

### 3.3. Cancer modeling

**Participants:** Angelo Iollo, Thierry Colin, Clair Poignard, Olivier Saut, Lisl Weynans.

Tumor growth, cancer, metastasis

As in microfluidics, the growth of a tumor is a low Reynolds number flow. Several kinds of interfaces are present (membranes, several populations of cells,...) The biological nature of the tissues impose the use of different models in order to describe the evolution of tumor growth. The complexity of the geometry, of the rheological properties and the coupling with multi-scale phenomena is high but not far away from those encountered in microfluidics and the models and methods are close.

The challenge is twofold. On one hand, we wish to understand the complexity of the coupling effects between the different levels (cellular, genetic, organs, membranes, molecular). Trying to be exhaustive is of course hopeless, however it is possible numerically to isolate some parts of the evolution in order to better understand the interactions. Another strategy is to test *in silico* some therapeutic innovations. An example of such a test is given in [81] where the efficacy of radiotherapy is studied and in [82] where the effects of anti-invasive agents is investigated. It is therefore useful to model a tumor growth at several stage of evolution. The macroscopic continuous model is based on Darcy's law which seems to be a good approximation to describe the flow of the tumor cells in the extra-cellular matrix [50], [72], [73]. It is therefore possible to develop a two-dimensional model for the evolution of the cell densities. We formulate mathematically the evolution of the cell densities in the tissue as advection equations for a set of unknowns representing the density of cells with position  $(x, y)$  at time  $t$  in a given cycle phase. Assuming that all cells move with the same velocity given by Darcy's law and applying the principle of mass balance, one obtains the advection equations with a source term given by a cellular automaton. We assume diffusion for the oxygen and the diffusion constant depends on the density of the cells. The source of oxygen corresponds to the spatial location of blood vessels. The available quantities of oxygen interact with the proliferation rate given by the cellular automaton [81].

A forthcoming investigation in cancer treatment simulation is the influence of the electrochemotherapy [76] on the tumor growth. Electrochemotherapy consists in imposing to the malignant tumor high voltage electric pulses so that the plasma membrane of carcinoma cells is permeabilized. Biologically active molecules such as bleomycin, which usually cannot diffuse through the membrane, may then be internalized. A work in progress (C.Poignard [80] in collaboration with the CNRS lab of physical vectorology at the Institut Gustave Roussy) consists in modelling electromagnetic phenomena at the cell scale. A coupling between the microscopic description of the electroporation of cells and its influence on the global tumor growth at the macroscopic scale is expected. Another key point is the parametrization of the models in order to produce image-based simulations.

The second challenge is more ambitious. Mathematical models of cancer have been extensively developed with the aim of understanding and predicting tumor growth and the effects of treatments. In vivo modeling of tumors is limited by the amount of information available. However, in the last few years there have been dramatic increases in the range and quality of information available from non-invasive imaging methods, so that several potentially valuable imaging measurements are now available to quantitatively measure tumor growth, assess tumor status as well as anatomical or functional details. Using different methods such as the CT scan, magnetic resonance imaging (MRI), or positron emission tomography (PET), it is now possible to evaluate and define tumor status at different levels: physiological, molecular and cellular.

In this context, the present project aims at supporting the decision process of oncologists in the definition of therapeutic protocols via quantitative methods. The idea is to build mathematically and physically sound phenomenological models that can lead to patient-specific full-scale simulations, starting from data collected typically through medical imagery like CT scans, MRIs and PET scans or by quantitative molecular biology for leukemia. Our ambition is to provide medical doctors with patient-specific tumor growth models able to estimate, on the basis of previously collected data and within the limits of phenomenological models, the evolution at subsequent times of the pathology and possibly the response to the therapies.

The final goal is to provide numerical tools in order to help to answer to the crucial questions for a clinician:

When to start a treatment?

When to change a treatment?

When to stop a treatment?

Also we intend to incorporate real-time model information for improving the precision and effectiveness of non-invasive or micro-invasive tumor ablation techniques like acoustic hyperthermia, electroporation, radiofrequency or cryo-ablation.

We will specifically focus on the following pathologies: Lung and liver metastasis of a distant tumor

Low grade and high grade gliomas, meningiomas

Chronic myelogenous leukemia

These pathologies have been chosen because of the existing collaborations between the applied mathematics department of University of Bordeaux and the Institut Bergonié.

Our approach. Our approach is deterministic and spatial: it is based on solving an inverse problem based on imaging data. Models are of partial differential equation (PDE) type. They are coupled with a process of data assimilation based on imaging. We already have undertaken test cases on patients that are followed at Bergonié for lung metastases of thyroid tumors. These patients have a slowly evolving, asymptomatic metastatic disease, monitored by CT scans. On two thoracic images relative to successive times, the volume of the tumor under investigation is extracted by segmentation. To test our method, we chose patients without treatment and for whom we had at least three successive.

### 3.4. Newtonian fluid flows simulations and their analysis

**Participants:** Charles-Henri Bruneau, Angelo Iollo, Iraj Mortazavi, Michel Bergmann, Lisl Weynans.

Simulation, Analysis

It is very exciting to model complex phenomena for high Reynolds flows and to develop methods to compute the corresponding approximate solutions, however a well-understanding of the phenomena is necessary. Classical graphic tools give us the possibility to visualize some aspects of the solution at a given time and to even see in some way their evolution. Nevertheless in many situations it is not sufficient to understand the mechanisms that create such a behavior or to find the real properties of the flow. It is then necessary to carefully analyze the flow, for instance the vortex dynamics or to identify the coherent structures to better understand their impact on the whole flow behavior.

The various numerical methods used or developed to approximate the flows depend on the studied phenomenon. Our goal is to compute the most reliable method for each situation.

The first method, which is affordable in 2D, consists in a directly solving of the genuine Navier-Stokes equations in primitive variables (velocity-pressure) on Cartesian domains [59]. The bodies, around which the flow has to be computed are modeled using the penalization method (also named Brinkman-Navier-Stokes equations). This is an immersed boundary method in which the bodies are considered as porous media with a very small intrinsic permeability [51]. This method is very easy to handle as it consists only in adding a mass term  $U/K$  in the momentum equations. The boundary conditions imposed on artificial boundaries of the computational domains avoid any reflections when vortices cross the boundary. To make the approximation efficient enough in terms of CPU time, a multi-grid solver with a cell by cell Gauss-Seidel smoother is used. The second type of methods is the vortex method. It is a Lagrangian technique that has been proposed as an alternative to more conventional grid-based methods. Its main feature is that the inertial nonlinear term in the flow equations is implicitly accounted for by the transport of particles. The method thus avoids to a large extent the classical stability/accuracy dilemma of finite-difference or finite-volume methods. This has been demonstrated in the context of computations for high Reynolds number laminar flows and for turbulent flows at moderate Reynolds numbers [66]. This method has recently enabled us to obtain new results concerning the three-dimensional dynamics of cylinder wakes.

The third method is to develop reduced order models (ROM) based on a Proper Orthogonal Decomposition (POD) [74]. The POD consists in approximating a given flow field  $U(x, t)$  with the decomposition

$$U(x, t) = \sum_i a_i(t) \phi_i(x),$$

where the basis functions are empirical in the sense that they derive from an existing data base given for instance by one of the methods above. Then the approximation of Navier-Stokes equations for instance is reduced to solving a low-order dynamical system that is very cheap in terms of CPU time. Nevertheless the ROM can only reconstitute what is contained in the basis. Our challenge is to extend its application in order to make it an actual prediction tool.

The fourth method is a finite volume method on cartesian grids to simulate compressible Euler or Navier Stokes Flows in complex domains. An immersed boundary-like technique is developed to take into account boundary conditions around the obstacles with order two accuracy.

### 3.5. Flow control and shape optimization

**Participants:** Charles-Henri Bruneau, Angelo Iollo, Iraj Mortazavi, Michel Bergmann.

Flow Control, Shape Optimization

Flow simulations, optimal design and flow control have been developed these last years in order to solve real industrial problems : vortex trapping cavities with CIRA (Centro Italiano Ricerche Aerospaziali), reduction of vortex induced vibrations on deep sea riser pipes with IFP (Institut Français du Pétrole), drag reduction of a ground vehicle with Renault or in-flight icing with Bombardier and Pratt-Wittney are some examples of possible applications of these researches. Presently the recent creation of the competitiveness cluster on aeronautics, space and embedded systems (AESE) based also in Aquitaine provides the ideal environment to extend our applied researches to the local industrial context. There are two main streams: the first need is to produce direct numerical simulations, the second one is to establish reliable optimization procedures.

In the next subsections we will detail the tools we will base our work on, they can be divided into three points: to find the appropriate devices or actions to control the flow; to determine an effective system identification technique based on the trace of the solution on the boundary; to apply shape optimization and system identification tools to the solution of inverse problems found in object imaging and turbomachinery.

#### 3.5.1. Control of flows

There are mainly two approaches: passive (using passive devices on some specific parts that modify the shear forces) or active (adding locally some energy to change the flow) control.

The passive control consists mainly in adding geometrical devices to modify the flow. One idea is to put a porous material between some parts of an obstacle and the flow in order to modify the shear forces in the boundary layer. This approach may pose remarkable difficulties in terms of numerical simulation since it would be necessary, a priori, to solve two models: one for the fluid, one for the porous medium. However, by using the penalization method it becomes a feasible task [55]. This approach has been now used in several contexts and in particular in the frame of a collaboration with IFP to reduce vortex induced vibrations [56]. Another technique we are interested in is to inject minimal amounts of polymers into hydrodynamic flows in order to stabilize the mechanisms which enhance hydrodynamic drag.

The active approach is addressed to conceive, implement and test automatic flow control and optimization aiming mainly at two applications : the control of unsteadiness and the control and optimization of coupled systems. Implementation of such ideas relies on several tools. The common challenges are infinite dimensional systems, Dirichlet boundary control, nonlinear tracking control, nonlinear partial state observation.

The bottom-line to obtain industrially relevant control devices is the energy budget. The energy required by the actuators should be less than the energy savings resulting from the control application. In this sense our research team has gained a certain experience in testing several control strategies with a doctoral thesis (E. Creusé) devoted to increasing the lift on a dihedral plane. Indeed the extension of these techniques to real world problems may reveal itself very delicate and special care will be devoted to implement numerical methods which permit on-line computing of actual practical applications. For instance the method can be successful to reduce the drag forces around a ground vehicle and a coupling with passive control is under consideration to improve the efficiency of each control strategy.

### 3.5.2. System identification

We remark that the problem of deriving an accurate estimation of the velocity field in an unsteady complex flow, starting from a limited number of measurements, is of great importance in many engineering applications. For instance, in the design of a feedback control, a knowledge of the velocity field is a fundamental element in deciding the appropriate actuator reaction to different flow conditions. In other applications it may be necessary or advisable to monitor the flow conditions in regions of space which are difficult to access or where probes cannot be fitted without causing interference problems.

The idea is to exploit ideas similar to those at the basis of the Kalman filter. The starting point is again a Galerkin representation of the velocity field in terms of empirical eigenfunctions. For a given flow, the POD modes can be computed once and for all based on Direct Numerical Simulation (DNS) or on highly resolved experimental velocity fields, such as those obtained by particle image velocimetry. An instantaneous velocity field can thus be reconstructed by estimating the coefficients  $a_i(t)$  of its Galerkin representation. One simple approach to estimate the POD coefficients is to approximate the flow measurements in a least square sense, as in [70].

A similar procedure is also used in the estimation based on gappy POD, see [85] and [89]. However, these approaches encounter difficulties in giving accurate estimations when three-dimensional flows with complicated unsteady patterns are considered, or when a very limited number of sensors is available. Under these conditions, for instance, the least squares approach cited above (LSQ) rapidly becomes ill-conditioned. This simply reflects the fact that more and more different flow configurations correspond to the same set of measurements.

Our challenge is to propose an approach that combines a linear estimation of the coefficients  $a_i(t)$  with an appropriate non-linear low-dimensional flow model, that can be readily implemented for real time applications.

### 3.5.3. Shape optimization and system identification tools applied to inverse problems found in object imaging and turbomachinery

We will consider two different objectives. The first is strictly linked to the level set methods that are developed for microfluidics. The main idea is to combine different technologies that are developed with our team: penalization methods, level sets, an optimization method that regardless of the model equation will be able to



solve inverse or optimization problems in 2D or 3D. For this we have started a project that is detailed in the research program. See also [62] for a preliminary application.

As for shape optimization in aeronautics, the aeroacoustic optimization problem of propeller blades is addressed by means of an inverse problem and its adjoint equations. This problem is divided into three subtasks:

i) formulation of an inverse problem for the design of propeller blades and determination of the design parameters ii) derivation of an aeroacoustic model able to predict noise levels once the blade geometry and the flow field are given iii) development of an optimization procedure in order to minimize the noise emission by controlling the design parameters.

The main challenge in this field is to move from simplified models [75] to actual 3D model. The spirit is to complete the design performed with a simplified tool with a fully three dimensional inverse problem where the load distribution as well as the geometry of the leading edge are those provided by the meridional plane analysis [84]. A 3D code will be based on the compressible Euler equations and an immersed boundary technique over a cartesian mesh. The code will be implicit and parallel, in the same spirit as what was done for the meridional plane. Further development include the extension of the 3D immersed boundary approach to time-dependent phenomena. This step will allow the designer to take into account noise sources that are typical of internal flows. The task will consist in including time dependent forcing on the inlet and/or outlet boundary under the form of Fourier modes and in computing the linearized response of the system. The optimization will then be based on a direct approach, i.e., an approach where the control is the geometry of the boundary. The computation of the gradient is performed by an adjoint method, which will be a simple "byproduct" of the implicit solver. The load distribution as well as the leading edge geometry obtained by the meridional plane approach will be considered as constraints of the optimization, by projection of the gradient on the constraint tangent plane. These challenges will be undertaken in collaboration with Politecnico di Torino and EC Lyon.

## MICMAC Project-Team

### 3. Scientific Foundations

#### 3.1. Scientific Foundations

Quantum Chemistry aims at understanding the properties of matter through the modeling of its behavior at a subatomic scale, where matter is described as an assembly of nuclei and electrons. At this scale, the equation that rules the interactions between these constitutive elements is the Schrödinger equation. It can be considered (except in few special cases notably those involving relativistic phenomena or nuclear reactions) as a universal model for at least three reasons. First it contains all the physical information of the system under consideration so that any of the properties of this system can in theory be deduced from the Schrödinger equation associated to it. Second, the Schrödinger equation does not involve any empirical parameters, except some fundamental constants of Physics (the Planck constant, the mass and charge of the electron, ...); it can thus be written for any kind of molecular system provided its chemical composition, in terms of natures of nuclei and number of electrons, is known. Third, this model enjoys remarkable predictive capabilities, as confirmed by comparisons with a large amount of experimental data of various types. On the other hand, using this high quality model requires working with space and time scales which are both very tiny: the typical size of the electronic cloud of an isolated atom is the Angström ( $10^{-10}$  meters), and the size of the nucleus embedded in it is  $10^{-15}$  meters; the typical vibration period of a molecular bond is the femtosecond ( $10^{-15}$  seconds), and the characteristic relaxation time for an electron is  $10^{-18}$  seconds. Consequently, Quantum Chemistry calculations concern very short time (say  $10^{-12}$  seconds) behaviors of very small size (say  $10^{-27}$  m<sup>3</sup>) systems. The underlying question is therefore whether information on phenomena at these scales is useful in understanding or, better, predicting macroscopic properties of matter. It is certainly not true that *all* macroscopic properties can be simply upscaled from the consideration of the short time behavior of a tiny sample of matter. Many of them derive from ensemble or bulk effects, that are far from being easy to understand and to model. Striking examples are found in solid state materials or biological systems. Cleavage, the ability minerals have to naturally split along crystal surfaces (e.g. mica yields to thin flakes) is an ensemble effect. Protein folding is also an ensemble effect that originates from the presence of the surrounding medium; it is responsible for peculiar properties (e.g. unexpected acidity of some reactive site enhanced by special interactions) upon which vital processes are based. However, it is undoubtedly true that *many* macroscopic phenomena originate from elementary processes which take place at the atomic scale. Let us mention for instance the fact that the elastic constants of a perfect crystal or the color of a chemical compound (which is related to the wavelengths absorbed or emitted during optic transitions between electronic levels) can be evaluated by atomic scale calculations. In the same fashion, the lubricative properties of graphite are essentially due to a phenomenon which can be entirely modeled at the atomic scale. It is therefore reasonable to simulate the behavior of matter at the atomic scale in order to understand what is going on at the macroscopic one. The journey is however a long one. Starting from the basic principles of Quantum Mechanics to model the matter at the subatomic scale, one finally uses statistical mechanics to reach the macroscopic scale. It is often necessary to rely on intermediate steps to deal with phenomena which take place on various *mesoscales*. It may then be possible to couple one description of the system with some others within the so-called *multiscale* models. The sequel indicates how this journey can be completed focusing on the first smallest scales (the subatomic one), rather than on the larger ones. It has already been mentioned that at the subatomic scale, the behavior of nuclei and electrons is governed by the Schrödinger equation, either in its time dependent form or in its time independent form. Let us only mention at this point that

- both equations involve the quantum Hamiltonian of the molecular system under consideration; from a mathematical viewpoint, it is a self-adjoint operator on some Hilbert space; *both* the Hilbert space and the Hamiltonian operator depend on the nature of the system;
- also present into these equations is the wavefunction of the system; it completely describes its state; its  $L^2$  norm is set to one.

The time dependent equation is a first order linear evolution equation, whereas the time-independent equation is a linear eigenvalue equation. For the reader more familiar with numerical analysis than with quantum mechanics, the linear nature of the problems stated above may look auspicious. What makes the numerical simulation of these equations extremely difficult is essentially the huge size of the Hilbert space: indeed, this space is roughly some symmetry-constrained subspace of  $L^2(\mathbb{R}^d)$ , with  $d = 3(M + N)$ ,  $M$  and  $N$  respectively denoting the number of nuclei and the number of electrons the system is made of. The parameter  $d$  is already 39 for a single water molecule and rapidly reaches  $10^6$  for polymers or biological molecules. In addition, a consequence of the universality of the model is that one has to deal at the same time with several energy scales. In molecular systems, the basic elementary interaction between nuclei and electrons (the two-body Coulomb interaction) appears in various complex physical and chemical phenomena whose characteristic energies cover several orders of magnitude: the binding energy of core electrons in heavy atoms is  $10^4$  times as large as a typical covalent bond energy, which is itself around 20 times as large as the energy of a hydrogen bond. High precision or at least controlled error cancellations are thus required to reach chemical accuracy when starting from the Schrödinger equation. Clever approximations of the Schrödinger problems are therefore needed. The main two approximation strategies, namely the Born-Oppenheimer-Hartree-Fock and the Born-Oppenheimer-Kohn-Sham strategies, end up with large systems of coupled *nonlinear* partial differential equations, each of these equations being posed on  $L^2(\mathbb{R}^3)$ . The size of the underlying functional space is thus reduced at the cost of a dramatic increase of the mathematical complexity of the problem: nonlinearity. The mathematical and numerical analysis of the resulting models has been the major concern of the project-team for a long time. In the recent years, while part of the activity still follows this path, the focus has progressively shifted to problems at other scales. Such problems are described in the following sections.

## NACHOS Project-Team

### 3. Scientific Foundations

#### 3.1. High order discretization methods

The applications in computational electromagnetics and computational geoseismics that are considered by the team lead to the numerical simulation of wave propagation in heterogeneous media or/and involve irregularly shaped objects or domains. The underlying wave propagation phenomena can be purely unsteady or they can be periodic (because the imposed source term follows a time harmonic evolution). In this context, the overall objective of the research activities undertaken by the team is to develop numerical methods putting the emphasis on several features:

- Accuracy. The foreseen numerical methods should ideally rely on discretization techniques that best fit to the geometrical characteristics of the problems at hand. For this reason, the team focuses on methods working on unstructured, locally refined, even non-conforming, simplicial meshes. These methods should also be capable to accurately describe the underlying physical phenomena that may involve highly variable space and time scales. With reference to this characteristic, two main strategies are possible: adaptive local refinement/coarsening of the mesh (i.e  $h$ -adaptivity) and adaptive local variation of the interpolation order (i.e  $p$ -adaptivity). Ideally, these two strategies are combined leading to the so-called  $hp$ -adaptive methods.
- Numerical efficiency. The simulation of unsteady problems most often relies on explicit time integration schemes. Such schemes are constrained by a stability criteria linking the space and time discretization parameters that can be very restrictive when the underlying mesh is highly non-uniform (especially for locally refined meshes). For realistic 3D problems, this can represent a severe limitation with regards to the overall computing time. In order to improve this situation, one possible approach consists in resorting to an implicit time scheme in regions of the computational domain where the underlying mesh is refined while an explicit time scheme is applied to the remaining part of the domain. The resulting hybrid explicit-implicit time integration strategy raises several challenging questions concerning both the mathematical analysis (stability and accuracy, especially for what concern numerical dispersion), and the computer implementation on modern high performance systems (data structures, parallel computing aspects). A second, more classical approach is to devise a local time strategy in the context of a fully explicit time integration scheme. Stability and accuracy are still important challenges in this case.

On the other hand, when considering time harmonic wave propagation problems, numerical efficiency is mainly linked to the solution of the system of algebraic equations resulting from the discretization in space of the underlying PDE model. Various strategies exist ranging from the more robust and efficient sparse direct solvers to the more flexible and cheaper (in terms of memory resources) iterative methods. Current trends tend to show that the ideal candidate will be a judicious mix of both approaches by relying on domain decomposition principles.

- Computational efficiency. Realistic 3D wave propagation problems lead to the processing of very large volumes of data. The latter results from two combined parameters: the size of the mesh i.e the number of mesh elements, and the number of degrees of freedom per mesh element which is itself linked to the degree of interpolation and to the number of physical variables (for systems of partial differential equations). Hence, numerical methods must be adapted to the characteristics of modern parallel computing platforms taking into account their hierarchical nature (e.g multiple processors and multiple core systems with complex cache and memory hierarchies). Besides, appropriate parallelization strategies need to be designed that combine SIMD and MIMD programming paradigms. Moreover, maximizing the effective floating point performances will require the design of numerical algorithms that can benefit from the optimized BLAS linear algebra kernels.

The discontinuous Galerkin method (DG) was introduced in 1973 by Reed and Hill to solve the neutron transport equation. From this time to the 90's a review on the DG methods would likely fit into one page. In the meantime, the finite volume approach has been widely adopted by computational fluid dynamics scientists and has now nearly supplanted classical finite difference and finite element methods in solving problems of non-linear convection. The success of the finite volume method is due to its ability to capture discontinuous solutions which may occur when solving non-linear equations or more simply, when convecting discontinuous initial data in the linear case. Let us first remark that DG methods share with finite volumes this property since a first order finite volume scheme can be viewed as a 0th order DG scheme. However a DG method may be also considered as a finite element one where the continuity constraint at an element interface is released. While it keeps almost all the advantages of the finite element method (large spectrum of applications, complex geometries, etc.), the DG method has other nice properties which explain the renewed interest it gains in various domains in scientific computing as witnessed by books or special issues of journals dedicated to this method [43]- [44]- [45]- [48]:

- It is naturally adapted to a high order approximation of the unknown field. Moreover, one may increase the degree of the approximation in the whole mesh as easily as for spectral methods but, with a DG method, this can also be done very locally. In most cases, the approximation relies on a polynomial interpolation method but the DG method also offers the flexibility of applying local approximation strategies that best fit to the intrinsic features of the modeled physical phenomena.
- When the discretization in space is coupled to an explicit time integration method, the DG method leads to a block diagonal mass matrix independently of the form of the local approximation (e.g the type of polynomial interpolation). This is a striking difference with classical, continuous finite element formulations. Moreover, the mass matrix is diagonal if an orthogonal basis is chosen.
- It easily handles complex meshes. The grid may be a classical conforming finite element mesh, a non-conforming one or even a hybrid mesh made of various elements (tetrahedra, prisms, hexahedra, etc.). The DG method has been proven to work well with highly locally refined meshes. This property makes the DG method more suitable to the design of a *hp*-adaptive solution strategy (i.e where the characteristic mesh size  $h$  and the interpolation degree  $p$  changes locally wherever it is needed).
- It is flexible with regards to the choice of the time stepping scheme. One may combine the DG spatial discretization with any global or local explicit time integration scheme, or even implicit, provided the resulting scheme is stable.
- It is naturally adapted to parallel computing. As long as an explicit time integration scheme is used, the DG method is easily parallelized. Moreover, the compact nature of DG discretization schemes is in favor of high computation to communication ratio especially when the interpolation order is increased.

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

For the numerical solution of the time domain Maxwell equations, we have first proposed a non-dissipative high order DGTD (Discontinuous Galerkin Time Domain) method working on unstructured conforming simplicial meshes [14]-[3]. This DG method combines a central numerical flux function for the approximation of the integral term at an interface between two neighboring elements with a second order leap-frog time integration scheme. Moreover, the local approximation of the electromagnetic field relies on a nodal (Lagrange type) polynomial interpolation method. Recent achievements by the team deal with the extension of these methods towards non-conforming meshes and *hp*-adaptivity [12]-[13], their coupling with hybrid explicit/implicit time integration schemes in order to improve their efficiency in the context of locally refined meshes [6]. A high

order DG method has also been proposed for the numerical resolution of the elastodynamic equations modeling the propagation of seismic waves [5]-[11]. For the numerical treatment of the time harmonic Maxwell equations, we have studied similar DG methods [7]-[18] and more recently, HDG (Hybridized Discontinuous Galerkin) methods [22].

### 3.2. Domain decomposition methods

Domain Decomposition (DD) methods are flexible and powerful techniques for the parallel numerical solution of systems of PDEs. As clearly described in [51], they can be used as a process of distributing a computational domain among a set of interconnected processors or, for the coupling of different physical models applied in different regions of a computational domain (together with the numerical methods best adapted to each model) and, finally as a process of subdividing the solution of a large linear system resulting from the discretization of a system of PDEs into smaller problems whose solutions can be used to devise a parallel preconditioner or a parallel solver. In all cases, DD methods (1) rely on a partitioning of the computational domain into subdomains, (2) solve in parallel the local problems using a direct or iterative solver and, (3) call for an iterative procedure to collect the local solutions in order to get the global solution of the original problem. Subdomain solutions are connected by means of suitable transmission conditions at the artificial interfaces between the subdomains. The choice of these transmission conditions greatly influences the convergence rate of the DD method. One generally distinguishes three kinds of DD methods:

- Overlapping methods use a decomposition of the computational domain in overlapping pieces. The so-called Schwarz method belongs to this class. Schwarz initially introduced this method for proving the existence of a solution to a Poisson problem. In the Schwarz method applied to the numerical resolution of elliptic PDEs, the transmission conditions at artificial subdomain boundaries are simple Dirichlet conditions. Depending on the way the solution procedure is performed, the iterative process is called a Schwarz multiplicative method (the subdomains are treated sequentially) or an additive method (the subdomains are treated in parallel).
- Non-overlapping methods are variants of the original Schwarz DD methods with no overlap between neighboring subdomains. In order to ensure convergence of the iterative process in this case, the transmission conditions are not trivial and are generally obtained through a detailed inspection of the mathematical properties of the underlying PDE or system of PDEs.
- Substructuring methods rely on a non-overlapping partition of the computational domain. They assume a separation of the problem unknowns in purely internal unknowns and interface ones. Then, the internal unknowns are eliminated thanks to a Schur complement technique yielding to the formulation of a problem of smaller size whose iterative resolution is generally easier. Nevertheless, each iteration of the interface solver requires the realization of a matrix/vector product with the Schur complement operator which in turn amounts to the concurrent solution of local subproblems.

Schwarz algorithms have enjoyed a second youth over the last decades, as parallel computers became more and more powerful and available. Fundamental convergence results for the classical Schwarz methods were derived for many partial differential equations, and can now be found in several books [51]- [50]- [53].

The research activities of the team on this topic aim at the formulation, analysis and evaluation of Schwarz type domain decomposition methods in conjunction with discontinuous Galerkin approximation methods on unstructured simplicial meshes for the solution of time domain and time harmonic wave propagation problems. Ongoing works in this direction are concerned with the design of non-overlapping Schwarz algorithms for the solution of the time harmonic Maxwell equations. A first achievement has been a Schwarz algorithm for the time harmonic Maxwell equations, where a first order absorbing condition is imposed at the interfaces between neighboring subdomains [9]. This interface condition is equivalent to a Dirichlet condition for characteristic variables associated to incoming waves. For this reason, it is often referred as a natural interface condition. Beside Schwarz algorithms based on natural interface conditions, the team also investigates algorithms that make use of more effective transmission conditions [10]. Recent contributions are concerned with the design and analysis of such optimized Schwarz algorithm for the solution of the time harmonic Maxwell equations with non-zero conductivity [17].

### **3.3. High performance numerical computing**

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

## NANO-D Team

### 3. Scientific Foundations

#### 3.1. Overview

The adaptive simulation algorithms we develop typically consist in two main components. The first one determines *which degrees of freedom are simulated* at a give time step, based on the current system's state, as well as user-defined precision or cost thresholds. The second component *incrementally updates the system's state* based on the set of active degrees of freedom. In particular, incremental algorithms update the system's potential energy and forces. This allows the user to smoothly trade between precision and cost.

We detail this approach in two important types of simulations: Cartesian quasi-statics and torsion-angle dynamics. A novel, very general approach for adaptive dynamics simulations of particles — that has a number of important benefits over previous approaches — is mentioned in more detail in Section 6.1 .

#### 3.2. Adaptive Cartesian mechanics

In order to focus computations on a specific set of atoms, when performing quasi-static simulations (minimizations), we have developed an adaptive Cartesian mechanics algorithm, which decides which atoms should move at each time step.

In the simplest approach, we simply examine the force applied on each atom. When the norm of the force is above a user-defined threshold, the atom is active. Else the atom position is frozen. A slightly more elaborate version consists in defining the threshold automatically based on the system state (it might be e.g., the average applied force, a percentage of the maximum norm, etc.).

In order to avoid the linear cost of determining the set of active atoms at each time step, a binary tree is used to represent the system. Each leaf node represents an individual atom, while each internal node represents a set of atoms. Each leaf node stores the norm of the force applied to the corresponding atom. Each non-leaf node stores the maximum of the two force norms of its children, as illustrated in Figure 2 . We use two tree passes in order to update tree nodes' values and to determine the new active atoms. In the first, bottom-up pass, force norms are updated in a sub-tree of the binary tree (only some atoms have moved since the previous time step, so only some forces have been updated), starting from the leaves with modified norms, in  $O(k^{old}(\log(\frac{n}{k^{old}}) + 1))$  times where  $k^{old}$  is the number of active atoms and  $n$  the total number of atoms. In the second, top-down pass, the new active atoms (i.e., the atoms with the force norms which are now the largest), are determined in  $O(k^{new}(\log(\frac{n}{k^{new}}) + 1))$  times where  $k^{new}$  is the new number of active atoms. This process is illustrated in Figure 2 as well.

Precisely, Figure 2 illustrates the procedure to determine the active zone, when the threshold is automatically set to half the largest atomic force norm. In this example, the four leaves correspond to atoms 1 to 4. The value indicated in each leaf node is the norm of the force applied to its corresponding atom. For internal nodes, this value is the maximum of the norms of the forces applied to atoms in the corresponding group. In step 0, the threshold is automatically set to 10. As a result, only atom 1 moves. In step 1, the potential is incrementally updated, and the norms of the forces applied to atoms 1 and 2 are updated. In step 2, the values associated to the tree nodes are incrementally updated through a bottom-up pass that starts from the modified leaf nodes values. Because of this bottom-up update, the adaptive threshold becomes equal to 4. In step 3, the new active atoms are determined through a top-down pass, by visiting only the nodes that have a value larger than the adaptive threshold.

#### 3.3. Adaptive torsion-angle mechanics

In many situations, it is preferable to represent molecular systems as articulated bodies, and perform so-called *torsion-angle* mechanics. This may be to allow for larger time step sizes in a simulation, or because the user wants to focus to e.g., protein backbone deformations.



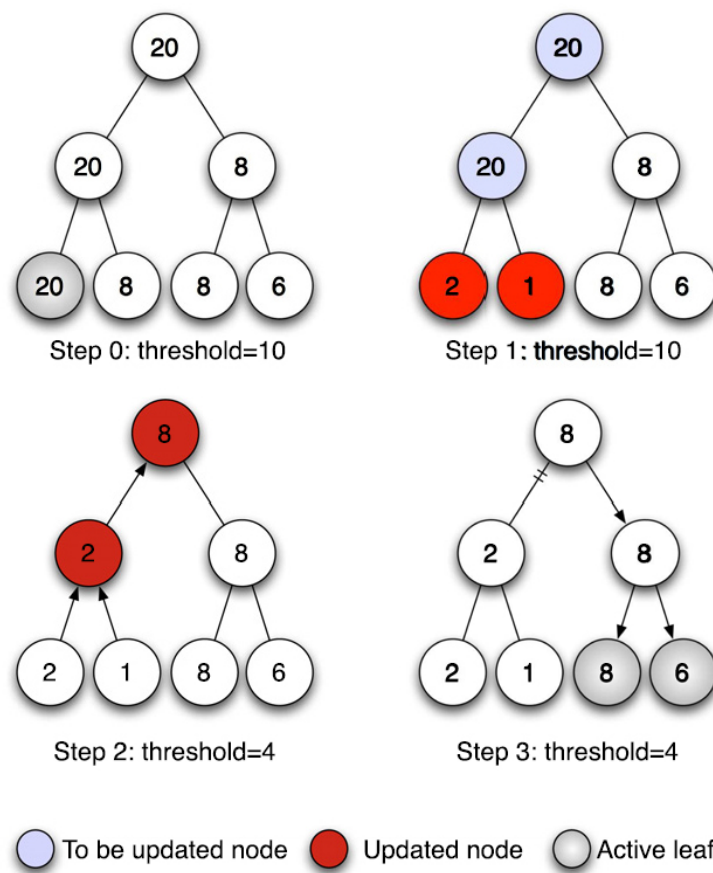


Figure 2. Adaptive Cartesian mechanics.

We have also developed an adaptive mechanics algorithm in the case of torsion-angle representations. In this case, a molecular system is recursively defined as the assembly of *two* molecular systems connected by a *joint* (when connecting two subassemblies which belong to the same molecule) or, more generally, by a *rigid body transform* (to assemble several molecules).

As in the Cartesian mechanics case, the complete molecular system is thus also represented by a binary tree, in which leaves are rigid bodies (a rigid body can be a single atom), internal nodes represent both sub-assemblies and connections between sub-assemblies, and the root represents the complete molecular system (see Figure 3 on the right, which shows an assembly tree associated to a short polyalanin). This hierarchical representation handles any branched molecule or groups of molecules, since the connections between two sub-molecular systems can be a rigid body transformation. In this representation, the positions of atoms are thus represented as superimposed rigid transformations: the position of any atom is obtained from the position of the whole set, to which is "added" the transformation from the complete set to the sub-set the atom belongs to, and so on until we reach the leaf node representing the atom. Similarly, the atomic motions are superimposed rigid motions.

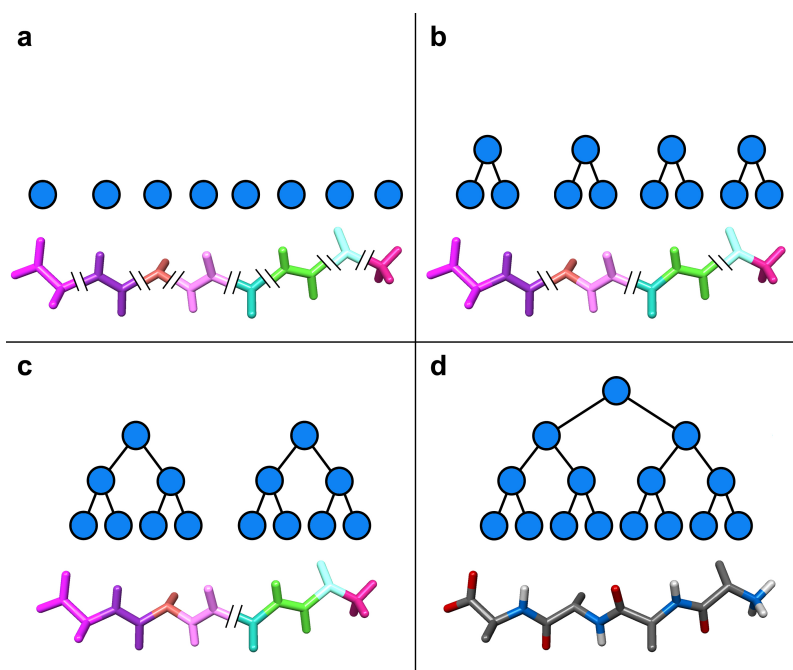


Figure 3. The assembly tree associated to a short polyalanin.

Our adaptive framework relies on two essential components. First, we associate a hierarchical set of reference frames to the assembly tree. Precisely, each node is associated to a local reference frame, in which all dynamical coefficients are expressed. This allows us to avoid updating these coefficients when a sub-assembly moves rigidly. Second, we have demonstrated that it is possible to determine a priori, at each time step, the set of joints which have the largest accelerations. Precisely, when going down the tree to compute joint accelerations, we are able to compute the weighted sum of the (squared) norms of joint accelerations in a sub-assembly  $C$  before computing joint accelerations themselves:

$$A(C) = (\mathbf{f}^C)^T \Psi^C \mathbf{f}^C + (\mathbf{f}^C)^T \mathbf{p}^C + \eta^C, \quad (14)$$

where the right part is a quadratic form of the spatial forces applied on the "handles" of node *C*. This allows us to cull away those sub-assemblies with (relatively) lower internal accelerations, and focus on the most mobile joints. Thus, at each time step, we can thus predict the set of joints with highest accelerations without computing all accelerations, and we simulate only a sub-tree of the assembly tree (the green nodes in the assembly tree, as in the figure above), based on a user-defined error threshold or computation time constraints. This sub-tree is called the active region, and may change at each time step.

We have exploited these two characteristics - hierarchical coordinate systems and adaptive motion refinement - to develop data structures and algorithms which enable adaptive molecular mechanics. The key observation in our approach is the following: all coefficients which only depend on relative atomic positions do not have to be updated when these relative positions do not change. We can thus store in each node of the assembly tree partial system states which hold information relative only to the node itself.

Precisely, each time step involves the following operations:

1. Adaptive acceleration update
  1. Determination of the acceleration update region: we determine the acceleration update region, i.e., the subset of nodes of the full articulated body which matter the most according to the acceleration metric, as indicated above. The union of the previous active region and the acceleration update region is the transient active region, i.e., the region temporarily considered as the active region.
  2. Joint accelerations projection: the acceleration is projected on the reduced motion space defined by the transient active region (to ensure that joint accelerations are consistent with both motion constraints and applied forces).
2. Adaptive velocity update
  1. Determination of the new active region: we update the joint velocities and the velocity metric values of the nodes in the transient active region. We then determine the set of nodes which are considered to be important according to the velocity metric (which is similar to the acceleration metric). This set becomes the new active region.
  2. Joint velocities projection: if one or more nodes become inactive due to the update of the active region, we determine a set of impulses that we must apply to the transient hybrid body to perform the rigidification of these nodes. This amounts to projecting joint velocities to the reduced motion space defined by the new active region.
3. Adaptive position update
  1. Position update: we update joint positions based on non-zero joint velocities in the active region.
  2. State update: once joint positions have been updated, we update the rest of the system's state: inverse inertias, acceleration metric coefficients, partial neighbor lists, partial force tables, etc.

Again, each of these steps involves a limited sub-tree of the assembly tree, which enables a fine control of the compromise between computation time and precision.

We have showed that our adaptive approach allows for a number of applications, some of which that were not possible for classical methods when using low-end desktop workstations. Indeed, by selecting a sufficiently small number of simultaneously active degrees of freedom, it becomes possible to perform interactive structural modifications of complex molecular systems.

## OPALE Project-Team

### 3. Scientific Foundations

#### 3.1. Functional and numerical analysis of PDE systems

Our common scientific background is the functional and numerical analysis of PDE systems, in particular with respect to nonlinear hyperbolic equations such as conservation laws of gas-dynamics.

Whereas the structure of weak solutions of the Euler equations has been thoroughly discussed in both the mathematical and fluid mechanics literature, in similar hyperbolic models, focus of new interest, such as those related to traffic, the situation is not so well established, except in one space dimension, and scalar equations. Thus, the study of such equations is one theme of emphasis of our research.

The well-developed domain of numerical methods for PDE systems, in particular finite volumes, constitute the sound background for PDE-constrained optimization.

#### 3.2. Numerical optimization of PDE systems

Partial Differential Equations (PDEs), finite volumes/elements, geometrical optimization, optimum shape design, multi-point/multi-criterion/multi-disciplinary optimization, shape parameterization, gradient-based/evolutionary/hybrid optimizers, hierarchical physical/numerical models, Proper Orthogonal Decomposition (POD)

Optimization problems involving systems governed by PDEs, such as optimum shape design in aerodynamics or electromagnetics, are more and more complex in the industrial setting.

In certain situations, the major difficulty resides in the costly evaluation of a functional by means of a simulation, and the numerical method to be used must exploit at best the problem characteristics (regularity or smoothness, local convexity).

In many other cases, several criteria are to be optimized and some are non differentiable and/or non convex. A large set of parameters, sometimes of different types (boolean, integer, real or functional), are to be taken into account, as well as constraints of various types (physical and geometrical, in particular). Additionally, today's most interesting optimization pre-industrial projects are multi-disciplinary, and this complicates the mathematical, physical and numerical settings. Developing *robust optimizers* is therefore an essential objective to make progress in this area of scientific computing.

In the area of numerical optimization algorithms, the project aims at adapting classical optimization methods (simplex, gradient, quasi-Newton) when applicable to relevant engineering applications, as well as developing and testing less conventional approaches such as Evolutionary Strategies (ES), including Genetic or Particle-Swarm Algorithms, or hybrid schemes, in contexts where robustness is a very severe constraint.

In a different perspective, the heritage from the former project Sinus in Finite-Volumes (or -Elements) for nonlinear hyperbolic problems, leads us to examine cost-efficiency issues of large shape-optimization applications with an emphasis on the PDE approximation; of particular interest to us:

- best approximation and shape-parameterization,
- convergence acceleration (in particular by multi-level methods),
- model reduction (e.g. by *Proper Orthogonal Decomposition*),
- parallel and grid computing; etc.

#### 3.3. Geometrical optimization

Jean-Paul Zolesio and Michel Delfour have developed, in particular in their book [6], a theoretical framework for geometrical optimization and shape control in Sobolev spaces.

In preparation to the construction of sound numerical techniques, their contribution remains a fundamental building block for the functional analysis of shape optimization formulations.

### 3.4. Integration platforms

Developing grid, cloud and high-performance computing for complex applications is one of the priorities of the IST chapter in the 7th Framework Program of the European Community. One of the challenges of the 21st century in the computer science area lies in the integration of various expertise in complex application areas such as simulation and optimization in aeronautics, automotive and nuclear simulation. Indeed, the design of the reentry vehicle of a space shuttle calls for aerothermal, aerostructure and aerodynamics disciplines which all interact in hypersonic regime, together with electromagnetics. Further, efficient, reliable, and safe design of aircraft involve thermal flows analysis, consumption optimization, noise reduction for environmental safety, using for example aeroacoustics expertise.

The integration of such various disciplines requires powerful computing infrastructures and particular software coupling techniques. Simultaneously, advances in computer technology militate in favor of the use of massively parallel clusters including hundreds of thousands of processors connected by high-speed gigabits/sec networks. This conjunction makes it possible for an unprecedented cross-fertilization of computational methods and computer science. New approaches including evolutionary algorithms, parameterization, multi-hierarchical decomposition lend themselves seamlessly to parallel implementations in such computing infrastructures. This opportunity is being dealt with by the OPALE project since its very beginning. A software integration platform has been designed by the OPALE project for the definition, configuration and deployment of multidisciplinary applications on a distributed heterogeneous infrastructure. Experiments conducted within European projects and industrial cooperations using CAST have led to significant performance results in complex aerodynamics optimization test-cases involving multi-elements airfoils and evolutionary algorithms, i.e. coupling genetic and hierarchical algorithms involving game strategies [70].

The main difficulty still remains however in the deployment and control of complex distributed applications by the end-users. Indeed, the deployment of the computing infrastructures and of the applications in such environments still requires specific expertise by computer science specialists. However, the users, which are experts in their particular application fields, e.g. aerodynamics, are not necessarily experts in distributed and grid computing. Being accustomed to Internet browsers, they want similar interfaces to interact with high-performance computing and problem-solving environments. A first approach to solve this problem is to define component-based infrastructures, e.g. the Corba Component Model, where the applications are considered as connection networks including various application codes. The advantage is here to implement a uniform approach for both the underlying infrastructure and the application modules. However, it still requires specific expertise not directly related to the application domains of each particular user. A second approach is to make use of web services, defined as application and support procedures to standardize access and invocation to remote support and application codes. This is usually considered as an extension of Web services to distributed infrastructures. A new approach, which is currently being explored by the OPALE project, is the design of a virtual computing environment able to hide the underlying high-performance-computing infrastructures to the users. The team is exploring the use of distributed workflows to define, monitor and control the execution of high-performance simulations on distributed clusters. The platform includes resilience, i.e., fault-tolerance features allowing for resource demanding and erroneous applications to be dynamically restarted safely, without user intervention.

## POEMS Project-Team

### 3. Scientific Foundations

#### 3.1. Mathematical analysis and simulation of wave propagation

Our activity relies on the existence of mathematical models established by physicists to model the propagation of waves in various situations. The basic ingredient is a partial differential equation (or a system of partial differential equations) of the hyperbolic type that are often (but not always) linear for most of the applications we are interested in. The prototype equation is the wave equation:

$$\frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = 0,$$

which can be directly applied to acoustic waves but which also constitutes a simplified scalar model for other types of waves (This is why the development of new numerical methods often begins by their application to the wave equation). Of course, taking into account more realistic physics will enrich and complexify the basic models (presence of sources, boundary conditions, coupling of models, integro-differential or non linear terms,...)

It is classical to distinguish between two types of problems associated with these models: the time domain problems and the frequency domain (or time harmonic) problems. In the first case, the time is one of the variables of which the unknown solution depends and one has to face an evolution problem. In the second case (which rigorously makes sense only for linear problems), the dependence with respect to time is imposed a priori (via the source term for instance): the solution is supposed to be harmonic in time, proportional to  $e^{i\omega t}$ , where  $\omega > 0$  denotes the pulsation (also commonly, but improperly, called the frequency). Therefore, the time dependence occurs only through this pulsation which is given a priori and plays the rôle of a parameter: the unknown is only a function of space variables. For instance, the wave equation leads to the Helmholtz wave equation (also called the reduced wave equation) :

$$-c^2 \Delta u - \omega^2 u = 0.$$

These two types of problems, although deduced from the same physical modelling, have very different mathematical properties and require the development of adapted numerical methods.

However, there is generally one common feature between the two problems: the existence of a dimension characteristic of the physical phenomenon: the wavelength. Intuitively, this dimension is the length along which the searched solution varies substantially. In the case of the propagation of a wave in an heterogeneous medium, it is necessary to speak of several wavelengths (the wavelength can vary from one medium to another). This quantity has a fundamental influence on the behavior of the solution and its knowledge will have a great influence on the choice of a numerical method.

Nowadays, the numerical techniques for solving the basic academic and industrial problems are well mastered. A lot of companies have at their disposal computational codes whose limits (in particular in terms of accuracy or robustness) are well known. However, the resolution of complex wave propagation problems close to real applications still poses (essentially open) problems which constitute a real challenge for applied mathematicians. A large part of research in mathematics applied to wave propagation problems is oriented towards the following goals:

- the conception of new numerical methods, more and more accurate and high performing.
- the treatment of more and more complex problems (non local models, non linear models, coupled systems, ...)

- the study of specific phenomena or features such as guided waves, resonances,...
- the development of approximate models in various situations,
- imaging techniques and inverse problems related to wave propagation.

## SCIPORT Team

### 3. Scientific Foundations

#### 3.1. Automatic Differentiation

**Participants:** Laurent Hascoët, Valérie Pascual.

**automatic differentiation** (AD) Automatic transformation of a program, that returns a new program that computes some derivatives of the given initial program, i.e. some combination of the partial derivatives of the program's outputs with respect to its inputs.

**adjoint** Mathematical manipulation of the Partial Derivative Equations that define a problem, obtaining new differential equations that define the gradient of the original problem's solution.

**checkpointing** General trade-off technique, used in adjoint-mode AD, that trades duplicate execution of a part of the program to save some memory space that was used to save intermediate results. Checkpointing a code fragment amounts to running this fragment without any storage of intermediate values, thus saving memory space. Later, when such an intermediate value is required, the fragment is run a second time to obtain the required values.

Automatic or Algorithmic Differentiation (AD) differentiates *programs*. An AD tool takes as input a source program  $P$  that, given a vector argument  $X \in \mathbb{R}^n$ , computes some vector result  $Y = F(X) \in \mathbb{R}^m$ . The AD tool generates a new source program  $P'$  that, given the argument  $X$ , computes some derivatives of  $F$ . The resulting  $P'$  reuses the control of  $P$ .

For any given control,  $P$  is equivalent to a sequence of instructions, which is identified with a composition of vector functions. Thus, if

$$\begin{aligned} P & \text{ is } \{I_1; I_2; \dots; I_p\}, \\ F & = f_p \circ f_{p-1} \circ \dots \circ f_1, \end{aligned} \quad (15)$$

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

$$F'(X) = f'_p(X_{p-1}) \cdot f'_{p-1}(X_{p-2}) \cdot \dots \cdot f'_1(X_0) \quad (16)$$

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

In practice, the Jacobian  $F'(X)$  is often too expensive to compute and store, but most applications only need projections of  $F'(X)$  such as:

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

$$F'(X) \cdot \dot{X} = f'_p(X_{p-1}) \cdot f'_{p-1}(X_{p-2}) \cdot \dots \cdot f'_1(X_0) \cdot \dot{X} \quad (17)$$

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

- **Adjoints**, defined for a given weighting  $\bar{Y}$  of the outputs as:



$$F'^*(X).\bar{Y} = f_1'^*(X_0).f_2'^*(X_1). \dots .f_{p-1}'^*(X_{p-2}).f_p'^*(X_{p-1}).\bar{Y} \quad (18)$$

Adjoint mode AD turns out to make a very efficient program, at least theoretically [30]. The computation time required for the gradient is only a small multiple of the run-time of  $P$ . It is independent from the number of parameters  $n$ . In contrast, computing the same gradient with the *tangent mode* would require running the tangent differentiated program  $n$  times.

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

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

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

## 3.2. Static Analysis and Transformation of programs

**Participants:** Laurent Hascoët, Valérie Pascual.

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

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

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

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

**data dependence analysis** Program analysis that studies the itinerary of values during program execution, from the place where a value is generated to the places where it is used, and finally to the place where it is overwritten. The collection of all these itineraries is often stored as a *data dependence graph*, and data flow analysis most often rely on this graph.

**data dependence graph** Directed graph that relates accesses to program variables, from the write access that defines a new value to the read accesses that use this value, and conversely from the read accesses to the write access that overwrites this value. Dependences express a partial order between operations, that must be preserved to preserve the program's result.

The most obvious example of a program transformation tool is certainly a compiler. Other examples are program translators, that go from one language or formalism to another, or optimizers, that transform a program to make it run better. AD is just one such transformation. These tools use sophisticated analysis [20] to improve the quality of the produced code. These tools share their technological basis. More importantly, there are common mathematical models to specify and analyze them.

An important principle is *abstraction*: the core of a compiler should not bother about syntactic details of the compiled program. The optimization and code generation phases must be independent from the particular input programming language. This is generally achieved using language-specific *front-ends* and *back-ends*. Abstraction can go further: the internal representation becomes more language independent, and semantic constructs can be unified. Analysis can then concentrate on the semantics of a small set of constructs. We advocate an internal representation composed of three levels.

- At the top level is the *call graph*, whose nodes are modules and procedures. Arrows relate nodes that call or import one another. Recursion leads to cycles.
- At the middle level is the *flow graph*, one per procedure. It captures the control flow between atomic instructions.
- At the lowest level are abstract *syntax trees* for the individual atomic instructions. Semantic transformations can benefit from the representation of expressions as directed acyclic graphs, sharing common sub-expressions.

To each level belong symbol tables, nested to capture scoping.

Static program analysis can be defined on this internal representation, which is largely language independent. The simplest analyses on trees can be specified with inference rules [23], [31], [21]. But many analyses are more complex, and better defined on graphs than on trees. This is the case for *data-flow analyses*, that look for run-time properties of variables. Since flow graphs are cyclic, these global analyses generally require an iterative resolution. *Data flow equations* is a practical formalism to describe data-flow analyses. Another formalism is described in [24], which is more precise because it can distinguish separate *instances* of instructions. However it is still based on trees, and its cost forbids application to large codes. *Abstract Interpretation* [25] is a theoretical framework to study complexity and termination of these analyses.

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

Even then, data flow analyses are limited, because they are static and thus have very little knowledge of actual run-time values. In addition to the very theoretical limit of *undecidability*, there are practical limitations to how much information one can infer from programs that use arrays [37], [26] or pointers. In general, conservative *over-approximations* are always made that lead to derivative code that is less efficient than possibly achievable.

### 3.3. Automatic Differentiation and Scientific Computing

**Participants:** Alain Dervieux, Laurent Hascoët, Bruno Koobus.

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

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

Scientific Computing provides reliable simulations of complex systems. For example it is possible to simulate the 3D air flow around a plane that captures the physical phenomena of shocks and turbulence. Next comes optimization, one degree higher in complexity because it repeatedly simulates and applies optimization steps until an optimum is reached. We focus on gradient-based optimization.

We investigate several approaches to obtain the gradient, between two extremes:

- One can write an *adjoint system* of mathematical equations, then discretize it and program it by hand. This is mathematically sound [28], but very costly in development time. It also does not produce an exact gradient of the discrete function, and this can be a problem if using optimization methods based on descent directions.
- One can apply adjoint-mode AD (cf 3.1) on the program that discretizes and solves the direct system. This gives in fact the adjoint of the discrete function computed by the program. Theoretical results [27] guarantee convergence of these derivatives when the direct program converges. This approach is highly mechanizable, but leads to massive use of storage and may require code transformation by hand [32], [35] to reduce memory usage.

If for instance the model is steady, one tradeoff can use the iterated states in the direct order [29]. Another tradeoff can use only the fully converged final state. Since tradeoff approaches can be error-prone, we advocate incorporating them into the AD model and into the AD tools.

## SIMPAF Project-Team

### 3. Scientific Foundations

#### 3.1. General framework

Partial Differential Equations, Kinetic Equations, Conservation Laws, Hyperbolic Systems, Fluid Mechanics, Parabolic Systems, Computational Fluid Dynamics, Plasma Physics, Asymptotic analysis

The scientific activity of the project is concerned with Partial Differential Equations (PDE) arising from the physical description of particles and fluids. It covers various viewpoints:

- At first, the words “particles and fluids” could simply mean that we are interested independently in models for particles, which can either be considered as individuals (which leads to “ $N$ -particle models”,  $N$  ranging from 1 to many) or through a statistical description (which leads to kinetic equations) as well as in models for fluids like Euler and Navier-Stokes equations or plasma physics.
- However, many particle systems can also be viewed as a fluid, via a passage from microscopic to macroscopic viewpoint, that is, a hydrodynamic limit.
- Conversely, a fruitful idea to build numerical solvers for hyperbolic conservation laws consists in coming back to a kinetic formulation. This approach has motivated the introduction of the so-called kinetic schemes.

By nature these problems describe multiscale phenomena and one of the major difficulties when studying them lies in the interactions between the various scales: number of particles, size, different time and length scales, coupling...

The originality of the project is to consider a wide spectrum of potential applications. In particular, the word “particles” covers various and very different physical situations and it has evolved with the composition of the team. One may think of:

- charged particles: description of semi-conductor devices or plasmas;
- bacteria, individuals or genes as in models motivated by biology or population dynamics;
- droplets and bubbles, as in Fluid/Particles Interaction models which arise in the description of sprays and aerosols, smoke and dust, combustion phenomena (aeronautics or engine design), industrial process in metallurgy...
- cross-links in polymer chains to describe rubber elasticity;
- oxyde molecules to model corrosion phenomena at the microscopic scale and derive effective macroscopic equations;
- cold atoms...

We aim at focusing on all the aspects of the problem:

- Modelling mathematically complex physics requires a deep discussion of the leading phenomena and the role of the physical parameters. With this respect, the asymptotic analysis is a crucial issue, the goal being to derive reduced models which can be solved with a reduced numerical cost but still provide accurate results in the physical situations that are considered.
- The mathematical analysis of the equations provides important qualitative properties of the solutions: well-posedness, stability, smoothness of the solutions, large time behavior... which in turn can motivate the design of numerical methods.
- Eventually, we aim at developing specific numerical methods and performing numerical simulations for these models, in order to validate the theoretical results and shed some light on the physics.

The team has been composed in order to study these various aspects simultaneously. In particular, we wish to keep a balance between modelling, analysis, development of original codes and simulations.

## 3.2. Interactions of Micro- and Macroscopic Scales and Simulations

Statistical Physics, Homogenization, Asymptotic Preserving Schemes

### 3.2.1. Homogenization methods

Homogenization methods aim at replacing a PDE with highly oscillatory coefficients by an effective PDE with smoother coefficients, whose solution captures the averaged behavior of the true oscillatory solution. The effective determination of the homogenized PDE is however not trivial (especially in the nonlinear or/and stochastic cases). Numerical approximations of the solution of the homogenized PDE is the heart of numerical homogenization.

Homogenization methods are used in many application fields. The two applications we are specifically interested in are material sciences (in particular the determination of macroscopic constitutive laws for rubber starting from polymer-chain networks) and nuclear waste storage (in particular the evolution of nuclear wastes in complex storage devices).

The team is interested in qualitative as well as quantitative results, and theoretical as well as numerical results. Challenging questions are mainly related to nonlinear problems (nonlinear elasticity for instance) and stochastic problems (especially regarding quantitative results).

### 3.2.2. Statistical physics : molecular dynamics

The team is concerned with the numerical simulation of stochastically perturbed Molecular Dynamics. The main goal is to handle in the same simulation the fastest time scales (e.g. the oscillations of molecular bindings), and the slowest time scales (e.g. the so-called reaction coordinates). Recently, M. Rousset co-authored a monograph [64] which summarizes standard and state-of-the-art free energy calculations, that are used to accelerate slow variables in MD simulations.

### 3.2.3. Statistical physics: dynamical friction, fluctuations and approach to equilibrium

In models of charge transport, say transport of electrons, a phenomenological friction force is generally introduced, which is proportional to the velocity  $v$ . The dissipation induced by such a term is essential for the description of phenomena such as Ohm's law and approach to equilibrium. Our idea is to go back to a microscopic framework, with a description of the energy exchanges between the electrons and the surrounding medium which is the ultimate source of the dissipation of energy by the medium and of an effective friction force. We have shown numerically and argued theoretically that the balance between the fluctuations and the dissipation by the medium drives the particle to thermal equilibrium. The goal is now to provide rigorous proof of this statement. As a first step in this program, results will be obtained in an appropriate weak coupling limit. This program requires efforts in modelling, probability and analysis, but the questions are also really challenging for numerics, due, notably, to the large number of degrees of freedom involved in the equation. The subject is at the heart of the PhD work of É. Soret, now in her second year as a PhD student.

### 3.2.4. Cold Atoms

In the framework of the Labex CEMPI, C. Besse, S. De Bièvre and G. Dujardin are working, in collaboration with J.-C. Garreau and the cold-atom team at PhLAM, on the mathematical analysis and the numerical simulation of kicked rotor systems. Such systems are experimentally realized at PhLAM. A triple goal is being pursued: understand the effect of non-linearities on dynamical localization, understand dynamical localization in systems other than kicked rotors, and exploring the limits of the analogy between kicked systems and the Anderson model.

## 3.3. Plasma

In the context of the Galileo satellite-positioning system, C. Besse and C. Yang, members of the ANR Iodissee project, developed a hierarchy of plasma models which describe ionospheric scintillations. This hierarchy involves many small parameters, and they introduced an asymptotic preserving scheme which allows one to take small parameters into account without solving the problem on a very fine grid [8]. The next step is to

understand the fading and phase variations when waves propagate in this medium. This is a work in progress with P. Lafitte and S. Minjeaud (CNRS, Université de Nice).

### 3.4. Finite element and finite volume methods

Conservation Laws, Anti-Diffusive Schemes, Viscous Flows, Control, Turbulence, Finite element methods, Finite volume methods

#### 3.4.1. Control in Fluid Mechanics

Flow control techniques are widely used to improve the performances of planes or vehicles, or to drive some internal flows arising for example in combustion chambers. Indeed, they can sensibly reduce energy consumption, noise disturbances, or prevent the flow from undesirable behaviors.

E. Creusé is involved in the development of open and closed active flow control, with applications to recirculation in engines or blood flows.

#### 3.4.2. Numerical Methods for Viscous Flows

Numerical investigations are very useful to check the behavior of systems of equations modelling very complicate dynamics. In order to simulate the motion of mixtures of immiscible fluids having different densities, a recent contribution of the team was to develop an hybrid Finite Element / Finite Volume scheme for the resolution of the variable density 2D incompressible Navier-Stokes equations. The main points of this work were to ensure the consistency of the new method [56] as well as its stability for high density ratios [54]. In order to answer these questions, we have developed a MATLAB code and a C++ code. In the following of this work, C. Calgaro and E. Creusé now have in mind the following objectives, in collaboration with T. Goudon (team COFFEE, Inria Sophia-Antipolis) :

- Distribute the matlab version of the code (with an accurate documentation and a graphic interface) to promote new collaborations in the domain and compare alternative numerical solution methods (for instance to compare updating LU factorizations, see [55]);
- To generalize the stability results obtained in [54] for the scalar transport equation to the full 2D Euler system, in particular very low density values density (near vacuum);
- Complete the C++ code to treat more general hydrodynamic models (combustion theory, transport of pollutants). We plan to check the behavior of the equations (typically the Kazhikhov-Smagulov model of powder-snow avalanches) in the regime when the current existence theory does not apply, and extend our kinetic asymptotic-based schemes to such problems.

#### 3.4.3. A posteriori error estimators for finite element methods

A posteriori estimates, finite element methods

The team works on a posteriori error estimators for finite element methods, applied to the resolution of several partial differential equations. The objective is to derive useful tools in order to control the global error between the exact solution and the approximated one (reliability of the estimator), and to control the local error leading to adaptive mesh refinement strategies (efficiency of the estimator).

More specifically, E. Creusé works on the derivation of some "reconstruction estimators" based on gradient averaging, for diffusion problems (with S. Nicaise, LAMAV, Valenciennes), the Reissner-Mindlin system (PhD of É. Verhille), and the Maxwell equations (PhD of Z. Tang).

### 3.5. Numerical analysis of Schrödinger equations

Dispersive equations, Schrödinger equations

### **3.5.1. Modelling of quantum dot-helium**

In collaboration with G. Reinish (Nice Observatory) and V. Guðmundsson (University of Reykjavik), C. Besse and G. Dujardin are working on the numerical computation of the ground state and the first bound states of the non linear Schrödinger-Poisson system with confining quadratic potential in 2 space dimensions. This models quantum dot helium (*i.e.* the behavior of a pair of quantum electrons in a strong confining potential). The goal is to perform after that numerical time stepping methods to simulate the dynamics of the NLSP system and compute accurately some quantities of physical interest as functions of time, in order to be able to compare the competition between the Coulomb (repulsive) interaction and the binding (attractive) forces due to the confinement in this model as well as in other quantum mechanics models.

### **3.5.2. Dispersive Schrödinger-like equations**

In collaboration with M. Taki (PhLAM laboratory, Lille), C. Besse and G. Dujardin are considering dispersive equations modelling the propagation of a laser beam in an optical fiber. They are trying to explain the possible ways of creating rogue waves in the propagation of laser beams. More generally, they are trying to explain which terms in the dispersive Schrödinger-like equations obtained by the physicists allow which physical behaviour of the solutions (e.g. the creation of rogue waves).

---

**APICS Project-Team**

## 3. Scientific Foundations

### 3.1. Introduction

Within the extensive field of inverse problems, much of the research by APICS deals with reconstructing solutions of classical elliptic PDEs from their boundary behaviour. Perhaps the most basic example of such a problem is harmonic identification of a stable linear dynamical system: the transfer-function  $f$  is holomorphic in the right half-plane, which means it satisfies there the Cauchy-Riemann equation  $\bar{\partial}f = 0$ , and in principle  $f$  can be recovered from its values on the imaginary axis, *e.g.* by Cauchy formula.

Practice is not nearly as simple, for  $f$  is only measured pointwise in the pass-band of the system which makes the problem ill-posed [69]. Moreover, the transfer function is usually sought in specific form, displaying the necessary physical parameters for control and design. For instance if  $f$  is rational of degree  $n$ , it satisfies  $\bar{\partial}f = \sum_{j=1}^n a_j \delta_{z_j}$  where the  $z_j$  are its poles, and finding the domain of holomorphy (*i.e.* locating the  $z_j$ ) amounts to solve a (degenerate) free-boundary inverse problem, this time on the left half-plane. To address these questions, the team has developed a two-step approach as follows.

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

Step 2: To compute a reduced order model. This typically consists of rational approximation of the complete model obtained in step 1, or phase-shift thereof to account for delays. Derivation of the complete model is important to achieve stability of the reduced one.

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

Iterating the previous steps coupled with a sensitivity analysis yields a tuning procedure which was first demonstrated in [77] on resonant microwave filters.

Similar steps can be taken to approach design problems in frequency domain, replacing measured behaviour by desired behaviour. However, describing achievable responses from the design parameters at hand is generally cumbersome, and most constructive techniques rely on rather specific criteria adapted to the physics of the problem. This is especially true of circuits and filters, whose design classically appeals to standard polynomial extremal problems and realization procedures from system theory [70], [55]. APICS is active in this field, where we introduced the use of Zolotarev-like problems for microwave multiband filter design. We currently favor interpolation techniques because of their transparency with respect to parameter use, see section 3.2.2 .

In another connection, the example of harmonic identification quickly suggests a generalization of itself. Indeed, on identifying  $\mathbb{C}$  with  $\mathbb{R}^2$ , holomorphic functions become conjugate-gradients of harmonic functions so that harmonic identification is, after all, a special case of a classical issue: to recover a harmonic function on a domain from partial knowledge of the Dirichlet-Neumann data; portion of the boundary where data are not available may be unknown, in which case we meet a free boundary problem. This framework for 2-D non-destructive control was first advocated in [59] and subsequently received considerable attention. This framework makes it clear how to state similar problems in higher dimensions and for more general operators than the Laplacian, provided solutions are essentially determined by the trace of their gradient on part of the boundary which is the case for elliptic equations <sup>1</sup> [79]. All these questions are particular instances of the

---



so-called inverse potential problem, where a measure  $\mu$  has to be recovered from knowledge of the gradient of its potential (*i.e.*, the field) on part of a hypersurface (a curve in 2-D) encompassing the support of  $\mu$ . For Laplace's operator, potentials are logarithmic in 2-D and Newtonian in higher dimensions. For elliptic operators with non constant coefficients, the potential depends on the form of fundamental solutions and is less manageable because it is no longer of convolution type. In any case, by construction, the operator applied to the potential yields back the measure.

Inverse potential problems are severely indeterminate because infinitely many measures within an open set produce the same field outside this set [68]. In step 1 above we implicitly removed this indeterminacy by requiring that the measure be supported on the boundary (because we seek a function holomorphic throughout the right half space), and in step 2 by requiring, say, in case of rational approximation that the measure be discrete in the left half-plane. The same discreteness assumption prevails in 3-D inverse source problems. To recap, the gist of our approach is to approximate boundary data by (boundary traces of) fields arising from potentials of measures with specific support. Note this is different from standard approaches to inverse problems, where descent algorithms are applied to integration schemes of the direct problem; in such methods, it is the equation which gets approximated (in fact: discretized).

Along these lines, the team initiated the use of steps 1 and 2 above, along with singularity analysis, to approach issues of nondestructive control in 2 and 3-D [41] [6], [2]. We are currently engaged in two kinds of generalization, further described in section 3.2.1. The first one deals with non-constant conductivities, where Cauchy-Riemann equations for holomorphic functions are replaced by conjugate Beltrami equations for pseudo-holomorphic functions; there we seek applications to plasma confinement. The other one lies with inverse source problems for Laplace's equation in 3-D, where holomorphic functions are replaced by harmonic gradients, developing applications to EEG/MEG and inverse magnetization problems in paleomagnetism, see section 4.2

The main approximation-theoretic tools developed by APICS to get to grips with issues mentioned so far are outlined in section 3.3. In section 3.2 to come, we make more precise which problems are considered and for which applications.

## 3.2. Range of inverse problems

### 3.2.1. Elliptic partial differential equations (PDE)

**Participants:** Laurent Baratchart, Slah Chaabi, Juliette Leblond, Ana-Maria Nicu, Dmitry Ponomarev, Elodie Pozzi.

This work is done in collaboration with Alexander Borichev (Univ. Provence).

Reconstructing Dirichlet-Neumann boundary conditions for a function harmonic in a plane domain when these are known on a strict subset  $E$  of the boundary, is equivalent to recover a holomorphic function in the domain from its boundary values on  $E$ . This is the problem raised on the half-plane in step 1 of section 3.1. It makes good sense in holomorphic Hardy spaces where functions are determined by their values on boundary subsets of positive linear measure, which is the framework for problem (P) in section 3.3.1. Such problems naturally arise in nondestructive testing of 2-D (or cylindrical) materials from partial electrical measurements on the boundary. Indeed, the ratio between tangential and normal currents (so-called Robin coefficient) tells about corrosion of the material. Solving problem (P) where  $\psi$  is chosen to be the response of some uncorroded piece with identical shape allows one to approach such questions, and this was an initial application of holomorphic extremal problems to non-destructive control [56], [52].

<sup>1</sup>There is a subtle difference here between dimension 2 and higher. Indeed, a function holomorphic on a plane domain is defined by its non-tangential limit on a boundary subset of positive linear measure, but there are non-constant harmonic functions in the 3-D ball,  $C^1$  up to the boundary sphere, yet having vanishing gradient on a subset of positive measure of the sphere

A recent application by the team deals with non-constant conductivity over a doubly connected domain,  $E$  being the outer boundary. Measuring Dirichlet-Neumann data on  $E$ , we wanted to check whether the solution is constant on the inner boundary. We first had to define and study Hardy spaces of the conjugate Beltrami equation, of which the conductivity equation is the compatibility condition (just like Laplace's equation is the compatibility condition of the Cauchy-Riemann system). This was done in references [5] and [35]. Then, solving an obvious modification of problem ( $P$ ) allows one to numerically check what we want. Further, the value of this extremal problem defines a criterion on inner boundaries, and subsequently a descent algorithm was set up to improve the initial boundary into one where the solution is closer to being constant, thereby trying to solve a free boundary problem..

When the domain is regarded as separating the edge of a tokamak's vessel from the plasma (rotational symmetry makes this a 2-D problem), the procedure just described suits plasma control from magnetic confinement. It was successfully applied in collaboration with CEA (the French nuclear agency) and the University of Nice (JAD Lab.) to data from *Tore Supra* [58], see section 6.2 . This procedure is fast because no numerical integration of the underlying PDE is needed, as an explicit basis of solutions to the conjugate Beltrami equation was found in this case.

Three-dimensional versions of step 1 in section 3.1 are also considered, namely to recover a harmonic function (up to a constant) in a ball or a half-space from partial knowledge of its gradient on the boundary. Such questions arise naturally in connection with neurosciences and medical imaging (electroencephalography, EEG) or in paleomagnetism (analysis of rocks magnetization) [2] [37], see section 6.1 . They are not yet as developed as the 2-D case where the power of complex analysis is at work, but considerable progress was made over the last years through methods of harmonic analysis and operator theory.

The team is also concerned with non-destructive control problems of localizing defaults such as cracks, sources or occlusions in a planar or 3-dimensional domain, from boundary data (which may correspond to thermal, electrical, or magnetic measurements). These defaults can be expressed as a lack of analyticity of the solution of the associated Dirichlet-Neumann problem and we approach them using techniques of best rational or meromorphic approximation on the boundary of the object [4] [16], see sections 3.3.2 and 4.2 . In fact, the way singularities of the approximant relate to the singularities of the approximated function is an all-pervasive theme in approximation theory, and for appropriate classes of functions the location of the poles of a best rational approximant can be used as an estimator of the singularities of the approximated function (see section 6.1 ). This circle of ideas is much in the spirit of step 2 in section 3.1 .

A genuine 3-dimensional theory of approximation by discrete potentials, though, is still in its infancy.

### 3.2.2. Systems, transfer and scattering

**Participants:** Laurent Baratchart, Sylvain Chevillard, Sanda Lefteriu, Martine Olivi, Fabien Seyfert.

Through initial contacts with CNES, the French space agency, the team came to work on identification-for-tuning of microwave electromagnetic filters used in space telecommunications (see section 4.3 ). The problem was to recover, from band-limited frequency measurements, the physical parameters of the device under examination. The latter consists of interconnected dual-mode resonant cavities with negligible loss, hence its scattering matrix is modelled by a  $2 \times 2$  unitary-valued matrix function on the frequency line, say the imaginary axis to fix ideas. In the bandwidth around the resonant frequency, a modal approximation of the Helmholtz equation in the cavities shows that this matrix is approximately rational, of Mc-Millan degree twice the number of cavities.

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

Hardy spaces, and in particular the Hilbert space  $H^2$ , provide a framework to transform this classical ill-posed issue into a series of well-posed analytic and meromorphic approximation problems. The procedure sketched in section 3.1 now goes as follows:

1. infer from the pointwise boundary data in the bandwidth a stable transfer function (*i.e.* one which is holomorphic in the right half-plane), that may be infinite dimensional (numerically: of high degree). This is done by solving in the Hardy space  $H^2$  of the right half-plane a problem analogous to (P) in section 3.3.1, taking into account prior knowledge on the decay of the response outside the bandwidth, see [18] for details.
2. From this stable model, a rational stable approximation of appropriate degree is computed. For this a descent method is used on the relatively compact manifold of inner matrices of given size and degree, using a novel parametrization of stable transfer functions [18].
3. From this rational model, realizations meeting certain constraints imposed by the technology in use are computed (see section 6.3). These constraints typically come from the nature and topology of the equivalent electrical network used to model the filter. This network is composed of resonators, coupled to each other by some specific coupling topology. Performing this realization step for given coupling topology can be recast, under appropriate compatibility conditions [8], as the problem of solving a zero-dimensional multivariate polynomial system. To tackle this problem in practice, we use Groebner basis techniques as well as continuation methods as implemented in the Dedale-HF software (5.4).

Let us also mention that extensions of classical coupling matrix theory to frequency-dependent (reactive) couplings have lately been carried-out [1] for wide-band design applications, but further study is needed to make them effective.

Subsequently APICS started investigating issues pertaining to filter design rather than identification. Given the topology of the filter, a basic problem is to find the optimal response with respect to amplitude specifications in frequency domain bearing on rejection, transmission and group delay of scattering parameters. Generalizing the approach based on Tchebychev polynomials for single band filters, we recast the problem of multi-band response synthesis in terms of a generalization of classical Zolotarev min-max problem [30] to rational functions [11]. Thanks to quasi-convexity, the latter can be solved efficiently using iterative methods relying on linear programming. These are implemented in the software easy-FF (see section 5.5).

Later, investigations by the team extended to design and identification of more complex microwave devices, like multiplexers and routers, which connect several filters through wave guides. Schur analysis plays an important role in such studies, which is no surprise since scattering matrices of passive systems are of Schur type (*i.e.* contractive in the stability region). The theory originates with the work of I. Schur [76], who devised a recursive test to check for contractivity of a holomorphic function in the disk. Generalizations thereof turned out to be very efficient to parametrize solutions to contractive interpolation problems subject to a well-known compatibility condition (positive definiteness of the so-called Pick matrix) [32]. Schur analysis became quite popular in electrical engineering, as the Schur recursion precisely describes how to chain two-port circuits.

Dwelling on this, members of the team contributed to differential parametrizations (atlases of charts) of lossless matrix functions to the theory [31][12], [10]. They are of fundamental use in our rational approximation software RARL2 (see section 5.1). Schur analysis is also instrumental to approach de-embedding issues considered in section 6.4, and provides further background to current studies by the team of synthesis and adaptation problems for multiplexers. At the heart of the latter lies a variant of contractive interpolation with degree constraint introduced in [62].

We also mention the role played by multipoint Schur analysis in the team's investigation of spectral representation for certain non-stationary discrete stochastic processes [3], [36].

Recently, in collaboration with UPV (Bilbao), our attention was driven by CNES, to questions of stability relative to high-frequency amplifiers, see section 7.2. Contrary to previously mentioned devices, these are *active* components. The amplifier can be linearized at a functioning point and admittances of the corresponding electrical network can be computed at various frequencies, using the so-called harmonic balance method. The

goal is to check for stability of this linearised model. The latter is composed of lumped electrical elements namely inductors, capacitors, negative *and* positive reactors, transmission lines, and commanded current sources. Research so far focused on determining the algebraic structure of admittance functions, and setting up a function-theoretic framework to analyse them. In particular, much effort was put on realistic assumptions under which a stable/unstable decomposition can be claimed in  $H^2 \oplus \overline{H^2}$  (see section 6.5 ). Under them, the unstable part of the elements under examination is rational and we expect to bring valuable estimates of stability to the designer using the general scheme in section 3.1 .

### 3.3. Approximation of boundary data

**Participants:** Laurent Baratchart, Sylvain Chevillard, Juliette Leblond, Martine Olivi, Dmitry Ponomarev, Elodie Pozzi, Fabien Seyfert.

The following people are collaborating with us on these topics: Bernard Hanzon (Univ. Cork, Ireland), Jean-Paul Marmorat (Centre de mathématiques appliquées (CMA), École des Mines de Paris), Jonathan Partington (Univ. Leeds, UK), Ralf Peeters (Univ. Maastricht, NL), Edward Saff (Vanderbilt University, Nashville, USA), Herbert Stahl (TFH Berlin), Maxim Yattselev (Univ. Oregon at Eugene, USA).

#### 3.3.1. Best constrained analytic approximation

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

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

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

Here  $\psi$  is a reference behaviour capturing *a priori* assumptions on the behaviour of the model off  $K$ , while  $M$  is some admissible deviation from them. The value of  $p$  reflects the type of stability which is sought and how much one wants to smoothen the data. The choice of  $L^p$  classes is well-adapted to handling pointwise measurements.

To fix terminology we refer to (P) as a *bounded extremal problem*. As shown in [40], [42], [47], for  $1 < p \leq \infty$ , the solution to this convex infinite-dimensional optimization problem can be obtained upon iterating with respect to a Lagrange parameter the solution to spectral equations for some appropriate Hankel and Toeplitz operators. These equations in turn involve the solution to the standard extremal problem below best approximation problem ( $P_0$ ) below [61]:

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

The case  $p = 1$  of (P) is essentially open.

Various modifications of (P) have been studied in order to meet specific needs. For instance when dealing with loss-less transfer functions (see section 4.3 ), one may want to express the constraint on  $T \setminus K$  in a pointwise manner:  $|g - \psi| \leq M$  a.e. on  $T \setminus K$ , see [43]. In this form, it comes close to (but still is different from)  $H^\infty$  frequency optimization methods for control [64], [75].

The analog of problem (P) on an annulus,  $K$  being now the outer boundary, can be seen as a means to regularize a classical inverse problem occurring in nondestructive control, namely recovering a harmonic function on the inner boundary from Dirichlet-Neumann data on the outer boundary (see sections 3.2.1, 4.2, 6.1.1, 6.2). For  $p = 2$  the solution is analysed in [66]. It may serve as a tool to approach Bernoulli type problems where we are given data on the outer boundary and we seek the inner boundary, knowing it is a level curve of the flux. Then, the Lagrange parameter indicates which deformation should be applied on the inner contour in order to improve data fitting.

This is discussed in sections 3.2.1 and 6.2 for more general equations than the Laplacian, namely isotropic conductivity equations of the form  $\operatorname{div}(\sigma \nabla u) = 0$  where  $\sigma$  is non constant. In this case Hardy spaces in problem (P) become those of a so-called conjugate or real Beltrami equation [65], which are studied for  $1 < p < \infty$  in [5], [35]. Expansions of solutions needed to constructively handle such issues have been carried out in [58].

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

When  $\Omega$  is a ball or a half-space, a convenient substitute of holomorphic Hardy spaces is provided by Stein-Weiss Hardy spaces of harmonic gradients [78]. Conformal maps are no longer available in  $\mathbb{R}^n$  for  $n > 2$  and other geometries have not been much studied so far. On the ball, the analog of problem (P) is

(P1) Let  $1 \leq p \leq \infty$  and  $B \subset \mathbb{R}^n$  the unit ball. Fix  $O$  an open subset of the unit sphere  $S \subset \mathbb{R}^n$ . Let further  $V \in L^p(O)$  and  $W \in L^p(S \setminus O)$  be  $\mathbb{R}^n$ -valued vector fields, and  $M > 0$ ; find a harmonic gradient  $G \in H^p(B)$  such that  $\|G - W\|_{L^p(S \setminus O)} \leq M$  and  $G - V$  is of minimal norm in  $L^p(O)$  under this constraint.

When  $p = 2$ , spherical harmonics offer a reasonable substitute to Fourier expansions and problem (P1) was solved in [2], together with its natural analog on a shell. The solution generalizes the Toeplitz operator approach to bounded extremal problems [40], and constructive aspects of the procedure (harmonic 3-D projection, Kelvin and Riesz transformation, spherical harmonics) were derived. An important ingredient is a refinement of the Hodge decomposition allowing us to express a  $\mathbb{R}^n$ -valued vector field in  $L^p(S)$ ,  $1 < p < \infty$ , as the sum of a vector field in  $H(B)$ , a vector field in  $H^p(\mathbb{R}^n \setminus \bar{B})$ , and a tangential divergence free vector field. If  $p = 1$  or  $p = \infty$ ,  $L^p$  must be replaced respectively by the real Hardy space  $H^1$  and the bounded mean oscillation space  $BMO$ , and  $H^\infty$  should be modified accordingly. This decomposition was fully discussed in [37] (for the case of the half-space) where it plays a fundamental role.

Problem (P1) is still under investigation in the case  $p = \infty$ , where even the case where  $O = S$  is pending because a substitute of the Adamjan-Arov-Krein theory [72] is still to be built in dimension greater than 2.

Such problems arise in connection with source recovery in electro/magneto encephalography and paleomagnetism, as discussed in sections 3.2.1 and 4.2.

### 3.3.2. Best meromorphic and rational approximation

The techniques explained in this section are used to solve step 2 in section 3.2 via conformal mapping and subsequently instrumental to approach inverse boundary value problems for Poisson equation  $\Delta u = \mu$ , where  $\mu$  is some (unknown) distribution.

#### 3.3.2.1. Scalar meromorphic and rational approximation

Let as before  $D$  designate the unit disk, and  $T$  the unit circle. We further put  $R_N$  for the set of rational functions with at most  $N$  poles in  $D$ , which allows us to define the meromorphic functions in  $L^p(T)$  as the traces of functions in  $H^p + R_N$ .

A natural generalization of problem (P<sub>0</sub>) is:

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

Only for  $p = \infty$  and continuous  $f$  it is known how to solve  $(P_N)$  in closed form. The unique solution is given by AAK theory (named after Adamjan, Arov and Krein), that connects the spectral decomposition of Hankel operators with best approximation in Hankel norm [72]. This theory allows one to express  $g_N$  in terms of the singular vectors of the Hankel operator with symbol  $f$ . The continuity of  $g_N$  as a function of  $f$  only holds for stronger norms than uniform.

The case  $p = 2$  is of special importance. In particular when  $f \in \overline{H}^2$ , the Hardy space of exponent 2 of the complement of  $D$  in the complex plane (by definition,  $h(z)$  belongs to  $\overline{H}^p$  if, and only if  $h(1/z)$  belongs to  $H^p$ ), then  $(P_N)$  reduces to rational approximation. Moreover, it turns out that the associated solution  $g_N \in R_N$  has no pole outside  $D$ , hence it is a *stable* rational approximant to  $f$ . However, in contrast with the situation when  $p = \infty$ , this approximant may *not* be unique.

The former Miaou project (predecessor of Apics) has designed an adapted steepest-descent algorithm for the case  $p = 2$  whose convergence to a *local minimum* is guaranteed; until now it seems to be the only procedure meeting this property. Roughly speaking, it is a gradient algorithm that proceeds recursively with respect to the order  $N$  of the approximant, in a compact region of the parameter space [34]. Although it has proved effective in all applications carried out so far (see sections 4.2, 4.3), it is not known whether the absolute minimum can always be obtained by choosing initial conditions corresponding to *critical points* of lower degree (as is done by the RARL2 software, section 5.1).

In order to establish global convergence results, APICS has undertaken a long-term study of the number and nature of critical points, in which tools from differential topology and operator theory team up with classical approximation theory. The main discovery is that the nature of the critical points (*e.g.*, local minima, saddles...) depends on the decrease of the interpolation error to  $f$  as  $N$  increases [44]. Based on this, sufficient conditions have been developed for a local minimum to be unique. These conditions are hard to use in practice because they require strong estimates of the approximation error. These are often difficult to obtain for a given function, and are usually only valid for large  $N$ . Examples where uniqueness or asymptotic uniqueness has been proved this way include transfer functions of relaxation systems (*i.e.* Markov functions) [48] and more generally Cauchy integrals over hyperbolic geodesic arcs [50] and certain entire functions [46].

An analog to AAK theory has been carried out for  $2 \leq p < \infty$  [47]. Although not computationally as powerful, it can be used to derive lower bounds and helps analysing the behaviour of poles. When  $1 \leq p < 2$ , problem  $(P_N)$  is still fairly open.

A common feature to all these problems is that critical point equations express non-Hermitian orthogonality relations for the denominator of the approximant. This makes connection with interpolation theory [51][7] and is used in an essential manner to assess the behaviour of the poles of the approximants to functions with branchpoint-type singularities, which is of particular interest for inverse source problems (*cf.* sections 5.6 and 6.1).

In higher dimensions, the analog of problem  $(P_N)$  is the approximation of a vector field with gradients of potentials generated by  $N$  point masses instead of meromorphic functions. The issue is by no means understood at present, and is a major endeavour of future research problems.

Certain constrained rational approximation problems, of special interest in identification and design of passive systems, arise when putting additional requirements on the approximant, for instance that it should be smaller than 1 in modulus. Such questions have become over years an increasingly significant part of the team's activity (see section 4.3). For instance, convergence properties of multipoint Schur approximants, which are rational interpolants preserving contractivity of a function, were analysed in [3]. Such approximants are useful in prediction theory of stochastic processes, but since they interpolate inside the domain of holomorphy they are of limited use in frequency design.

In another connection, the generalization to several arcs of classical Zolotarev problems [74] is an achievement by the team which is useful for multiband synthesis [11]. Still, though the modulus of the response is the first concern in filter design, variation of the phase must nevertheless remain under control to avoid unacceptable distortion of the signal. This specific but important issue has less structure and was approached

using constrained optimization; a dedicated code has been developed under contract with the CNES (see section 5.5).

### 3.3.2.2. Matrix-valued rational approximation

Matrix-valued approximation is necessary for handling systems with several inputs and outputs, and it generates substantial additional difficulties with respect to scalar approximation, theoretically as well as algorithmically. In the matrix case, the McMillan degree (*i.e.* the degree of a minimal realization in the System-Theoretic sense) generalizes the degree.

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

The scalar approximation algorithm [34], mentioned in section 3.3.2.1, generalizes to the matrix-valued situation [60]. The first difficulty here consists in the parametrization of transfer matrices of given McMillan degree  $n$ , and the inner matrices (*i.e.* matrix-valued functions that are analytic in the unit disk and unitary on the circle) of degree  $n$ . The latter enter the picture in an essential manner as they play the role of the denominator in a fractional representation of transfer matrices (using the so-called Douglas-Shapiro-Shields factorization). The set of inner matrices of given degree has the structure of a smooth manifold that allows one to use differential tools as in the scalar case. In practice, one has to produce an atlas of charts (parametrization valid in a neighborhood of a point), and we must handle changes of charts in the course of the algorithm. Such parametrization can be obtained from interpolation theory and Schur type algorithms, the parameters being interpolation vectors or matrices ([31], [10], [12]). Some of them are particularly interesting to compute realizations and achieve filter synthesis ([10] [12]). Rational approximation software codes have been developed in the team (see sections 5.1).

Difficulties relative to multiple local minima naturally arise in the matrix-valued case as well, and deriving criteria that guarantee uniqueness is even more difficult than in the scalar case. The case of rational functions of sought degree or small perturbations thereof (the consistency problem) was solved in [45]. The case of matrix-valued Markov functions, the first example beyond rational functions, was treated in [33].

Let us stress that the algorithms mentioned above are first to handle rational approximation in the matrix case in a way that converges to local minima, while meeting stability constraints on the approximant.

### 3.3.3. Behavior of poles of meromorphic approximants

**Participant:** Laurent Baratchart.

The following people collaborate with us on this subject: Herbert Stahl (TFH Berlin), Maxim Yattselev (Univ. Oregon at Eugene, USA).

We refer here to the behaviour of poles of best meromorphic approximants, in the  $L^p$ -sense on a closed curve, to functions  $f$  defined as Cauchy integrals of complex measures whose support lies inside the curve. If one normalizes the contour to be the unit circle  $T$ , we are back to the framework of section 3.3.2.1 and to problem  $(P_N)$ ; invariance of the problem under conformal mapping was established in [6]. Research so far has focused on functions whose singular set inside the contour is zero or one-dimensional.

Generally speaking, the behaviour of poles is particularly important in meromorphic approximation to obtain error rates as the degree goes large and to tackle constructive issues like uniqueness. As explained in section 3.2.1, we consider this issue in connection with approximation of the solution to a Dirichlet-Neumann problem, so as to extract information on the singularities. The general theme is thus *how do the singularities of the approximant reflect those of the approximated function?* This approach to inverse problem for the 2-D Laplacian turns out to be attractive when singularities are zero- or one-dimensional (see section 4.2). It can be used as a computationally cheap initialization of more precise but heavier numerical optimizations.

As regards crack detection or source recovery, the approach in question boils down to analysing the behaviour of best meromorphic approximants of a function with branch points. For piecewise analytic cracks, or in the case of sources, We were able to prove ([6], [38], [14]) that the poles of the approximants accumulate on some extremal contour of minimum weighted energy linkings the singular points of the crack, or the sources [41]. Moreover, the asymptotic density of the poles turns out to be the Green equilibrium distribution of this contour in  $D$ , hence puts heavy charge around the singular points (in particular at the endpoints) which are therefore well localized if one is able to approximate in sufficiently high degree (this is where the method could fail).

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

It is interesting that inverse source problems inside a sphere or an ellipsoid in 3-D can be attacked with the above 2-D techniques, as applied to planar sections (see section 6.1 ).

### 3.3.4. Miscellaneous

**Participant:** Sylvain Chevillard.

Sylvain Chevillard, joined team in November 2010. His coming resulted in APICS hosting a research activity in certified computing, centered around the software *Sollya* of which S. Chevillard is a co-author, see section 5.7 . On the one hand, *Sollya* is an Inria software which still requires some tuning to a growing community of users. On the other hand, approximation-theoretic methods at work in *Sollya* are potentially useful for certified solutions to constrained analytic problems described in section 3.3.1 . However, developing *Solya* is not a long-term objective of APICS.



## BIPOP Project-Team

### 3. Scientific Foundations

#### 3.1. Dynamic non-regular systems

mechanical systems, impacts, unilateral constraints, complementarity, modeling, analysis, simulation, control, convex analysis

Dynamical systems (we limit ourselves to finite-dimensional ones) are said to be *non-regular* whenever some nonsmoothness of the state arises. This nonsmoothness may have various roots: for example some outer impulse, entailing so-called *differential equations with measure*. An important class of such systems can be described by the complementarity system

$$\left\{ \begin{array}{l} \dot{x} = f(x, u, \lambda), \\ 0 \leq y \perp \lambda \geq 0, \\ g(y, \lambda, x, u, t) = 0, \\ \text{re-initialization law of the state } x(\cdot), \end{array} \right. \quad (19)$$

where  $\perp$  denotes orthogonality;  $u$  is a control input. Now (1) can be viewed from different angles.

- Hybrid systems: it is in fact natural to consider that (1) corresponds to different models, depending whether  $y_i = 0$  or  $y_i > 0$  ( $y_i$  being a component of the vector  $y$ ). In some cases, passing from one mode to the other implies a jump in the state  $x$ ; then the continuous dynamics in (1) may contain distributions.
- Differential inclusions:  $0 \leq y \perp \lambda \geq 0$  is equivalent to  $-\lambda \in N_K(y)$ , where  $K$  is the nonnegative orthant and  $N_K(y)$  denotes the normal cone to  $K$  at  $y$ . Then it is not difficult to reformulate (1) as a differential inclusion.
- Dynamic variational inequalities: such a formalism reads as  $\langle \dot{x}(t) + F(x(t), t), v - x(t) \rangle \geq 0$  for all  $v \in K$  and  $x(t) \in K$ , where  $K$  is a nonempty closed convex set. When  $K$  is a polyhedron, then this can also be written as a complementarity system as in (1).

Thus, the 2nd and 3rd lines in (1) define the modes of the hybrid systems, as well as the conditions under which transitions occur from one mode to another. The 4th line defines how transitions are performed by the state  $x$ . There are several other formalisms which are quite related to complementarity. A tutorial-survey paper has been published [4], whose aim is to introduce the dynamics of complementarity systems and the main available results in the fields of mathematical analysis, analysis for control (controllability, observability, stability), and feedback control.

#### 3.2. Nonsmooth optimization

optimization, numerical algorithm, convexity, Lagrangian relaxation, combinatorial optimization.

Here we are dealing with the minimization of a function  $f$  (say over the whole space  $\mathbb{R}^n$ ), whose derivatives are discontinuous. A typical situation is when  $f$  comes from dualization, if the primal problem is not strictly convex – for example a large-scale linear program – or even nonconvex – for example a combinatorial optimization problem. Also important is the case of spectral functions, where  $f(x) = F(\lambda(A(x)))$ ,  $A$  being a symmetric matrix and  $\lambda$  its spectrum.

---

For these types of problems, we are mainly interested in developing efficient resolution algorithms. Our basic tool is bundling (Chap. XV of [10]) and we act along two directions:

- To explore application areas where nonsmooth optimization algorithms can be applied, possibly after some tailoring. A rich field of such application is combinatorial optimization, with all forms of relaxation [12], [11].
- To explore the possibility of designing more sophisticated algorithms. This implies an appropriate generalization of second derivatives when the first derivative does not exist, and we use advanced tools of nonsmooth analysis, for example [13].

## COMMANDS Project-Team

### 3. Scientific Foundations

#### 3.1. Historical aspects

The roots of deterministic optimal control are the “classical” theory of the calculus of variations, illustrated by the work of Newton, Bernoulli, Euler, and Lagrange (whose famous multipliers were introduced in [69]), with improvements due to the “Chicago school”, Bliss [44] during the first part of the 20th century, and by the notion of relaxed problem and generalized solution (Young [77]).

*Trajectory optimization* really started with the spectacular achievement done by Pontryagin’s group [75] during the fifties, by stating, for general optimal control problems, nonlocal optimality conditions generalizing those of Weierstrass. This motivated the application to many industrial problems (see the classical books by Bryson and Ho [50], Leitmann [71], Lee and Markus [70], Ioffe and Tihomirov [66]). Since then, various theoretical achievements have been obtained by extending the results to nonsmooth problems, see Aubin [40], Clarke [51], Ekeland [58].

*Dynamic programming* was introduced and systematically studied by R. Bellman during the fifties. The HJB equation, whose solution is the value function of the (parameterized) optimal control problem, is a variant of the classical Hamilton-Jacobi equation of mechanics for the case of dynamics parameterized by a control variable. It may be viewed as a differential form of the dynamic programming principle. This nonlinear first-order PDE appears to be well-posed in the framework of *viscosity solutions* introduced by Crandall and Lions [53], [54], [52]. These tools also allow to perform the numerical analysis of discretization schemes. The theoretical contributions in this direction did not cease growing, see the books by Barles [42] and Bardi and Capuzzo-Dolcetta [41].

#### 3.2. Trajectory optimization

The so-called *direct methods* consist in an optimization of the trajectory, after having discretized time, by a nonlinear programming solver that possibly takes into account the dynamic structure. So the two main problems are the choice of the discretization and the nonlinear programming algorithm. A third problem is the possibility of refinement of the discretization once after solving on a coarser grid.

In the *full discretization approach*, general Runge-Kutta schemes with different values of control for each inner step are used. This allows to obtain and control high orders of precision, see Hager [62], Bonnans [47]. In an interior-point algorithm context, controls can be eliminated and the resulting system of equation is easily solved due to its band structure. Discretization errors due to constraints are discussed in Dontchev et al. [57]. See also Malanowski et al. [72].

In the *indirect* approach, the control is eliminated thanks to Pontryagin’s maximum principle. One has then to solve the two-points boundary value problem (with differential variables state and costate) by a single or multiple shooting method. The questions are here the choice of a discretization scheme for the integration of the boundary value problem, of a (possibly globalized) Newton type algorithm for solving the resulting finite dimensional problem in  $IR^n$  ( $n$  is the number of state variables), and a methodology for finding an initial point.

For state constrained problems or singular arcs, the formulation of the shooting function may be quite elaborate [45], [46], [39]. As initiated in [61], we focus more specifically on the handling of discontinuities, with ongoing work on the geometric integration aspects (Hamiltonian conservation).

### 3.3. Hamilton-Jacobi-Bellman approach

This approach consists in calculating the value function associated with the optimal control problem, and then synthesizing the feedback control and the optimal trajectory using Pontryagin's principle. The method has the great particular advantage of reaching directly the global optimum, which can be very interesting, when the problem is not convex.

*Characterization of the value function* From the dynamic programming principle, we derive a characterization of the value function as being a solution (in viscosity sense) of an Hamilton-Jacobi-Bellman equation, which is a nonlinear PDE of dimension equal to the number  $n$  of state variables. Since the pioneer works of Crandall and Lions [53], [54], [52], many theoretical contributions were carried out, allowing an understanding of the properties of the value function as well as of the set of admissible trajectories. However, there remains an important effort to provide for the development of effective and adapted numerical tools, mainly because of numerical complexity (complexity is exponential with respect to  $n$ ).

*Numerical approximation for continuous value function* Several numerical schemes have been already studied to treat the case when the solution of the HJB equation (the value function) is continuous. Let us quote for example the Semi-Lagrangian methods [60], [59] studied by the team of M. Falcone (La Sapienza, Rome), the high order schemes WENO, ENO, Discrete galerkin introduced by S. Osher, C.-W. Shu, E. Harten [63], [64], [65], [73], and also the schemes on nonregular grids by R. Abgrall [38], [37]. All these schemes rely on finite differences or/and interpolation techniques which lead to numerical diffusions. Hence, the numerical solution is unsatisfying for long time approximations even in the continuous case.

One of the (nonmonotone) schemes for solving the HJB equation is based on the Ultrabee algorithm proposed, in the case of advection equation with constant velocity, by Roe [76] and recently revisited by Després-Lagoutière [56], [55]. The numerical results on several academic problems show the relevance of the antidiffusive schemes. However, the theoretical study of the convergence is a difficult question and is only partially done.

*Optimal stochastic control problems* occur when the dynamical system is uncertain. A decision typically has to be taken at each time, while realizations of future events are unknown (but some information is given on their distribution of probabilities). In particular, problems of economic nature deal with large uncertainties (on prices, production and demand). Specific examples are the portfolio selection problems in a market with risky and non-risky assets, super-replication with uncertain volatility, management of power resources (dams, gas). Air traffic control is another example of such problems.

*Nonsmoothness of the value function.* Sometimes the value function is smooth (e.g. in the case of Merton's portfolio problem, Oksendal [78]) and the associated HJB equation can be solved explicitly. Still, the value function is not smooth enough to satisfy the HJB equation in the classical sense. As for the deterministic case, the notion of viscosity solution provides a convenient framework for dealing with the lack of smoothness, see Pham [74], that happens also to be well adapted to the study of discretization errors for numerical discretization schemes [67], [43].

*Numerical approximation for optimal stochastic control problems.* The numerical discretization of second order HJB equations was the subject of several contributions. The book of Kushner-Dupuis [68] gives a complete synthesis on the Markov chain schemes (i.e Finite Differences, semi-Lagrangian, Finite Elements, ...). Here a main difficulty of these equations comes from the fact that the second order operator (i.e. the diffusion term) is not uniformly elliptic and can be degenerated. Moreover, the diffusion term (covariance matrix) may change direction at any space point and at any time (this matrix is associated the dynamics volatility).

For solving stochastic control problems, we studied the so-called Generalized Finite Differences (GFD), that allow to choose at any node, the stencil approximating the diffusion matrix up to a certain threshold [49]. Determining the stencil and the associated coefficients boils down to a quadratic program to be solved at each point of the grid, and for each control. This is definitely expensive, with the exception of special structures where the coefficients can be computed at low cost. For two dimensional systems, we designed a (very) fast algorithm for computing the coefficients of the GFD scheme, based on the Stern-Brocot tree [48].

---

## CORIDA Project-Team

### 3. Scientific Foundations

#### 3.1. Analysis and control of fluids and of fluid-structure interactions

**Participants:** Thomas Chambrion, Antoine Henrot, Alexandre Munnier, Lionel Rosier, Jean-François Scheid, Takeo Takahashi, Marius Tucsnak, Jean-Claude Vivalda.

The problems we consider are modeled by the Navier-Stokes, Euler or Korteweg de Vries equations (for the fluid) coupled to the equations governing the motion of the solids. One of the main difficulties of this problem comes from the fact that the domain occupied by the fluid is one of the unknowns of the problem. We have thus to tackle a *free boundary problem*.

The control of fluid flows is a major challenge in many applications: aeronautics, pollution issues, regulation of irrigation channels or of the flow in pipelines, etc. All these problems cannot be easily reduced to finite dimensional models so a methodology of analysis and control based on PDE's is an essential issue. In a first approximation the motion of fluid and of the solids can be decoupled. The most used models for an incompressible fluid are given by the Navier-Stokes or by the Euler equations.

The optimal open loop control approach of these models has been developed from both the theoretical and numerical points of view. Controllability issues for the equations modeling the fluid motion are by now well understood (see, for instance, Imanuvilov [59] and the references therein). The feedback control of fluid motion has also been recently investigated by several research teams (see, for instance Barbu [54] and references therein) but this field still contains an important number of open problems (in particular those concerning observers and implementation issues). One of our aims is to develop efficient tools for computing feedback laws for the control of fluid systems.

In real applications the fluid is often surrounded by or it surrounds an elastic structure. In the above situation one has to study fluid-structure interactions. This subject has been intensively studied during the last years, in particular for its applications in noise reduction problems, in lubrication issues or in aeronautics. In this kind of problems, a PDE's system modeling the fluid in a cavity (Laplace equation, wave equation, Stokes, Navier-Stokes or Euler systems) is coupled to the equations modeling the motion of a part of the boundary. The difficulties of this problem are due to several reasons such as the strong nonlinear coupling and the existence of a free boundary. This partially explains the fact that applied mathematicians have only recently tackled these problems from either the numerical or theoretical point of view. One of the main results obtained in our project concerns the global existence of weak solutions in the case of a two-dimensional Navier-Stokes fluid (see [8]). Another important result gives the existence and the uniqueness of strong solutions for two or three-dimensional Navier-Stokes fluid (see [9]). In that case, the solution exists as long as there is no contact between rigid bodies, and for small data in the three-dimensional case.

#### 3.2. Frequency domain methods for the analysis and control of systems governed by PDE's

**Participants:** Xavier Antoine, Bruno Pinçon, Karim Ramdani, Bertrand Thierry.

We use frequency tools to analyze different types of problems. The first one concerns the control, the optimal control and the stabilization of systems governed by PDE's, and their numerical approximations. The second one concerns time-reversal phenomena, while the last one deals with numerical approximation of high-frequency scattering problems.

### 3.2.1. Control and stabilization for skew-adjoint systems

The first area concerns theoretical and numerical aspects in the control of a class of PDE's. More precisely, in a semigroup setting, the systems we consider have a skew-adjoint generator. Classical examples are the wave, the Bernoulli-Euler or the Schrödinger equations. Our approach is based on an original characterization of exact controllability of second order conservative systems proposed by K. Liu [63]. This characterization can be related to the Hautus criterion in the theory of finite dimensional systems (cf. [58]). It provides for time-dependent problems exact controllability criteria **that do not depend on time, but depend on the frequency variable** conjugated to time. Studying the controllability of a given system amounts then to establishing uniform (with respect to frequency) estimates. In other words, the problem of exact controllability for the wave equation, for instance, comes down to a high-frequency analysis for the Helmholtz operator. This frequency approach has been proposed first by K. Liu for bounded control operators (corresponding to internal control problems), and has been recently extended to the case of unbounded control operators (and thus including boundary control problems) by L. Miller [64]. Using the result of Miller, K. Ramdani, T. Takahashi, M. Tucsnak have obtained in [5] a new spectral formulation of the criterion of Liu [63], which is valid for boundary control problems. This frequency test can be seen as an observability condition for packets of eigenvectors of the operator. This frequency test has been successfully applied in [5] to study the exact controllability of the Schrödinger equation, the plate equation and the wave equation in a square. Let us emphasize here that one further important advantage of this frequency approach lies in the fact that it can also be used for the analysis of space semi-discretized control problems (by finite element or finite differences). The estimates to be proved must then be uniform with respect to **both the frequency and the mesh size**.

In the case of finite dimensional systems one of the main applications of frequency domain methods consists in designing robust controllers, in particular of  $H^\infty$  type. Obtaining the similar tools for systems governed by PDE's is one of the major challenges in the theory of infinite dimensional systems. The first difficulty which has to be tackled is that, even for very simple PDE systems, no method giving the parametrisation of all stabilizing controllers is available. One of the possible remedies consists in considering known families of stabilizing feedback laws depending on several parameters and in optimizing the  $H^\infty$  norm of an appropriate transfer function with respect to this parameters. Such families of feedback laws yielding computationally tractable optimization problems are now available for systems governed by PDE's in one space dimension.

### 3.2.2. Time-reversal

The second area in which we make use of frequency tools is the analysis of time-reversal for harmonic acoustic waves. This phenomenon described in Fink [56] is a direct consequence of the reversibility of the wave equation in a non dissipative medium. It can be used to **focus an acoustic wave** on a target through a complex and/or unknown medium. To achieve this, the procedure followed is quite simple. First, time-reversal mirrors are used to generate an incident wave that propagates through the medium. Then, the mirrors measure the acoustic field diffracted by the targets, time-reverse it and back-propagate it in the medium. Iterating the scheme, we observe that the incident wave emitted by the mirrors focuses on the scatterers. An alternative and more original focusing technique is based on the so-called D.O.R.T. method [57]. According to this experimental method, the eigenelements of the time-reversal operator contain important information on the propagation medium and on the scatterers contained in it. More precisely, the number of nonzero eigenvalues is exactly the number of scatterers, while each eigenvector corresponds to an incident wave that selectively focuses on each scatterer.

Time-reversal has many applications covering a wide range of fields, among which we can cite **medicine** (kidney stones destruction or medical imaging), **sub-marine communication** and **non destructive testing**. Let us emphasize that in the case of time-harmonic acoustic waves, time-reversal is equivalent to phase conjugation and involves the Helmholtz operator.

In [2], we proposed the first far field model of time reversal in the time-harmonic case.

### 3.2.3. Numerical approximation of high-frequency scattering problems

This subject deals mainly with the numerical solution of the Helmholtz or Maxwell equations for open region scattering problems. This kind of situation can be met e.g. in radar systems in electromagnetism or in acoustics for the detection of underwater objects like submarines.

Two particular difficulties are considered in this situation

- the wavelength of the incident signal is small compared to the characteristic size of the scatterer,
- the problem is set in an unbounded domain.

These two problematics limit the application range of most common numerical techniques. The aim of this part is to develop new numerical simulation techniques based on microlocal analysis for modeling the propagation of rays. The importance of microlocal techniques in this situation is that it makes possible a local analysis both in the spatial and frequency domain. Therefore, it can be seen as a kind of asymptotic theory of rays which can be combined with numerical approximation techniques like boundary element methods. The resulting method is called the On-Surface Radiation Condition method.

### 3.3. Observability, controllability and stabilization in the time domain

**Participants:** Fatiha Alabau, Xavier Antoine, Thomas Chambrión, Antoine Henrot, Karim Ramdani, Marius Tucsnak, Jean-Claude Vivalda.

Controllability and observability have been set at the center of control theory by the work of R. Kalman in the 1960's and soon they have been generalized to the infinite-dimensional context. The main early contributors have been D.L. Russell, H. Fattorini, T. Seidman, R. Triggiani, W. Littman and J.-L. Lions. The latter gave the field an enormous impact with his book [61], which is still a main source of inspiration for many researchers. Unlike in classical control theory, for infinite-dimensional systems there are many different (and not equivalent) concepts of controllability and observability. The strongest concepts are called exact controllability and exact observability, respectively. In the case of linear systems exact controllability is important because it guarantees stabilizability and the existence of a linear quadratic optimal control. Dually, exact observability guarantees the existence of an exponentially converging state estimator and the existence of a linear quadratic optimal filter. An important feature of infinite dimensional systems is that, unlike in the finite dimensional case, the conditions for exact observability are no longer independent of time. More precisely, for simple systems like a string equation, we have exact observability only for times which are large enough. For systems governed by other PDE's (like dispersive equations) the exact observability in arbitrarily small time has been only recently established by using new frequency domain techniques. A natural question is to estimate the energy required to drive a system in the desired final state when the control time goes to zero. This is a challenging theoretical issue which is critical for perturbation and approximation problems. In the finite dimensional case this issue has been first investigated in Seidman [66]. In the case of systems governed by linear PDE's some similar estimates have been obtained only very recently (see, for instance Miller [64]). One of the open problems of this field is to give sharp estimates of the observability constants when the control time goes to zero.

Even in the finite-dimensional case, despite the fact that the linear theory is well established, many challenging questions are still open, concerning in particular nonlinear control systems.

In some cases it is appropriate to regard external perturbations as unknown inputs; for these systems the synthesis of observers is a challenging issue, since one cannot take into account the term containing the unknown input into the equations of the observer. While the theory of observability for linear systems with unknown inputs is well established, this is far from being the case in the nonlinear case. A related active field of research is the uniform stabilization of systems with time-varying parameters. The goal in this case is to stabilize a control system with a control strategy independent of some signals appearing in the dynamics, i.e., to stabilize simultaneously a family of time-dependent control systems and to characterize families of control systems that can be simultaneously stabilized.

One of the basic questions in finite- and infinite-dimensional control theory is that of motion planning, i.e., the explicit design of a control law capable of driving a system from an initial state to a prescribed final one. Several techniques, whose suitability depends strongly on the application which is considered, have been and are being developed to tackle such a problem, as for instance the continuation method, flatness, tracking or optimal control. Preliminary to any question regarding motion planning or optimal control is the issue of controllability, which is not, in the general nonlinear case, solved by the verification of a simple algebraic criterion. A further motivation to study nonlinear controllability criteria is given by the fact that techniques developed in the domain of (finite-dimensional) geometric control theory have been recently applied successfully to study the controllability of infinite-dimensional control systems, namely the Navier–Stokes equations (see Agrachev and Sarychev [53]).

### **3.4. Implementation**

This is a transverse research axis since all the research directions presented above have to be validated by giving control algorithms which are aimed to be implemented in real control systems. We stress below some of the main points which are common (from the implementation point of view) to the application of the different methods described in the previous sections.

For many infinite dimensional systems the use of co-located actuators and sensors and of simple proportional feed-back laws gives satisfying results. However, for a large class of systems of interest it is not clear that these feedbacks are efficient, or the use of co-located actuators and sensors is not possible. This is why a more general approach for the design of the feedbacks has to be considered. Among the techniques in finite dimensional systems theory those based on the solutions of infinite dimensional Riccati equation seem the most appropriate for a generalization to infinite dimensional systems. The classical approach is to approximate an LQR problem for a given infinite dimensional system by finite dimensional LQR problems. As it has been already pointed out in the literature this approach should be carefully analyzed since, even for some very simple examples, the sequence of feedbacks operators solving the finite dimensional LQR is not convergent. Roughly speaking this means that by refining the mesh we obtain a closed loop system which is not exponentially stable (even if the corresponding infinite dimensional system is theoretically stabilized). In order to overcome this difficulty, several methods have been proposed in the literature : filtering of high frequencies, multigrid methods or the introduction of a numerical viscosity term. We intend to first apply the numerical viscosity method introduced in Tcheougoue Tebou – Zuazua [67], for optimal and robust control problems.



---

## DISCO Project-Team

### 3. Scientific Foundations

#### 3.1. Modeling of complex environment

We want to model phenomena such as a temporary loss of connection (e.g. synchronisation of the movements through haptic interfaces), a nonhomogeneous environment (e.g. case of cryogenic systems) or the presence of the human factor in the control loop (e.g. grid systems) but also problems involved with technological constraints (e.g. range of the sensors). The mathematical models concerned include integro-differential, partial differential equations, algebraic inequalities with the presence of several time scales, whose variables and/or parameters must satisfy certain constraints (for instance, positivity).

#### 3.2. Analysis of interconnected systems

- Algebraic analysis of linear systems

Study of the structural properties of linear differential time-delay systems and linear infinite-dimensional systems (e.g. invariants, controllability, observability, flatness, reductions, decomposition, decoupling, equivalences) by means of constructive algebra, module theory, homological algebra, algebraic analysis and symbolic computation [8], [9], [85], [106], [87], [90].

- Robust stability of linear systems

Within an interconnection context, lots of phenomena are modelled directly or after an approximation by delay systems. These systems might have fixed delays, time-varying delays, distributed delays...

For various infinite-dimensional systems, particularly delay and fractional systems, input-output and time-domain methods are jointly developed in the team to characterize stability. This research is developed at four levels: analytic approaches ( $H_\infty$ -stability, BIBO-stability, robust stability, robustness metrics) [1], [2], [5], [6], symbolic computation approaches (SOS methods are used for determining easy-to-check conditions which guarantee that the poles of a given linear system are not in the closed right half-plane, certified CAD techniques), numerical approaches (root-loci, continuation methods) and by means of softwares developed in the team [5], [6].

- Robustness/fragility of biological systems

Deterministic biological models describing, for instance, species interactions, are frequently composed of equations with important disturbances and poorly known parameters. To evaluate the impact of the uncertainties, we use the techniques of designing of global strict Lyapunov functions or functional developed in the team.

However, for other biological systems, the notion of robustness may be different and this question is still in its infancy (see, e.g. [98]). Unlike engineering problems where a major issue is to maintain stability in the presence of disturbances, a main issue here is to maintain the system response in the presence of disturbances. For instance, a biological network is required to keep its functioning in case of a failure of one of the nodes in the network. The team, which has a strong expertise in robustness for engineering problems, aims at contributing at the development of new robustness metrics in this biological context.

#### 3.3. Stabilization of interconnected systems

- Linear systems: Analytic and algebraic approaches are considered for infinite-dimensional linear systems studied within the input-output framework.

In the recent years, the Youla-Kučera parametrization (which gives the set of all stabilizing controllers of a system in terms of its coprime factorizations) has been the cornerstone of the success of the  $H_\infty$ -control since this parametrization allows one to rewrite the problem of finding the optimal stabilizing controllers for a certain norm such as  $H_\infty$  or  $H_2$  as affine, and thus, convex problem.

A central issue studied in the team is the computation of such factorizations for a given infinite-dimensional linear system as well as establishing the links between stabilizability of a system for a certain norm and the existence of coprime factorizations for this system. These questions are fundamental for robust stabilization problems [1], [2], [8], [9].

We also consider simultaneous stabilization since it plays an important role in the study of reliable stabilization, i.e. in the design of controllers which stabilize a finite family of plants describing a system during normal operating conditions and various failed modes (e.g. loss of sensors or actuators, changes in operating points) [9]. Moreover, we investigate strongly stabilizable systems [9], namely systems which can be stabilized by stable controllers, since they have a good ability to track reference inputs and, in practice, engineers are reluctant to use unstable controllers especially when the system is stable.

- Nonlinear systems

The project aims at developing robust stabilization theory and methods for important classes of nonlinear systems that ensure good controller performance under uncertainty and time delays. The main techniques include techniques called backstepping and forwarding, constructions of strict Lyapunov functions through so-called "strictification" approaches [3] and construction of Lyapunov-Krasovskii functionals [4], [5], [6].

- Predictive control

For highly complex systems described in the time-domain and which are submitted to constraints, predictive control seems to be well-adapted. This model based control method (MPC: Model Predictive Control) is founded on the determination of an optimal control sequence over a receding horizon. Due to its formulation in the time-domain, it is an effective tool for handling constraints and uncertainties which can be explicitly taken into account in the synthesis procedure [7]. The team considers how multiparametric optimization can help to reduce the computational load of this method, allowing its effective use on real world constrained problems.

The team also investigates stochastic optimization methods such as genetic algorithm, particle swarm optimization or ant colony [10] as they can be used to optimize any criterion and constraint whatever their mathematical structure is. The developed methodologies can be used by non specialists.

### 3.4. Synthesis of reduced complexity controllers

- PID controllers

Even though the synthesis of control laws of a given complexity is not a new problem, it is still open, even for finite-dimensional linear systems. Our purpose is to search for good families of "simple" (e.g. low order) controllers for infinite-dimensional dynamical systems. Within our approach, PID candidates are first considered in the team [2], [31].

- Predictive control

The synthesis of predictive control laws is concerned with the solution of multiparametric optimization problems. Reduced order controller constraints can be viewed as non convex constraints in the synthesis procedure. Such constraints can be taken into account with stochastic algorithms.

Finally, the development of algorithms based on both symbolic computation and numerical methods, and their implementations in dedicated Scilab/Matlab/Maple toolboxes are important issues in the project.

## GECO Team

### 3. Scientific Foundations

#### 3.1. Geometric control theory

The main research topic of the project-team will be **geometric control**, with a special focus on **control design**. The application areas that we target are control of quantum mechanical systems, neurogeometry and switched systems.

Geometric control theory provides a viewpoint and several tools, issued in particular from differential geometry, to tackle typical questions arising in the control framework: controllability, observability, stabilization, optimal control... [27], [62] The geometric control approach is particularly well suited for systems involving nonlinear and nonholonomic phenomena. We recall that nonholonomicity refers to the property of a velocity constraint that is not equivalent to a state constraint.

The expression **control design** refers here to all phases of the construction of a control law, in a mainly open-loop perspective: modeling, controllability analysis, output tracking, motion planning, simultaneous control algorithms, tracking algorithms, performance comparisons for control and tracking algorithms, simulation and implementation.

We recall that

- **controllability** denotes the property of a system for which any two states can be connected by a trajectory corresponding to an admissible control law ;
- **output tracking** refers to a control strategy aiming at keeping the value of some functions of the state arbitrarily close to a prescribed time-dependent profile. A typical example is **configuration tracking** for a mechanical system, in which the controls act as forces and one prescribes the position variables along the trajectory, while the evolution of the momenta is free. One can think for instance at the lateral movement of a car-like vehicle: even if such a movement is unfeasible, it can be tracked with arbitrary precision by applying a suitable control strategy;
- **motion planning** is the expression usually denoting the algorithmic strategy for selecting one control law steering the system from a given initial state to an attainable final one;
- **simultaneous control** concerns algorithms that aim at driving the system from two different initial conditions, with the same control law and over the same time interval, towards two given final states (one can think, for instance, at some control action on a fluid whose goal is to steer simultaneously two floating bodies.) Clearly, the study of which pairs (or  $n$ -uples) of states can be simultaneously connected thanks to an admissible control requires an additional controllability analysis with respect to the plain controllability mentioned above.

At the core of control design is then the notion of motion planning. Among the motion planning methods, a preeminent role is played by those based on the Lie algebra associated with the control system ([83], [70], [76]), those exploiting the possible flatness of the system ([56]) and those based on the continuation method ([95]). Optimal control is clearly another method for choosing a control law connecting two states, although it generally introduces new computational and theoretical difficulties.

Control systems with special structure, which are very important for applications are those for which the controls appear linearly. When the controls are not bounded, this means that the admissible velocities form a distribution in the tangent bundle to the state manifold. If the distribution is equipped with a smoothly varying norm (representing a cost of the control), the resulting geometrical structure is called *sub-Riemannian*. Sub-Riemannian geometry thus appears as the underlying geometry of the nonholonomic control systems, playing the same role as Euclidean geometry for linear systems. As such, its study is fundamental for control design. Moreover its importance goes far beyond control theory and is an active field of research both in differential geometry ([82]), geometric measure theory ([57], [31]) and hypoelliptic operator theory ([43]).

Other important classes of control systems are those modeling mechanical systems. The dynamics are naturally defined on the tangent or cotangent bundle of the configuration manifold, they have Lagrangian or Hamiltonian structure, and the controls act as forces. When the controls appear linearly, the resulting model can be seen somehow as a second-order sub-Riemannian structure (see [49]).

The control design topics presented above naturally extend to the case of distributed parameter control systems. The geometric approach to control systems governed by partial differential equations is a novel subject with great potential. It could complement purely analytical and numerical approaches, thanks to its more dynamical, qualitative and intrinsic point of view. An interesting example of this approach is the paper [28] about the controllability of Navier–Stokes equation by low forcing modes.

## MAXPLUS Project-Team

### 3. Scientific Foundations

#### 3.1. L'algèbre max-plus/Max-plus algebra

Le semi-corps *max-plus* est l'ensemble  $\mathbb{R} \cup \{-\infty\}$ , muni de l'addition  $(a, b) \mapsto a \oplus b = \max(a, b)$  et de la multiplication  $(a, b) \mapsto a \otimes b = a + b$ . Cette structure algébrique diffère des structures de corps classiques par le fait que l'addition n'est pas une loi de groupe, mais est idempotente:  $a \oplus a = a$ . On rencontre parfois des variantes de cette structure: par exemple, le semi-corps *min-plus* est l'ensemble  $\mathbb{R} \cup \{+\infty\}$  muni des lois  $a \oplus b = \min(a, b)$  et  $a \otimes b = a + b$ , et le semi-anneau *tropical* est l'ensemble  $\mathbb{N} \cup \{+\infty\}$  munis des mêmes lois. L'on peut se poser la question de généraliser les constructions de l'algèbre et de l'analyse classique, qui reposent pour une bonne part sur des anneaux ou des corps tels que  $\mathbb{Z}$  ou  $\mathbb{R}$ , au cas de semi-anneaux de type max-plus: tel est l'objet de ce qu'on appelle un peu familièrement "l'algèbre max-plus".

Il est impossible ici de donner une vue complète du domaine. Nous nous bornerons à indiquer quelques références bibliographiques. L'intérêt pour les structures de type max-plus est contemporain de la naissance de la théorie des treillis [102]. Depuis, les structures de type max-plus ont été développées indépendamment par plusieurs écoles, en relation avec plusieurs domaines. Les motivations venant de la Recherche Opérationnelle (programmation dynamique, problèmes de plus court chemin, problèmes d'ordonnancement, optimisation discrète) ont été centrales dans le développement du domaine [95], [122], [173], [178], [179]. Les semi-anneaux de type max-plus sont bien sûr reliés aux algèbres de Boole [82]. L'algèbre max-plus apparaît de manière naturelle en contrôle optimal et dans la théorie des équations aux dérivées partielles d'Hamilton-Jacobi [161], [160], [144], [128], [119], [165], [138], [120], [105], [66]. Elle apparaît aussi en analyse asymptotique (asymptotiques de type WKB [143], [144], [128], grandes déviations [159], asymptotiques à température nulle en physique statistique [84]), puisque l'algèbre max-plus apparaît comme limite de l'algèbre usuelle. La théorie des opérateurs linéaires max-plus peut être vue comme faisant partie de la théorie des opérateurs de Perron-Frobenius non-linéaires, ou de la théorie des applications contractantes ou monotones sur les cônes [129], [149], [141], [73], laquelle a de nombreuses motivations, telles l'économie mathématique [146], et la théorie des jeux [162], [56]. Dans la communauté des systèmes à événements discrets, l'algèbre max-plus a été beaucoup étudiée parce qu'elle permet de représenter de manière linéaire les phénomènes de synchronisation, lesquels déterminent le comportement temporel de systèmes de production ou de réseaux, voir [6]. Parmi les développements récents du domaine, on peut citer le calcul des réseaux [83], [133], qui permet de calculer des bornes pire des cas de certaines mesures de qualité de service. En informatique théorique, l'algèbre max-plus (ou plutôt le semi-anneau tropical) a joué un rôle décisif dans la résolution de problèmes de décision en théorie des automates [168], [125], [169], [130], [151]. Notons finalement, pour information, que l'algèbre max-plus est apparue récemment en géométrie algébrique [118], [172], [145], [171] et en théorie des représentations [107], [76], sous les noms de géométrie et combinatoire tropicales.

Nous décrivons maintenant de manière plus détaillée les sujets qui relèvent directement des intérêts du projet, comme la commande optimale, les asymptotiques, et les systèmes à événements discrets.

#### *English version*

The *max-plus* semifield is the set  $\mathbb{R} \cup \{-\infty\}$ , equipped with the addition  $(a, b) \mapsto a \oplus b = \max(a, b)$  and the multiplication  $(a, b) \mapsto a \otimes b = a + b$ . This algebraic structure differs from classical structures, like fields, in that addition is idempotent:  $a \oplus a = a$ . Several variants have appeared in the literature: for instance, the *min-plus* semifield is the set  $\mathbb{R} \cup \{+\infty\}$  equipped with the laws  $a \oplus b = \min(a, b)$  and  $a \otimes b = a + b$ , and the *tropical* semiring is the set  $\mathbb{N} \cup \{+\infty\}$  equipped with the same laws. One can ask the question of extending to max-plus type structures the classical constructions and results of algebra and analysis: this is what is often called in a wide sense "max-plus algebra" or "tropical algebra".

It is impossible to give in this short space a fair view of the field. Let us, however, give a few references. The interest in max-plus type structures is contemporaneous with the early developments of lattice theory [102]. Since that time, max-plus structures have been developed independently by several schools, in relation with several fields. Motivations from Operations Research (dynamic programming, shortest path problems, scheduling problems, discrete optimisation) were central in the development of the field [95], [122], [173], [178], [179]. Of course, max-plus type semirings are related to Boolean algebras [82]. Max-plus algebras arises naturally in optimal control and in the theory of Hamilton-Jacobi partial differential equations [161], [160], [144], [128], [119], [165], [138], [120], [105], [66]. It arises in asymptotic analysis (WKB asymptotics [143], [144], [128], large deviation asymptotics [159], or zero temperature asymptotics in statistical physics [84]), since max-plus algebra appears as a limit of the usual algebra. The theory of max-plus linear operators may be thought of as a part of the non-linear Perron-Frobenius theory, or of the theory of nonexpansive or monotone operators on cones [129], [149], [141], [73], a theory with numerous motivations, including mathematical economy [146] and game theory [162], [56]. In the discrete event systems community, max-plus algebra has been much studied since it allows one to represent linearly the synchronisation phenomena which determine the time behaviour of manufacturing systems and networks, see [6]. Recent developments include the network calculus of [83], [133] which allows one to compute worst case bounds for certain measures of quality of service. In theoretical computer science, max-plus algebra (or rather, the tropical semiring) played a key role in the solution of decision problems in automata theory [168], [125], [169], [130], [151]. We finally note for information that max-plus algebra has recently arisen in algebraic geometry [118], [172], [145], [171] and in representation theory [107], [76], under the names of tropical geometry and combinatorics.

We now describe in more details some parts of the subject directly related to our interests, like optimal control, asymptotics, and discrete event systems.

### 3.2. Algèbre max-plus, programmation dynamique, et commande optimale/Max-plus algebra, dynamic programming, and optimal control

L'exemple le plus simple d'un problème conduisant à une équation min-plus linéaire est le problème classique du plus court chemin. Considérons un graphe dont les nœuds sont numérotés de 1 à  $n$  et dont le coût de l'arc allant du nœud  $i$  au nœud  $j$  est noté  $M_{ij} \in \mathbb{R} \cup \{+\infty\}$ . Le coût minimal d'un chemin de longueur  $k$ , allant de  $i$  à  $j$ , est donné par la quantité:

$$v_{ij}(k) = \min_{\ell: \ell_0=i, \ell_k=j} \sum_{r=0}^{k-1} M_{\ell_r \ell_{r+1}} \quad , \quad (20)$$

où le minimum est pris sur tous les chemins  $\ell = (\ell_0, \dots, \ell_k)$  de longueur  $k$ , de nœud initial  $\ell_0 = i$  et de nœud final  $\ell_k = j$ . L'équation classique de la programmation dynamique s'écrit:

$$v_{ij}(k) = \min_{1 \leq s \leq n} (M_{is} + v_{sj}(k-1)) \quad . \quad (21)$$

On reconnaît ainsi une équation linéaire min-plus :

$$v(k) = Mv(k-1) \quad , \quad (22)$$

où on note par la concaténation le produit matriciel induit par la structure de l'algèbre min-plus. Le classique *problème de Lagrange* du calcul des variations,

$$v(x, T) = \inf_{X(\cdot), X(0)=x} \int_0^T L(X(t), \dot{X}(t)) dt + \phi(X(T)) \quad , \quad (23)$$

où  $X(t) \in \mathbb{R}^n$ , pour  $0 \leq t \leq T$ , et  $L : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  est le Lagrangien, peut être vu comme une version continue de (1), ce qui permet de voir l'équation d'Hamilton-Jacobi que vérifie  $v$ ,

$$v(\cdot, 0) = \phi, \quad \frac{\partial v}{\partial T} + H(x, \frac{\partial v}{\partial x}) = 0, \quad H(x, p) = \sup_{y \in \mathbb{R}^n} (-p \cdot y - L(x, y)) , \quad (24)$$

comme une équation min-plus linéaire. En particulier, les solutions de (5) vérifient un principe de superposition min-plus: si  $v$  et  $w$  sont deux solutions, et si  $\lambda, \mu \in \mathbb{R}$ ,  $\inf(\lambda + v, \mu + w)$  est encore solution de (5). Ce point de vue, inauguré par Maslov, a conduit au développement de l'école d'Analyse Idempotente (voir [144], [128], [138]).

La présence d'une structure algébrique sous-jacente permet de voir les solutions stationnaires de (2) et (5) comme des vecteurs propres de la matrice  $M$  ou du semi-groupe d'évolution de l'équation d'Hamilton-Jacobi. La valeur propre associée fournit le coût moyen par unité de temps (coût ergodique). La représentation des vecteurs propres (voir [161], [173], [95], [121], [89], [72], [6] pour la dimension finie, et [144], [128] pour la dimension infinie) est intimement liée au théorème de l'autoroute qui décrit les trajectoires optimales quand la durée ou la longueur des chemins tend vers l'infini. Pour l'équation d'Hamilton-Jacobi, des résultats reliés sont apparus récemment en théorie d'"Aubry-Mather" [105].

#### English version

The most elementary example of a problem leading to a min-plus linear equation is the classical shortest path problem. Consider a graph with nodes  $1, \dots, n$ , and let  $M_{ij} \in \mathbb{R} \cup \{+\infty\}$  denote the cost of the arc from node  $i$  to node  $j$ . The minimal cost of a path of a given length,  $k$ , from  $i$  to  $j$ , is given by (1), where the minimum is taken over all paths  $\ell = (\ell_0, \dots, \ell_k)$  of length  $k$ , with initial node  $\ell_0 = i$  and final node  $\ell_k = j$ . The classical dynamic programming equation can be written as in (2). We recognise the min-plus linear equation (3), where concatenation denotes the matrix product induced by the min-plus algebraic structure. The classical *Lagrange problem* of calculus of variations, given by (4) where  $X(t) \in \mathbb{R}^n$ , for  $0 \leq t \leq T$ , and  $L : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  is the Lagrangian, may be thought of as a continuous version of (1), which allows us to see the Hamilton-Jacobi equation (5) satisfied by  $v$ , as a min-plus linear equation. In particular, the solutions of (5) satisfy a min-plus superposition principle: if  $v$  and  $w$  are two solutions, and if  $\lambda, \mu \in \mathbb{R}$ , then  $\inf(\lambda + v, \mu + w)$  is also a solution of (5). This point of view, due to Maslov, led to the development of the school of Idempotent Analysis (see [144], [128], [138]).

The underlying algebraic structure allows one to see stationary solutions of (2) and (5) as eigenvectors of the matrix  $M$  or of the evolution semigroup of the Hamilton-Jacobi equation. The associated eigenvalue gives the average cost per time unit (ergodic cost). The representation of eigenvectors (see [161], [173], [121], [89], [95], [72], [6] for the finite dimension case, and [144], [128] for the infinite dimension case) is intimately related to turnpike theorems, which describe optimal trajectories as the horizon, or path length, tends to infinity. For the Hamilton-Jacobi equation, related results have appeared recently in the "Aubry-Mather" theory [105].

### 3.3. Applications monotones et théorie de Perron-Frobenius non-linéaire, ou l'approche opératorielle du contrôle optimal et des jeux/Monotone maps and non-linear Perron-Frobenius theory, or the operator approach to optimal control and games

On sait depuis le tout début des travaux en décision markovienne que les opérateurs de la programmation dynamique  $f$  de problèmes de contrôle optimal ou de jeux (à somme nulle et deux joueurs), avec critère additif, ont les propriétés suivantes :

$$\begin{array}{ll} \text{monotonie/monotonicity} & x \leq y \Rightarrow f(x) \leq f(y) , \\ \text{contraction/nonexpansiveness} & \|f(x) - f(y)\|_\infty \leq \|x - y\|_\infty . \end{array} \quad (25)$$

Ici, l'opérateur  $f$  est une application d'un certain espace de fonctions à valeurs réelles dans lui-même,  $\leq$  désigne l'ordre partiel usuel, et  $\|\cdot\|_\infty$  désigne la norme sup. Dans le cas le plus simple, l'ensemble des états est  $\{1, \dots, n\}$  et  $f$  est une application de  $\mathbb{R}^n$  dans lui-même. Les applications monotones qui sont contractantes pour la norme du sup peuvent être vues comme des généralisations non-linéaires des matrices sous-stochastiques. Une sous-classe utile, généralisant les matrices stochastiques, est formée des applications qui sont monotones et commutent avec l'addition d'une constante [94] (celles ci sont parfois appelées fonctions topicales). Les problèmes de programmation dynamique peuvent être traduits en termes d'opérateurs : l'équation de la programmation dynamique d'un problème de commande optimale à horizon fini s'écrit en effet  $x(k) = f(x(k-1))$ , où  $x(k)$  est la fonction valeur en horizon  $k$  et  $x(0)$  est donné; la fonction valeur  $y$  d'un problème à horizon infini (y compris le cas d'un problème d'arrêt optimal) vérifie  $y = f(y)$ ; la fonction valeur  $z$  d'un problème avec facteur d'actualisation  $0 < \alpha < 1$  vérifie  $z = f(\alpha z)$ , etc. Ce point de vue abstrait a été très fructueux, voir par exemple [56]. Il permet d'inclure la programmation dynamique dans la perspective plus large de la théorie de Perron-Frobenius non-linéaire, qui, depuis l'extension du théorème de Perron-Frobenius par Krein et Rutman, traite des applications non linéaires sur des cônes vérifiant des conditions de monotonie, de contraction ou d'homogénéité. Les problèmes auxquels on s'intéresse typiquement sont la structure de l'ensemble des points fixes de  $f$ , le comportement asymptotique de  $f^k$ , en particulier l'existence de la limite de  $f^k(x)/k$  lorsque  $k$  tends vers l'infini (afin d'obtenir le coût ergodique d'un problème de contrôle optimal ou de jeux), l'asymptotique plus précise de  $f^k$ , à une normalisation près (afin d'obtenir le comportement précis de l'itération sur les valeurs), etc. Nous renvoyons le lecteur à [149] pour un panorama. Signalons que dans [111],[7], des algorithmes inspirés de l'algorithme classique d'itérations sur les politiques du contrôle stochastique ont pu être introduits dans le cas des opérateurs monotones contractants généraux, en utilisant des résultats de structure de l'ensemble des points fixes de ces opérateurs. Les applications de la théorie des applications monotones contractantes ne se limitent pas au contrôle optimal et aux jeux. En particulier, on utilise la même classe d'applications dans la modélisation des systèmes à événements discrets, voir le §3.5 ci-dessous, et une classe semblable d'applications en analyse statique de programmes, voir le §4.4 ci-dessous.

#### English version

Since the very beginning of Markov decision theory, it has been observed that dynamic programming operators  $f$  arising in optimal control or (zero-sum, two player) game problems have Properties (6). Here, the operator  $f$  is a self-map of a certain space of real valued functions, equipped with the standard ordering  $\leq$  and with the sup-norm  $\|\cdot\|_\infty$ . In the simplest case, the set of states is  $\{1, \dots, n\}$ , and  $f$  is a self-map of  $\mathbb{R}^n$ . Monotone maps that are nonexpansive in the sup norm may be thought of as nonlinear generalisations of substochastic matrices. A useful subclass, which generalises stochastic matrices, consists of those maps which are monotone and commute with the addition of a constant [94] (these maps are sometimes called topical functions). Dynamic programming problems can be translated in operator terms: the dynamic programming equation for a finite horizon problem can be written as  $x(k) = f(x(k-1))$ , where  $x(k)$  is the value function in horizon  $k$  and  $x(0)$  is given; the value function  $y$  of a problem with an infinite horizon (including the case of optimal stopping) satisfies  $y = f(y)$ ; the value function  $z$  of a problem with discount factor  $0 < \alpha < 1$  satisfies  $z = f(\alpha z)$ , etc. This abstract point of view has been very fruitful, see for instance [56]. It allows one to put dynamic programming in the wider perspective of nonlinear Perron-Frobenius theory, which, after the extension of the Perron-Frobenius theorem by Krein and Rutman, studies non-linear self-maps of cones, satisfying various monotonicity, nonexpansiveness, and homogeneity conditions. Typical problems of interests are the structure of the fixed point set of  $f$ , the asymptotic behaviour of  $f^k$ , including the existence of the limit of  $f^k(x)/k$  as  $k$  tends to infinity (which yields the ergodic cost in control or games problems), the finer asymptotic behaviour of  $f^k$ , possibly up to a normalisation (which yields precise results on value iteration), etc. We shall not attempt to survey this theory here, and will only refer the reader to [149] for more background. In [111],[7], algorithms inspired from the classical policy iterations algorithm of stochastic control have been introduced for general monotone nonexpansive operators, using structural results for the fixed point set of these operators. Applications of monotone or nonexpansive maps are not limited to optimal control and game theory. In particular, we also use the same class of maps as models of discrete event dynamics systems,



see §3.5 below, and we shall see in §4.4 that related classes of maps are useful in the static analysis of computer programs.

### 3.4. Processus de Bellman/Bellman processes

Un autre point de vue sur la commande optimale est la théorie des *processus de Bellman* [160], [97], [96], [66],[1], qui fournit un analogue max-plus de la théorie des probabilités. Cette théorie a été développée à partir de la notion de *mesure idempotente* introduite par Maslov [143]. Elle établit une correspondance entre probabilités et optimisation, dans laquelle les variables aléatoires deviennent des variables de coût (qui permettent de paramétrer les problèmes d'optimisation), la notion d'espérance conditionnelle est remplacée par celle de coût conditionnel (pris sur un ensemble de solutions faisables), la propriété de Markov correspond au principe de la programmation dynamique de Bellman, et la convergence faible à une convergence de type épigraphe. Les théorèmes limites pour les processus de Bellman (loi des grands nombres, théorème de la limite centrale, lois stables) fournissent des résultats asymptotiques en commande optimale. Ces résultats généraux permettent en particulier de comprendre qualitativement les difficultés d'approximation des solutions d'équations d'Hamilton-Jacobi retrouvés en particulier dans le travail de thèse d'Asma Lakhoua [131], [63].

#### *English version*

Another point of view on optimal control is the theory of *Bellman processes* [160], [97], [96], [66], [1] which provides a max-plus analogue of probability theory, relying on the theory of *idempotent measures* due to Maslov [143]. This establishes a correspondence between probability and optimisation, in which random variables become cost variables (which allow to parametrise optimisation problems), the notion of conditional expectation is replaced by a notion of conditional cost (taken over a subset of feasible solutions), the Markov property corresponds to the Bellman's dynamic programming principle, and weak convergence corresponds to an epigraph-type convergence. Limit theorems for Bellman processes (law of large numbers, central limit theorems, stable laws) yield asymptotic results in optimal control. Such general results help in particular to understand qualitatively the difficulty of approximation of Hamilton-Jacobi equations found again in particular in the PhD thesis work of Asma Lakhoua [131], [63].

### 3.5. Systèmes à événements discrets/Discrete event systems

Des systèmes dynamiques max-plus linéaires, de type (2), interviennent aussi, avec une interprétation toute différente, dans la modélisation des systèmes à événements discrets. Dans ce contexte, on associe à chaque tâche répétitive,  $i$ , une fonction *compteur*,  $v_i : \mathbb{R} \rightarrow \mathbb{N}$ , telle que  $v_i(t)$  compte le nombre cumulé d'occurrences de la tâche  $i$  jusqu'à l'instant  $t$ . Par exemple, dans un système de production,  $v_i(t)$  compte le nombre de pièces d'un certain type produites jusqu'à l'instant  $t$ . Dans le cas le plus simple, qui dans le langage des réseaux de Petri, correspond à la sous-classe très étudiée des graphes d'événements temporisés [85], on obtient des équations min-plus linéaires analogues à (2). Cette observation, ou plutôt, l'observation duale faisant intervenir des fonctions dateurs, a été le point de départ [89] de l'approche max-plus des systèmes à événements discrets [6], qui fournit un analogue max-plus de la théorie des systèmes linéaires classiques, incluant les notions de représentation d'état, de stabilité, de séries de transfert, etc. En particulier, les valeurs propres fournissent des mesures de performance telles que le taux de production. Des généralisations non-linéaires, telles que les systèmes dynamiques min-max [150], [124], ont aussi été étudiées. Les systèmes dynamiques max-plus linéaires aléatoires sont particulièrement utiles dans la modélisation des réseaux [71]. Les modèles d'automates à multiplicités max-plus [109], incluant certains versions temporisées des modèles de traces ou de tas de pièces [113], permettent de représenter des phénomènes de concurrence ou de partage de ressources. Les automates à multiplicités max-plus ont été très étudiés par ailleurs en informatique théorique [168], [125], [137], [169], [130], [151]. Ils fournissent des modèles particulièrement adaptés à l'analyse de problèmes d'ordonnancement [136].

#### *English version*

Dynamical systems of type (2) also arise, with a different interpretation, in the modelling of discrete event systems. In this context, one associates to every repetitive task,  $i$ , a counter function,  $v_i : \mathbb{R} \rightarrow \mathbb{N}$ , such that  $v_i(t)$  gives the total number of occurrences of task  $i$  up to time  $t$ . For instance, in a manufacturing system,  $v_i(t)$  will count the number of parts of a given type produced up to time  $t$ . In the simplest case, which, in the vocabulary of Petri nets, corresponds to the much studied subclass of timed event graphs [85], we get min-plus linear equations similar to (2). This observation, or rather, the dual observation concerning dater functions, was the starting point [89] of the max-plus approach of discrete event systems [6], which provides some analogue of the classical linear control theory, including notions of state space representations, stability, transfer series, etc. In particular, eigenvalues yield performance measures like the throughput. Nonlinear generalisations, like min-max dynamical systems [150], [124], have been particularly studied. Random max-plus linear dynamical systems are particularly useful in the modelling of networks [71]. Max-plus automata models [109], which include some timed version of trace or heaps of pieces models [113], allow to represent phenomena of concurrency or resource sharing. Note that max-plus automata have been much studied in theoretical computer science [168], [125], [137], [169], [130], [151]. Such automata models are particularly adapted to the analysis of scheduling problems [136].

### 3.6. Algèbre linéaire max-plus/Basic max-plus algebra

Une bonne partie des résultats de l'algèbre max-plus concerne l'étude des systèmes d'équations linéaires. On peut distinguer trois familles d'équations, qui sont traitées par des techniques différentes : 1) Nous avons déjà évoqué dans les sections 3.2 et 3.3 le problème spectral max-plus  $Ax = \lambda x$  et ses généralisations. Celui-ci apparaît en contrôle optimal déterministe et dans l'analyse des systèmes à événements discrets. 2) Le problème  $Ax = b$  intervient en commande juste-à-temps (dans ce contexte, le vecteur  $x$  représente les dates de démarrage des tâches initiales,  $b$  représente certaines dates limites, et on se contente souvent de l'inégalité  $Ax \leq b$ ). Le problème  $Ax = b$  est intimement lié au problème d'affectation optimale, et plus généralement au problème de transport optimal. Il se traite via la théorie des correspondances de Galois abstraites, ou théorie de la résiduation [102], [78], [173], [178],[6]. Les versions dimension infinie du problème  $Ax = b$  sont reliées aux questions d'analyse convexe abstraite [170], [163], [61] et de dualité non convexe. 3) Le problème linéaire général  $Ax = Bx$  conduit à des développements combinatoires intéressants (polyèdres max-plus, déterminants max-plus, symétrisation [123], [152],[6]). Le sujet fait l'objet d'un intérêt récemment renouvelé [98].

#### *English version*

An important class of results in max-plus algebra concerns the study of max-plus linear equations. One can distinguish three families of equations, which are handled using different techniques: 1) We already mentioned in Sections 3.2 and 3.3 the max-plus spectral problem  $Ax = \lambda x$  and its generalisations, which appears in deterministic optimal control and in performance analysis of discrete event systems. 2) The  $Ax = b$  problem arises naturally in just in time problems (in this context, the vector  $x$  represents the starting times of initial tasks,  $b$  represents some deadlines, and one is often content with the inequality  $Ax \leq b$ ). The  $Ax = b$  problem is intimately related with optimal assignment, and more generally, with optimal transportation problems. Its theory relies on abstract Galois correspondences, or residuation theory [102], [78], [173], [178],[6]. Infinite dimensional versions of the  $Ax = b$  problem are related to questions of abstract convex analysis [170], [163], [61] and nonconvex duality. 3) The general linear system  $Ax = Bx$  leads to interesting combinatorial developments (max-plus polyedra, determinants, symmetrisation [123], [152],[6]). The subject has attracted recently a new attention [98].

### 3.7. Algèbre max-plus et asymptotiques/Using max-plus algebra in asymptotic analysis

Le rôle de l'algèbre min-plus ou max-plus dans les problèmes asymptotiques est évident si l'on écrit

$$e^{-a/\epsilon} + e^{-b/\epsilon} \asymp e^{-\min(a,b)/\epsilon}, \quad e^{-a/\epsilon} \times e^{-b/\epsilon} = e^{-(a+b)/\epsilon}, \quad (26)$$

lorsque  $\epsilon \rightarrow 0^+$ . Formellement, l'algèbre min-plus peut être vue comme la limite d'une déformation de l'algèbre classique, en introduisant le semi-anneau  $\mathbb{R}_\epsilon$ , qui est l'ensemble  $\mathbb{R} \cup \{+\infty\}$ , muni de l'addition  $(a, b) \mapsto -\epsilon \log(e^{-a/\epsilon} + e^{-b/\epsilon})$  et de la multiplication  $(a, b) \mapsto a + b$ . Pour tout  $\epsilon > 0$ ,  $\mathbb{R}_\epsilon$  est isomorphe au semi-corps usuel des réels positifs,  $(\mathbb{R}_+, +, \times)$ , mais pour  $\epsilon = 0^+$ ,  $\mathbb{R}_\epsilon$  n'est autre que le semi-anneau min-plus. Cette idée a été introduite par Maslov [143], motivé par l'étude des asymptotiques de type WKB d'équations de Schrödinger. Ce point de vue permet d'utiliser des résultats algébriques pour résoudre des problèmes d'asymptotiques, puisque les équations limites ont souvent un caractère min-plus linéaire.

Cette déformation apparaît classiquement en théorie des grandes déviations à la loi des grands nombres : dans ce contexte, les objets limites sont des mesures idempotentes au sens de Maslov. Voir [1], [159], [62], pour les relations entre l'algèbre max-plus et les grandes déviations, voir aussi [60], [59], [58] pour des applications de ces idées aux perturbations singulières de valeurs propres. La même déformation est à l'origine de nombreux travaux actuels en géométrie tropicale, à la suite de Viro [172].

#### *English version*

The role of min-plus algebra in asymptotic problems becomes obvious when writing Equations (7) when  $\epsilon \rightarrow 0^+$ . Formally, min-plus algebra may be thought of as the limit of a deformation of classical algebra, by introducing the semi-field  $\mathbb{R}_\epsilon$ , which is the set  $\mathbb{R} \cup \{+\infty\}$ , equipped with the addition  $(a, b) \mapsto -\epsilon \log(e^{-a/\epsilon} + e^{-b/\epsilon})$  and the multiplication  $(a, b) \mapsto a + b$ . For all  $\epsilon > 0$ ,  $\mathbb{R}_\epsilon$  is isomorphic to the semi-field of usual real positive numbers,  $(\mathbb{R}_+, +, \times)$ , but for  $\epsilon = 0^+$ ,  $\mathbb{R}_\epsilon$  coincides with the min-plus semiring. This idea was introduced by Maslov [143], motivated by the study of WKB-type asymptotics of Schrödinger equations. This point of view allows one to use algebraic results in asymptotics problems, since the limit equations have often some kind of min-plus linear structure.

This deformation appears classically in large deviation theory: in this context, the limiting objects are idempotent measures, in the sense of Maslov. See [1], [159], [62] for the relation between max-plus algebra and large deviations. See also [60], [59], [58] for the application of such ideas to singular perturbation problems for matrix eigenvalues. The same deformation is at the origin of many current works in tropical geometry, in the line initiated by Viro [172].

## MCTAO Team

### 3. Scientific Foundations

#### 3.1. Control Systems

Our effort is directed toward efficient methods for the *control* of real (physical) systems, based on a *model* of the system to be controlled. *System* refers to the physical plant or device, whereas *model* refers to a mathematical representation of it.

We mostly investigate nonlinear systems whose nonlinearities admit a strong structure derived from the physics; the equations governing their behavior is then rather well known, and the modeling part consists in choosing what phenomena are to be retained in the model used for control design; the other phenomena being treated as perturbations; a more complete model may be used for simulations, for instance. We focus on systems that admit a reliable finite-dimensional model, in continuous time; this means that models are ordinary differential equations, often nonlinear.

Choosing accurate models yet simple enough to allow control design is in itself a key issue; however, modeling or identification as a theory is not per se in the scope of our project.

The extreme generality and versatility of linear control do not contradict the often heard sentence “most real life systems are nonlinear”. Indeed, for many control problems, a linear model is sufficient to capture the important features for control. The reason is that most control objectives are local, first order variations around an operating point or a trajectory are governed by a linear control model, and except in degenerate situations (non-controllability of this linear model), the local behavior of a nonlinear dynamic phenomenon is dictated by the behavior of first order variations. Linear control is the hard core of control theory and practice; it has been pushed to a high degree of achievement –see for instance some classics: [56], [42]– that leads to big successes in industrial applications (PID, Kalman filtering, frequency domain design,  $H^\infty$  robust control, etc...). It must be taught to future engineers, and it is still a topic of ongoing research.

Linear control by itself however reaches its limits in some important situations:

1. **Non local control objectives.** For instance, steering the system from a region to a reasonably remote other one (path planning and optimal control); in this case, local linear approximation cannot be sufficient.  
It is also the case when some domain of validity (e.g. stability) is prescribed and larger than the region where the linear approximation is dominant.
2. **Local control at degenerate equilibria.** Linear control yields local stabilization of an equilibrium point based on the tangent linear approximation if the latter is controllable. When it is *not*, and this occurs in some physical systems at interesting operating points, linear control is irrelevant and specific nonlinear techniques have to be designed.  
This is in a sense an extreme case of the second paragraph in point 1 : the region where the linear approximation is dominant vanishes.
3. **Small controls.** In some situations, actuators only allow a very small magnitude of the effect of control compared to the effect of other phenomena. Then the behavior of the system without control plays a major role and we are again outside the scope of linear control methods.
4. **Local control around a trajectory.** Sometimes a trajectory has been selected (this appeals to point 1), and local regulation around this reference is to be performed. Linearization in general yields, when the trajectory is not a single equilibrium point, a *time-varying* linear system. Even if it is controllable, time-varying linear systems are not in the scope of most classical linear control methods, and it is better to incorporate this local regulation in the nonlinear design, all the more so as the linear approximation along optimal trajectories is, by nature, often non controllable.

Let us discuss in more details some specific problems that we are studying or plan to study: classification and structure of control systems in section 3.2 , optimal control, and its links with feedback, in section 3.3 , the problem of optimal transport in section 3.4 , and finally problems relevant to a specific class of systems where the control is “small” in section 3.5 .

## 3.2. Structure of nonlinear control systems

In most problems, choosing the adapted coordinates, or the right quantities that describe a phenomenon, sheds light on a path to the solution. In control systems, it is often crucial to analyze the structure of the model, deduced from physical principles, of the plant to be controlled; this may lead to putting it via some transformations in a simpler form, or a form that is most suitable for control design. For instance, equivalence to a linear system may allow to use linear control; also, the so-called “flatness” property drastically simplifies path planning [48], [63].

A better understanding of the “set of nonlinear models”, partly classifying them, has another motivation than facilitating control design for a given system and its model: it may also be a necessary step towards a theory of “nonlinear identification” and modeling. Linear identification is a mature area of control science; its success is mostly due to a very fine knowledge of the structure of the class of linear models: similarly, any progress in the understanding of the structure of the class of nonlinear models would be a contribution to a possible theory of nonlinear identification.

These topics are central in control theory, but raise very difficult mathematical questions: static feedback classification is a geometric problem feasible in principle, although describing invariants explicitly is technically very difficult; and conditions for dynamic feedback equivalence and linearization raise unsolved mathematical problems, that make one wonder about decidability <sup>1</sup>.

## 3.3. Optimal control and feedback control, stabilization

### 3.3.1. Optimal control.

Mathematically speaking, optimal control is the modern branch of the calculus of variations, rather well established and mature [24], [60], [34], [72]. Relying on Hamiltonian dynamics is now prevalent, instead of the standard Lagrangian formalism of the calculus of variations. Also, coming from control engineering, constraints on the control (for instance the control is a force or a torque, naturally bounded) or the state (for example in the shuttle atmospheric re-entry problem there is a constraint on the thermal flux) are imposed; the ones on the state are usual but these on the state yield more complicated necessary optimality conditions and an increased intrinsic complexity of the optimal solutions. Also, in the modern treatment, adapted numerical schemes have to be derived for effective computations of the optimal solutions.

What makes optimal control an applied field is the necessity of computing these optimal trajectories, or rather the controls that produce these trajectories (or, of course, close-by trajectories). Computing a given optimal trajectory and its control as a function of time is a demanding task, with non trivial numerical difficulties: roughly speaking, the Pontryagin Maximum Principle gives candidate optimal trajectories as solutions of a two point boundary value problem (for an ODE) which can be analyzed using mathematical tools from geometric control theory or solved numerically using shooting methods. Obtaining the *optimal synthesis* –the optimal control as a function of the state– is of course a more intricate problem [34], [37]. On the other hand the *value function* –the minimum of the criteria at each point– would give the solution to both questions if it were differentiable, but it is usually not, and is only a viscosity solution of the Hamilton-Jacobi-Bellman (**HJB**) equation (a PDE), whose study requires sophisticated non-smooth and singularity analysis.

<sup>1</sup>Consider the simple system with state  $(x, y, z) \in \mathbb{R}^3$  and two controls that reads  $\dot{z} = (\dot{y} - z\dot{x})^2 \dot{x}$  after elimination of the controls; it is not known whether it is equivalent to a linear system, or flat; this is because the property amounts to existence of a formula giving the general solution as a function of two arbitrary functions of time and their derivatives up to a certain order, but no bound on this order is known a priori, even for this very particular example.

These questions are not only academic for minimizing a cost is *very* relevant in many control engineering problems. However, modern engineering textbooks in nonlinear control systems like the “best-seller” [51] hardly mention optimal control, and rather put the emphasis on designing a feedback control, as regular and explicit as possible, satisfying some qualitative (and extremely important!) objectives: disturbance attenuation, decoupling, output regulation or stabilization. Optimal control is sometimes viewed as disconnected from automatic control... we shall come back to this unfortunate point.

### 3.3.2. Feedback, control Lyapunov functions, stabilization.

A control Lyapunov function (CLF) is a function that can be made a Lyapunov function (roughly speaking, a function that decreases along all trajectories, some call this an “artificial potential”) for the closed-loop system corresponding to *some* feedback law. This can be translated into a partial differential relation sometimes called “Artstein’s (in)equation” [27]. There is a definite parallel between a CLF for stabilization, solution of this differential inequation on the one hand, and the value function of an optimal control problem for the system, solution of a HJB equation on the other hand. Now, optimal control is a quantitative objective while stabilization is a qualitative objective; it is not surprising that Artstein (in)equation is very under-determined and has many more solutions than HJB equation, and that it may (although not always) even have smooth ones.

We have, in the team, a longstanding research record on the topic, construction of CLFs and stabilizing feedback controls. This is all the more interesting as our line of research has been pointing in almost opposite directions. [43], [52], [53], [67], [68], [70], [71], [44] insist on the construction of continuous feedback, hence smooth CLFs whereas, on the contrary, [41], [74], [75], [76], [77] proceed with a very fine study of non-smooth CLFs, yet good enough (semi-concave) that they can produce a reasonable discontinuous feedback with reasonable properties.

## 3.4. Optimal Transport

The study of optimal mass transport problems in the Euclidean or Riemannian setting has a long history which goes from the pioneer works of Monge [65] and Kantorovitch [57] to the recent revival initiated by fundamental contributions due to Brenier [38], [39] and McCann [64]. However, the study of the same transportation problems in the presence of non-holonomic constraints –(like being an admissible trajectory for a control system– is quite new. The first contributors were Ambrosio and Rigot [25] who proved the existence and uniqueness of an optimal transport map for the Monge problem associated with the squared canonical sub-Riemannian distance on the Heisenberg groups. This result was extended later by Agrachev and Lee [22], then by Figalli and Rifford [46] who showed that the Ambrosio-Rigot theorem holds indeed true on many sub-Riemannian manifolds satisfying reasonable assumptions. The problem of existence and uniqueness of an optimal transport map for the squared sub-Riemannian distance on a general complete sub-Riemannian manifold remains open; it is strictly related to the regularity of the sub-Riemannian distance in the product space, and remains a formidable challenge. Generalized notions of Ricci curvatures (bounded from below) in metric spaces have been developed recently by Lott and Villani [61] and Sturm [79], [80]. A pioneer work has been work in the Heisenberg group by Juillet [54] who captured the right notion of curvature in this setting. Agrachev and Lee [23] have elaborated on this work to define new notions of curvatures in three dimensional sub-Riemannian structures. The optimal transport approach happened to be very fruitful in this context. Many things remain to do in a more general context.

One of the results of A. Hindawi’s PhD under the supervision of L. Rifford and J.-B. Pomet was to extend regularity theory established in the Euclidean case to the more general quadratic costs associated with linear optimal control problems (LQR), see [50]. This successful result opens a new range of optimal transport problems associated with cost coming from optimal control problems. We can nowadays expect regularity properties for optimal transport maps associated with reasonable optimal control problems with constraints on the state or on the velocities.

We believe that matching optimal transport with geometric control theory is one originality of our team. We expect interactions in both ways.

### 3.5. Small controls and conservative systems, averaging

Using averaging techniques to study small perturbations of integrable Hamiltonian systems dates back to H. Poincaré or earlier; it gives an approximation of the (slow) evolution of quantities that are preserved in the non-perturbed system. It is very subtle in the case of multiple periods but more elementary in the single period case, here it boils down to taking the average of the perturbation along each periodic orbit; see for instance [26], [78].

When the “perturbation” is a control, these techniques may be used after deciding how the control will depend on time and state and other quantities, for instance it may be used after applying the Pontryagin Maximum Principle as in [30], [31], [40], [49]. Without deciding the control a priori, an “average control system” may be defined as in [2].

The focus is then on studying into details this simpler “averaged” problem, that can often be described by a Riemannian metric for quadratic costs or by a Finsler metric for costs lime minimum time.

This line of research stemmed out of applications to space engineering, see section 4.1 . For orbit transfer in the two-body problem, an important contribution was made by B. Bonnard, J.-B. Caillaud and J. Gergaud [31] in explicitly computing the solutions of the average system obtained after applying Pontryagin Maximum Principle to minimizing a quadratic integral cost; this yields an explicit calculation of the optimal control law itself. Studying the Finsler metric issued from the time-minimal case is in progress.

## NECS Project-Team

### 3. Scientific Foundations

#### 3.1. Multi-disciplinary nature of the project

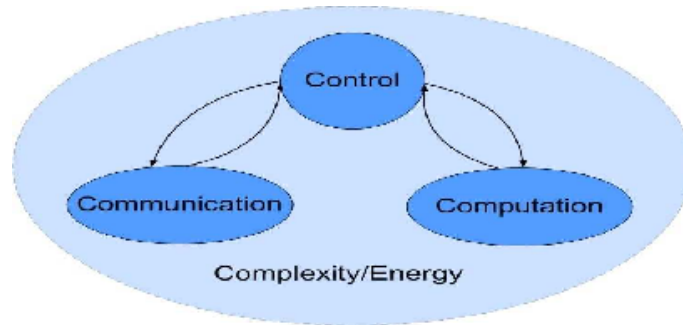


Figure 2. Relation of the NCS area with the fields of: Control, Communication, Computation.

The team's project is to investigate problems in the area of NCS with the originality of integrated aspects on computation, communication and control. The combination of these three disciplines requires the interplay of the multi-disciplinary fields of: communication, real-time computation, and systems theory (control). Figure 2, shows the natural interaction between disciplines that concern the NeCS project. The arrows describe the direction in which these areas interact, i.e.

- (a) *Control in Communication*
- (b) *Communication in Control*
- (c) *Computation in Control*
- (d) *Control in Computation*

Complexity and energy-management are additional features to be considered as well. Complexity here refers to the problems coming from: wireless networks with varying interconnection topologies, multi-agent systems coordination, scaling with respect to a growing number of sensors. Energy management concerns in particular the efficient handling of energy in wireless sensors, and means an efficient way to handle both information transmission and computation.

##### 3.1.1. (a) *Control in Communication*

This topic is the study of how control-theoretic methods can be applied in order to solve some problems found in the communication field. Examples are: the Power control in cell telephones, and the optimal routing of messages in communication networks (Internet, sensor networks).

##### 3.1.2. (b) *Communication in Control*

This area concerns problems where communication and information theory interact with systems theory (control). As an example of a classical paradigm we can mention the stabilization problem under channel (communication) constraints. A key result here [75] was to show that it was generically impossible to stabilize a linear system in any reasonable sense, if the feedback channel's Shannon classical capacity  $C$  was smaller than the sum of the logarithms, base 2, of the unstable eigenvalues. In other words, in order to be able to cope with the stabilization problem under communication constraints, we need that



$$C > \sum_i \log_2 \lambda_i$$

where the  $\lambda_i$ 's are unstable eigenvalues of the open loop system. Intuitively, this means that the rate of information production (for discrete-time linear systems, the intrinsic rate bits/time equals  $\sum_i \log_2 \lambda_i$ ) should be smaller than the rate of information that can be transmitted throughout the channel. In that way, a potentially growing signal can be cached out, if the information of the signal is sent via a channel with fast enough transmission rate. In relation to this, a problem of interest is the coding and control co-design. This issue is motivated by applications calling for data-compression algorithms aiming at reducing the amount of information that may be transmitted throughout the communication channel, and therefore allowing for a better resource allocation and/or for an improvement of the permissible closed loop system bandwidth (data-rate). Networked controlled systems also constitute a new class of control systems including specific problems concerned by delays. In NCS, the communication between two agents leads unavoidably to transmission delays. Also, transmission usually happens in discrete time, whereas most controlled processes evolve in continuous time. Moreover, communication can induce loss of information. Our objectives concern the stabilization of systems where the sensor, actuator and system are assumed to be remotely commissioned by a controller that interchanges measurements and control signals through a communication network. Additional dynamics are introduced due to time-varying communication delays, asynchronous samplings, packets losses or lack of synchronization. All those phenomena can be modeled as the introduction of time-delays in the closed loop system. Even if these time-delay approaches can be easily proposed, they require careful attention and more complex analysis. In general, the introduction of delays in a controlled loop leads to a reduction of the performance with respect to the delay-free situation and could even make the systems unstable. Our objective is to provide specific modeling of these phenomena and to develop dedicated tools and methodologies to cope with stability and stabilization of such systems.

### 3.1.3. (c) Computation in Control

This area concerns the problem of redesigning the control law such as to account for variations due to the resource allocation constraints. Computation tasks having different levels of priority may be handled by asynchronous time executions. Hence controllers need to be re-designed as to account for non-uniform sampling times resulting in this framework. Questions on how to redesign the control laws while preserving its stability properties are in order. This category of problems can arise in embedded systems with low computation capacity or low level resolution.

### 3.1.4. (d) Control in Computation

The use of control methods to solve or to optimize the use of computational resources is the key problem in this area. This problem is also known as a scheduling control. The resource allocations are decided by the controller that tries to regulate the total computation load to a prefixed value. Here, the system to be regulated is the process that generates and uses the resources, and not any physical system. Hence, internal states are computational tasks, the control signal is the resource allocation, and the output is the period allowed to each task.

### 3.1.5. (c + d) Integrated control/scheduling co-design

Control and Computation co-design describes the possibility to study the interaction or coupling between the flows (c) and (d). It is possible, as shown in Fig. 3, to re-frame both problems as a single one, or to interpret such an interconnection as the cascade connection between a computational system, and a physical system. In our framework the feedback scheduling is designed w.r.t. a QoC (Quality of Control) measure. The QoC criterion captures the control performance requirements, and the problem can be stated as QoC optimization under constraint of available computing resources. However, preliminary studies suggest that a direct synthesis of the scheduling regulator as an optimal control problem leads, when it is tractable, to a solution too costly to be implemented in real-time applications [64]. Practical solutions will be found in the currently available control theory and tools or in enhancements and adaptation of current control theory. We propose in Fig. 3

a hierarchical control structure: besides the usual process control loops we add an outer control loop whose goal is to manage the execution of the real-time application through the control of the scheduling parameters of the inner loops. Together with the outer loop (working on a periodic sampled time scale) we also need a scheduling manager working on a discrete events time scale to process exception handling and admission control. The task periods directly affect the computing load, they have been chosen as actuators. They can be implemented through software variable clocks. As timing uncertainties cannot be avoided and are difficult to model or measure, we currently design robust control algorithms using the  $H_\infty$  control theory, which have been successfully simulated and experimentally validated [74]. This methodology is supported by the software ORCCAD (see Section 5.1) where a run-time library for multi-rate multitasking has been developed and integrated. It will be further improved using a QoS-based management of the timing constraints to fully benefit from the intrinsic robustness of closed loop controllers w.r.t. timing uncertainties.

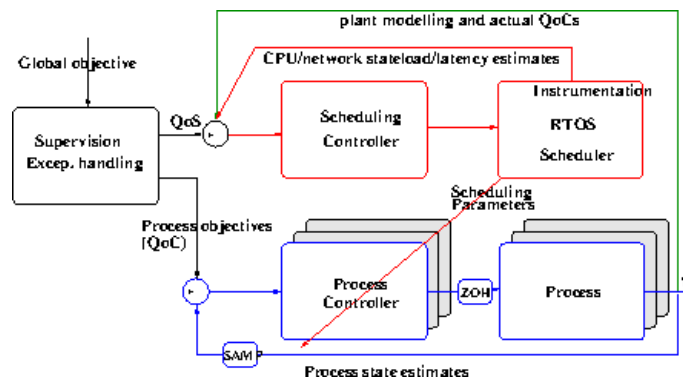


Figure 3. Hierarchical control structure.

### 3.2. Main Research Directions

The main objective of the project is to develop a unified control, communication, computing co-design methodology explicitly accounting for all the components involved in the system controlled over a network. This includes quantifier properties, scheduling parameters, encoder/decoder, alphabet length, bandwidth of the transmission media (wire or wireless), delays, resource allocation, jitter, etc. These components, including the control laws, should be designed so as to optimize performance/stability trade-offs resulting from the ceiling of the computing resources, the channel capacity limitations and the quality of the send/received information protocols. More informations about the main research directions of the team can be found in [1], [3],[2], [4] [5], [6], [7], [8], [9] and [10].

In short, the project is centered along the following 3 main axes:

1. **Control under Communications Constraints.** One well established topic along this axis concerns the coding and control co-design. That is, the design of new code alphabets simultaneously than the design of the control law. Or equivalently, the ability of designing codes containing information pertained to the system model and the control law. The objective being the improvements of the overall closed loop performances. Besides this matter, additional improvements pertain to the field of the information theory are also in order.
2. **Control under Computational resources constraints.** The main objective here is the design of control loops by explicitly accounting for the network and/or the computing resources. Dynamic allocation of such resources depends on the desired controlled systems specifications. Keys aspects to be considered are: the design of controllers with variable sampling time, the robustness with

respect to time uncertainties such as the input/output latencies, the global control of resources and its impact over the performance and the robustness of the system to be controlled. We aim to provide an integrated control and scheduling co-design approach [1].

3. **Controlling Complexity.** Design and control of partially cooperative networked (possible also multi-agent) systems subject to communication and computational constraints. Here, a large number of entities (agents), having each its own goal share limited common resources. In this context, if there is no minimum coordination, dramatic consequences may follow, on the other hand, total coordination would be impossible because of the lack of exhaustive, reliable and synchronous information. Finally, a local network of strategies that are based on worst-case assumptions is clearly far from being realistic for a well designed system. The aim of this topic is to properly define key concepts and the relevant variables associated to the above problem (sub-system, partial objective, constraints on the exchanged data and computational resources, level of locally shared knowledge, key parameters for the central level, etc).

## NON-A Project-Team

### 3. Scientific Foundations

#### 3.1. Fast parametric estimation and its applications

Parametric estimation may often be formalized as follows:

$$y = F(x, \Theta) + n, \quad (27)$$

where:

- the measured signal  $y$  is a functional  $F$  of the "true" signal  $x$ , which depends on a set  $\Theta$  of parameters,
- $n$  is a noise corrupting the observation.

Finding a "good" approximation of the components of  $\Theta$  has been the subject of a huge literature in various fields of applied mathematics. Most of those researches have been done in a probabilistic setting, which necessitates a good knowledge of the statistical properties of  $n$ . Our project is devoted to a new standpoint, which does not require this knowledge and which is based on the following tools, which are of algebraic flavor:

- differential algebra <sup>2</sup>, which plays with respect to differential equations a similar role that the commutative algebra plays with respect to algebraic equations;
- module theory, i.e. linear algebra over rings, which are not necessarily commutative;
- operational calculus, which is the most classical tool among control and mechanical engineers <sup>3</sup>.

##### 3.1.1. Linear identifiability

In the most problems, which appear in linear control as well as in signal processing, the unknown parameters are *linearly identifiable*: standard elimination procedures are yielding the following matrix equation

$$P \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_r \end{pmatrix} = Q, \quad (28)$$

where:

- $\theta_i, 1 \leq i \leq r$  represents unknown parameter,
- $P$  is a  $r \times r$  square matrix and  $Q$  is a  $r \times 1$  column matrix,
- the entries of  $P$  and  $Q$  are finite linear combinations of terms of the form  $t^\nu \frac{d^\mu \xi}{dt^\mu}$ ,  $\mu, \nu \geq 0$ , where  $\xi$  is an input or output signal,
- the matrix  $P$  is *generically* invertible, i.e.,  $\det(P) \neq 0$ .

##### 3.1.2. How to deal with perturbations and noises?

With noisy measurements equation (2) becomes:

<sup>2</sup>Differential algebra was introduced in nonlinear control theory by one of us almost twenty years ago for understanding some specific questions like input-output inversion. It allowed us to recast the whole of nonlinear control into a more realistic light. The best example is of course the discovery of *flat* systems, which are now quite popular in industry.

<sup>3</sup>Operational calculus is often formalized *via* the Laplace transform whereas the Fourier transform is today the cornerstone in estimation. Note that the one-sided Laplace transform is causal, but the Fourier transform over  $\mathbb{R}$  is not.

$$P \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_r \end{pmatrix} = Q + R, \quad (29)$$

where  $R$  is a  $r \times 1$  column matrix, whose entries are finite linear combinations of terms of the form  $t^\nu \frac{d^\mu \eta}{dt^\mu}$ ,  $\mu, \nu \geq 0$ , where  $\eta$  is a perturbation or a noise.

### 3.1.2.1. Structured perturbations

A perturbation  $\pi$  is said to be *structured* if, and only if, it can be annihilated by a linear differential operator of the form  $\sum_{\text{finite}} a_k(t) \frac{d^k}{dt^k}$ , where  $a_k(t)$  is a rational function of  $t$ , i.e.  $\left( \sum_{\text{finite}} a_k(t) \frac{d^k}{dt^k} \right) \pi = 0$ . Note that many classical perturbations, like a constant bias, are annihilated by such an operator. An *unstructured* noise cannot be annihilated by a non-zero differential operator.

By well-known properties of the non-commutative ring of differential operators, we can multiply both sides of equation (3) by a suitable differential operator  $\Delta$  such that equation (3) becomes:

$$\Delta P \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_r \end{pmatrix} = \Delta Q + R', \quad (30)$$

where the entries of the  $r \times 1$  column matrix  $R'$  are unstructured noises.

### 3.1.2.2. Attenuating unstructured noises

Unstructured noises are usually dealt with stochastic processes like white Gaussian noises. They are considered here as highly fluctuating phenomena, which may therefore be attenuated *via* low pass filters. Note that no precise knowledge of the statistical properties of the noises is required.

### 3.1.2.3. Comments

Although the previous noise attenuation procedure<sup>4</sup> may be fully explained *via* formula (4), its theoretical comparison<sup>5</sup> with today's literature<sup>6</sup> has yet to be done. It will require a complete resetting of the notions of noises and perturbations. Besides some connections with physics, it might lead to quite new "epistemological" issues [101].

## 3.1.3. Some hints on the calculations

The time derivatives of the input and output signals appearing in equations (2), (3), (4) can be suppressed in the two following ways which might be combined:

- integrate both sides of the equation a sufficient number of times,
- take the convolution product of both sides by a suitable low pass filter.

The numerical values of the unknown parameters  $\Theta = (\theta_1, \dots, \theta_r)$  can be obtained by integrating both sides of the modified equation (4) during a very short time interval.

### 3.1.4. A first, very simple example

Let us illustrate on a very basic example, the grounding ideas of the algebraic approach. For this purpose consider the first order linear system:

<sup>4</sup>It is reminiscent to that the most practitioners in electronics are doing.

<sup>5</sup>Let us stress again that many computer simulations and several laboratory experiments have been already successfully achieved and can be quite favorably compared with the existing techniques.

<sup>6</sup>Especially in signal processing.

$$\dot{y}(t) = ay(t) + u(t) + \gamma_0, \quad (31)$$

where  $a$  is an unknown parameter to be identified and  $\gamma_0$  is an unknown constant perturbation. With the notations of operational calculus and  $y_0 = y(0)$ , equation (5) reads:

$$s\hat{y}(s) = a\hat{y}(s) + \hat{u}(s) + y_0 + \frac{\gamma_0}{s} \quad (32)$$

where  $\hat{y}(s)$  represents the Laplace transform of  $y(t)$ .

In order to eliminate the term  $\gamma_0$ , multiply first the both hand-sides of this equation by  $s$  and next take their derivatives with respect to  $s$ :

$$\frac{d}{ds} \left[ s \left\{ s\hat{y}(s) = a\hat{y}(s) + \hat{u}(s) + y_0 + \frac{\gamma_0}{s} \right\} \right] \quad (33)$$

$$\Rightarrow 2s\hat{y}(s) + s^2\hat{y}'(s) = a(s\hat{y}'(s) + \hat{y}(s)) + s\hat{u}'(s) + \hat{u}(s) + y_0. \quad (34)$$

Recall that  $\hat{y}'(s) \triangleq \frac{d\hat{y}(s)}{ds}$  corresponds to  $-ty(t)$ . Assume  $y_0 = 0$  for simplicity of presentation <sup>7</sup>. Then for any  $\nu > 0$ ,

$$s^{-\nu} [2s\hat{y}(s) + s^2\hat{y}'(s)] = s^{-\nu} [a(s\hat{y}'(s) + \hat{y}(s)) + s\hat{u}'(s) + \hat{u}(s)]. \quad (35)$$

For  $\nu = 3$ , we obtained the estimated value  $a$ :

$$a = \frac{2 \int_0^T d\lambda \int_0^\lambda y(t)dt - \int_0^T ty(t)dt + \int_0^T d\lambda \int_0^\lambda tu(t)dt - \int_0^T d\lambda \int_0^\lambda d\sigma \int_0^\sigma u(t)dt}{\int_0^T d\lambda \int_0^\lambda d\sigma \int_0^\sigma y(t)dt - \int_0^T d\lambda \int_0^\lambda ty(t)dt} \quad (36)$$

Since  $T > 0$  can be very small, estimation *via* (10) is very fast.

Note that equation (10) represents an on-line algorithm, which involves only two kinds of operations on  $u$  and  $y$ : (1) multiplications by  $t$ , and (2) integrations over a pre-selected time interval.

If we now consider an additional noise of zero mean in (5), say:

$$\dot{y}(t) = ay(t) + u(t) + \gamma_0 + n(t), \quad (37)$$

it can be considered as a fast fluctuating signal. The order  $\nu$  in (9) determines the order of iterations in the integrals (3 integrals in (10)). Those iterated integrals are low-pass filters which are attenuating the fluctuations.

This example, even simple, clearly demonstrates how algebraic techniques proceed:

- they are algebraic: operations on  $s$ -functions;
- they are non-asymptotic: parameter  $a$  is obtained from (10) in a finite time;
- they are deterministic: no knowledge of the statistical properties of the noise  $n$  is required.

<sup>7</sup>If  $y_0 \neq 0$  one has to take above derivatives of order 2 with respect to  $s$ , in order to eliminate the initial condition.

### 3.1.5. A second simple example, with delay

Consider the first order, linear system with constant input delay <sup>8</sup>:

$$\dot{y}(t) + ay(t) = y(0)\delta + \gamma_0 H + bu(t - \tau). \quad (38)$$

Here we use a distributional-like notation, where  $\delta$  denotes the Dirac impulse and  $H$  is its integral, i.e. the Heaviside function (unit step) <sup>9</sup>. Still for simplicity, we suppose that the parameter  $a$  is known. The parameter to be identified is now the delay  $\tau$ . As previously,  $\gamma_0$  is a constant perturbation,  $a$ ,  $b$ , and  $\tau$  are constant parameters. Consider also a step input  $u = u_0 H$ . A first order derivation yields:

$$\ddot{y} + a\dot{y} = \varphi_0 + \gamma_0 \delta + b u_0 \delta_\tau, \quad (39)$$

where  $\delta_\tau$  denotes the delayed Dirac impulse and  $\varphi_0 = (\dot{y}(0) + ay(0))\delta + y(0)\delta^{(1)}$ , of order 1 and support  $\{0\}$ , contains the contributions of the initial conditions. According to Schwartz theorem, multiplication by a function  $\alpha$  such that  $\alpha(0) = \alpha'(0) = 0$ ,  $\alpha(\tau) = 0$  yields interesting simplifications. For instance, choosing  $\alpha(t) = t^3 - \tau t^2$  leads to the following equalities (to be understood in the distributional framework):

$$\begin{aligned} t^3 [\ddot{y} + a\dot{y}] &= \tau t^2 [\ddot{y} + a\dot{y}], \\ b u_0 t^3 \delta_\tau &= b u_0 \tau t^2 \delta_\tau. \end{aligned} \quad (40)$$

The delay  $\tau$  becomes available from  $k \geq 1$  successive integrations (represented by the operator  $H$ ), as follows:

$$\tau = \frac{H^k(w_0 + a w_3)}{H^k(w_1 + a w_2)}, \quad t > \tau, \quad (41)$$

where the  $w_i$  are defined using the notation  $z_i = t^i y$  by:

$$\begin{aligned} w_0 &= t^3 y^{(2)} = -6 z_1 + 6 z_2^{(1)} - z_3^{(2)}, \\ w_1 &= t^2 y^{(2)} = -2 z_0 + 4 z_1^{(1)} - z_2^{(2)}, \\ w_2 &= t^2 y^{(1)} = 2 z_1 - z_2^{(1)}, \\ w_3 &= t^3 y^{(1)} = 3 z_2 - z_3^{(1)}. \end{aligned}$$

These coefficients show that  $k \geq 2$  integrations avoid any derivation in the delay identification.

Figure 1 gives a numerical simulation with  $k = 2$  integrations and  $a = 2, b = 1, \tau = 0.6, y(0) = 0.3, \gamma_0 = 2, u_0 = 1$ . Due to the non identifiability over  $(0, \tau)$ , the delay  $\tau$  is set to zero until the numerator or the denominator in the right hand side of (15) reaches a significant nonzero value.

Again, note the realization algorithm (15) involves two kinds of operators: (1) integrations and (2) multiplications by  $t$ . It relies on the measurement of  $y$  and on the knowledge of  $a$ . If  $a$  is also unknown, the same approach can be utilized for a simultaneous identification of  $a$  and  $\tau$ . The following relation is derived from (14):

<sup>8</sup>This example is taken from [93]. For further details, we suggest the reader to refer to it.

<sup>9</sup>In this document, for the sake of simplicity, we make an abuse of the language since we merge in a single notation the Heaviside function  $H$  and the integration operator. To be rigorous, the iterated integration ( $k$  times) corresponds, in the operational domain, to a division by  $s^k$ , whereas the convolution with  $H$  ( $k$  times) corresponds to a division by  $s^k / (k - 1)!$ . For  $k = 0$ , there is no difference and  $H * y$  realizes the integration of  $y$ . More generally, since we will always apply these operations to complete equations (left- and right-hand sides), the factor  $(k - 1)!$  makes no difference.

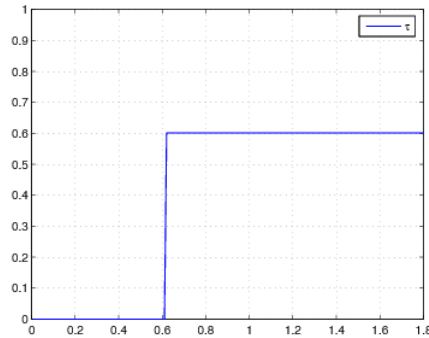


Figure 1. Delay  $\tau$  identification from algorithm (15)

$$\tau(H^k w_1) + a \tau(H^k w_2) - a(H^k w_3) = H^k w_0, \tag{42}$$

and a linear system with unknown parameters  $(\tau, a, \tau, a)$  is obtained by using different integration orders:

$$\begin{pmatrix} H^2 w_1 & H^2 w_2 & H^2 w_3 \\ H^3 w_1 & H^3 w_2 & H^3 w_3 \\ H^4 w_1 & H^4 w_2 & H^4 w_3 \end{pmatrix} \begin{pmatrix} \hat{\tau} \\ \hat{a}\tau \\ -\hat{a} \end{pmatrix} = \begin{pmatrix} H^2 w_0 \\ H^3 w_0 \\ H^4 w_0 \end{pmatrix}.$$

The resulting numerical simulations are shown in Figure 2. For identifiability reasons, the obtained linear system may be not consistent for  $t < \tau$ .

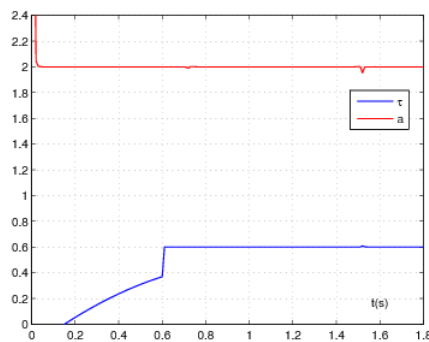


Figure 2. Simultaneous identification of  $a$  and  $\tau$  from algorithm (16)



## 3.2. Finite time estimation of derivatives

Numerical differentiation, i.e. determining the time derivatives of various orders of a noisy time signal, is an important but difficult ill-posed theoretical problem. This fundamental issue has attracted a lot of attention in many fields of engineering and applied mathematics (see, e.g. in the recent control literature [94], [95], [109], [108], [115], [116], and the references therein).

### 3.2.1. Model-free techniques for numerical differentiation

A common way of estimating the derivatives of a signal is to resort to a least squares fitting and then take the derivatives of the resulting function. In [119], [117] this problem was revised through our algebraic approach. The approach can be briefly explained as follows:

- The coefficients of a polynomial time function are linearly identifiable. Their estimation can therefore be achieved as above. Indeed, consider a real-valued polynomial function  $x_N(t) = \sum_{\nu=0}^N x^{(\nu)}(0) \frac{t^\nu}{\nu!} \in \mathbb{R}[t]$ ,  $t \geq 0$ , of degree  $N$ . Rewrite it in the well-known notations of operational calculus:

$$X_N(s) = \sum_{\nu=0}^N \frac{x^{(\nu)}(0)}{s^{\nu+1}}.$$

Here we use  $\frac{d}{ds}$ , which corresponds in the time domain to the multiplication by  $-t$ . Multiply both sides by  $\frac{d^\alpha}{ds^\alpha} s^{N+1}$ ,  $\alpha = 0, 1, \dots, N$ . The quantities  $x^{(\nu)}(0)$ ,  $\nu = 0, 1, \dots, N$  are given by the triangular system of linear equations:

$$\frac{d^\alpha s^{N+1} X_N}{ds^\alpha} = \frac{d^\alpha}{ds^\alpha} \left( \sum_{\nu=0}^N x^{(\nu)}(0) s^{N-\nu} \right). \quad (43)$$

The time derivatives, i.e.  $s^\mu \frac{d^\mu X_N}{ds^\mu}$ ,  $\mu = 1, \dots, N$ ,  $0 \leq \mu \leq N$  are removed by multiplying both sides of Equation (17) by  $s^{-N}$ ,  $N > N$ .

- For an arbitrary analytic time function, let us apply the preceding calculations to a suitable truncated Taylor expansion. Consider a real-valued analytic time function defined by the convergent power series  $x(t) = \sum_{\nu=0}^{\infty} x^{(\nu)}(0) \frac{t^\nu}{\nu!}$ , where  $0 \leq t < \rho$ . Approximate  $x(t)$  in the interval  $(0, \varepsilon)$ ,  $0 < \varepsilon \leq \rho$  by its truncated Taylor expansion  $x_N(t) = \sum_{\nu=0}^N x^{(\nu)}(0) \frac{t^\nu}{\nu!}$  of order  $N$ . Introduce the operational analogue of  $x(t)$ , i.e.  $X(s) = \sum_{\nu \geq 0} \frac{x^{(\nu)}(0)}{s^{\nu+1}}$ . Denote by  $[x^{(\nu)}(0)]_{e_N}(t)$ ,  $0 \leq \nu \leq N$ , the numerical estimate of  $x^{(\nu)}(0)$ , which is obtained by replacing  $X_N(s)$  by  $X(s)$  in Eq. (17). It can be shown [104] that a good estimate is obtained in this way.

Thus using elementary differential algebraic operations, we derive an explicit formulae yielding point-wise derivative estimation for each given order. Interesting enough, it turns out that the Jacobi orthogonal polynomials [129] are inherently connected with the developed algebraic numerical differentiators. A least-squares interpretation then naturally follows [118], [119] and this leads to a key result: the algebraic numerical differentiation is as efficient as an appropriately chosen time delay. Though, such a delay may not be tolerable in some real-time applications. Moreover, instability generally occurs when introducing delayed signals in a control loop. Note however that since the delay is known *a priori*, it is always possible to derive a control law, which compensates for its effects (see [127]). A second key feature of the algebraic numerical differentiators is its very low complexity which allows for a real-time implementation. Indeed, the  $n^{th}$  order derivative estimate (that can be directly managed for  $n \geq 2$ , without using  $n$  cascaded estimators) is expressed as the output of the linear time-invariant filter, with finite support impulse response  $h_{\kappa, \mu, n, r}(\cdot)$ . Implementing such a stable and causal filter is easy and simple. This is achieved either in continuous-time or in discrete-time when only discrete-time samples of the observation are available. In the latter case, we obtain a tapped delay line digital filter by considering any numerical integration method with equally-spaced abscissas.

### 3.2.2. Model-based estimation of derivatives

If we assume that the derivatives to be estimated are unmeasured states of a process that generates the signal, then the differentiation techniques can be considered as left invertibility algorithms. In this sense, the previous algebraic estimation achieves a “model-free” left inversion. Now, when such a model is available, the *finite-time observers* relying on higher order sliding modes [123] and homogeneity properties [124], [120] also represent possible non-asymptotic algorithms for differentiation<sup>10</sup>. Using such model-based techniques appears to be complementary<sup>11</sup>. The left-inversion results have been already obtained for several classes of models: linear systems [106], nonlinear systems [92], delay systems [2] and hybrid systems [114].

---

<sup>10</sup>Usually, observer design yields asymptotic convergence of the estimation error dynamics. The main advantages of such a technique in the case of linear systems are simplicity of design, estimation with a filtering action and global stability property. Nevertheless, the filtering property is not ensured for nonlinear systems and the stability property is generally obtained only locally. For these reasons, in the case of nonlinear systems, finite-time observers and estimators have been proposed in the literature [116], [124], [125], [105]...

<sup>11</sup>The choice between the two approaches will be done after comparison with respect to the indicators 1–3, and taking into account the application (for instance, the system bandwidth, system dimension), the kind of discontinuity, the observer in the control loop or not...

## CLASSIC Project-Team

### 3. Scientific Foundations

#### 3.1. Regression models of supervised learning

The most obvious contribution of statistics to machine learning is to consider the supervised learning scenario as a special case of regression estimation: given  $n$  independent pairs of observations  $(X_i, Y_i)$ ,  $i = 1, \dots, n$ , the aim is to “learn” the dependence of  $Y_i$  on  $X_i$ . Thus, classical results about statistical regression estimation apply, with the caveat that the hypotheses we can reasonably assume about the distribution of the pairs  $(X_i, Y_i)$  are much weaker than what is usually considered in statistical studies. The aim here is to assume very little, maybe only independence of the observed sequence of input-output pairs, and to validate model and variable selection schemes. These schemes should produce the best possible approximation of the joint distribution of  $(X_i, Y_i)$  within some restricted family of models. Their performance is evaluated according to some measure of discrepancy between distributions, a standard choice being to use the Kullback-Leibler divergence.

##### 3.1.1. PAC-Bayes inequalities

One of the specialties of the team in this direction is to use PAC-Bayes inequalities to combine thresholded exponential moment inequalities. The name of this theory comes from its founder, David McAllester, and may be misleading. Indeed, its cornerstone is rather made of non-asymptotic entropy inequalities, and a perturbative approach to parameter estimation. The team has made major contributions to the theory, first focussed on classification [6], then on regression [1]. It has introduced the idea of combining the PAC-Bayesian approach with the use of thresholded exponential moments, in order to derive bounds under very weak assumptions on the noise.

##### 3.1.2. Sparsity and $\ell_1$ -regularization

Another line of research in regression estimation is the use of sparse models, and its link with  $\ell_1$ -regularization. Regularization is the joint minimization of some empirical criterion and some penalty function; it should lead to a model that not only fits well the data but is also as simple as possible.

For instance, the Lasso uses a  $\ell^1$ -regularization instead of a  $\ell^0$ -one; it is popular mostly because it leads to *sparse* solutions (the estimate has only a few nonzero coordinates), which usually have a clear interpretation in many settings (e.g., the influence or lack of influence of some variables). In addition, unlike  $\ell^0$ -penalization, the Lasso is *computationally feasible* for high-dimensional data.

##### 3.1.3. Pushing it to the extreme: no assumption on the data

The next brick of our scientific foundations explains why and how, in certain cases, we may formulate absolutely no assumption on the data  $(x_i, y_i)$ ,  $i = 1, \dots, n$ , which is then considered a deterministic set of input-output pairs.

#### 3.2. On-line aggregation of predictors for the prediction of time series, with or without stationarity assumptions

We are concerned here with *sequential prediction* of outcomes, given some base predictions formed by *experts*. We distinguish two settings, depending on how the sequence of outcomes is generated: it is either

- the realization of some stationary process,
- or is not modeled at all as the realization of any underlying stochastic process (these sequences are called *individual sequences*).

The aim is to predict almost as well as the best expert. Typical good forecasters maintain one weight per expert, update these weights depending on the past performances, and output at each step the corresponding weighted linear combination of experts' advices.

The difference between the cumulative prediction error of the forecaster and the one of the best expert is called the regret. The goal here is to upper bound the regret by a quantity as small as possible.

### **3.3. Multi-armed bandit problems, prediction with limited feedback**

We are interested in settings in which the feedback obtained on the predictions is limited, in the sense that it does not fully reveal what actually happened.

#### **3.3.1. Bandit problems**

This is also a sequential problem in which some regret is to be minimized.

However, this problem is a stochastic problem: a large number of arms, possibly indexed by a continuous set like  $[0, 1]$ , is available. Each arm is associated with a fixed but unknown distribution. At each round, the player chooses an arm, a payoff is drawn at random according to the distribution that is associated with it, and the only feedback that the player gets is the value of this payoff. The key quantity to study this problem is the mean-payoff function  $f$ , that indicates for each arm  $x$  the expected payoff  $f(x)$  of the distribution that is associated with it. The target is to minimize the regret, i.e., ensure that the difference between the cumulative payoff obtained by the player and the one of the best arm is small.

#### **3.3.2. A generalization of the regret: the approachability of sets**

Approachability is the ability to control random walks. At each round, a vector payoff is obtained by the first player, depending on his action and on the action of the opponent player. The aim is to ensure that the average of the vector payoffs converges to some convex set. Necessary and sufficient conditions were obtained by Blackwell and others to ensure that such strategies exist, both in the full information and in the bandit cases.

Some of these results can be extended to the case of games with signals (games with partial monitoring), where at each round the only feedback obtained by the first player is a random signal drawn according to a distribution that depends on the action profile taken by the two players, while the opponent player still has a full monitoring.

## **DOLPHIN Project-Team**

### **3. Scientific Foundations**

#### **3.1. Modeling and landscape analysis**

The modeling of problems, the analysis of structures (landscapes) of MOPs and the performance assessment of resolution methods are significant topics in the design of optimization methods. The effectiveness of metaheuristics depends on the properties of the problem and its landscape (roughness, convexity, etc). The notion of landscape has been first described in [89] by the way of the study of species evolution. Then, this notion has been used to analyze combinatorial optimization problems.

##### **3.1.1. Modeling of problems**

Generally there are several ways of modeling a given problem. First, one has to find the most suitable model for the type of resolution he or she plans to use. The choice can be made after a theoretical analysis of the model, or after computational experiments. The choice of the model depends on the type of method used. For example, a major issue in the design of exact methods is to find tight relaxations for the problem considered.

Let us note that many combinatorial optimization problems of the literature have been studied in their mono-objective form even if a lot of them are naturally of a multi-objective nature.

Therefore, in the DOLPHIN project, we address the modeling of MOPs in two phases. The first one consists in studying the mono-objective version of the problem, where all objectives but one are considered as constraints. In the second phase, we propose methods to adapt the mono-objective models or to create hand-tailored models for the multi-objective case. The models used may come from the first phase, or from the literature.

##### **3.1.2. Analysis of the structure of a problem**

The landscape is defined by a neighborhood operator and can be represented by a graph  $G = (V, E)$ . The vertices represent the solutions of the problem and an edge  $(e_1, e_2)$  exists if the solution  $e_2$  can be obtained by an application of the neighborhood operator on the solution  $e_1$ . Then, considering this graph as the ground floor, we elevate each solution to an altitude equals to its cost. We obtain a surface, or landscape, made of peaks, valleys, plateaus, cliffs, etc. The problem lies in the difficulty to have a realistic view of this landscape.

Like others, we believe that the main point of interest in the domain of combinatorial optimization is not the design of the best algorithm for a large number of problems but the search for the most adapted method to an instance or a set of instances of a given problem. Therefore, we are convinced that no ideal metaheuristic, designed as a black-box, may exist.

Indeed, the first studies realized in our research group on the analysis of landscapes of different mono-objective combinatorial optimization problems (traveling salesman problem, quadratic assignment problem) have shown that not only different problems correspond to different structures but also that different instances of the same problem correspond to different structures.

For instance, we have realized a statistical study of the landscapes of the quadratic assignment problem. Some indicators that characterize the landscape of an instance have been proposed and a taxonomy of the instances including three classes has been deduced. Hence it is not enough to adapt the method to the problem under study but it is necessary to specialize it according to the type of the treated instance.

So in its studies of mono-objective problems, the DOLPHIN research group has introduced into the resolution methods some information about the problem to be solved. The landscapes of some combinatorial problems have been studied in order to investigate the intrinsic natures of their instances. The resulting information has been inserted into an optimization strategy and has allowed the design of efficient and robust hybrid methods. The extension of these studies to multi-objective problems is a part of the DOLPHIN project [87], [88].

### 3.1.3. Performance assessment

The DOLPHIN project is also interested in the performance assessment of multi-objective optimization methods. Nowadays, statistical techniques developed for mono-objective problems can be adapted to the multi-objective case. Nevertheless, specific tools are necessary in many situations: for example, the comparison of two different algorithms is relatively easy in the mono-objective case - we compare the quality of the best solution obtained in a fixed time, or the time needed to obtain a solution of a certain quality. The same idea cannot be immediately transposed to the case where the output of the algorithms is a set of solutions having several quality measures, and not a single solution.

Various indicators have been proposed in the literature for evaluating the performance of multi-objective optimization methods but no indicator seems to outperform the others [90]. The DOLPHIN research group has proposed two indicators: the *contribution* and the *entropy* [82]. The contribution evaluates the supply in term of Pareto-optimal solutions of a front compared to another one. The entropy gives an idea of the diversity of the solutions found. These two metrics are used to compare the different metaheuristics in the research group, for example in the resolution of the bi-objective flow-shop problem, and also to show the contribution of the various mechanisms introduced in these metaheuristics.

### 3.1.4. Goals

One of the main issues in the DOLPHIN project is the study of the landscape of multi-objective problems and the performance assessment of multi-objective optimization methods to design efficient and robust resolution methods:

- *Landscape study*: The goal here is to extend the study of landscapes of the mono-objective combinatorial optimization problems to multi-objective problems in order to determine the structure of the Pareto frontier and to integrate this knowledge about the problem structure in the design of resolution methods.

This study has been initiated for the bi-objective flow-shop problem. We have studied the convexity of the frontiers obtained in order to show the interest of our Pareto approach compared to an aggregation approach, which only allows one to obtain the Pareto solutions situated on the convex hull of the Pareto front (supported solutions).

Our preliminary study of the landscape of the bi-objective flow-shop problem shows that the supported solutions are very closed to each other. This remark leads us to improve an exact method initially proposed for bi-objective problems. Furthermore, a new exact method able to deal with any number of objectives has been designed.

- *Performance assessment*: The goal here is to extend *GUIMOO* in order to provide efficient visual and metric tools for evaluating the assessment of multi-objective resolution methods.

## 3.2. Hybrid multi-objective optimization methods

The success of metaheuristics is based on their ability to find efficient solutions in a reasonable time [76]. But with very large problems and/or multi-objective problems, efficiency of metaheuristics may be compromised. Hence, in this context it is necessary to integrate metaheuristics in more general schemes in order to develop even more efficient methods. For instance, this can be done by different strategies such as cooperation and parallelization.

The DOLPHIN project deals with “*a posteriori*” multi-objective optimization where the set of Pareto solutions (solutions of best compromise) have to be generated in order to give the decision maker the opportunity to choose the solution that interests him/her.

Population-based methods, such as evolutionary algorithms, are well fitted for multi-objective problems, as they work with a set of solutions [59], [74]. To be convinced one may refer to the list of references on Evolutionary Multi-objective Optimization maintained by Carlos A. Coello Coello <sup>4</sup>, which contains more

<sup>4</sup><http://www.lania.mx/~ccoello/EMOO/EMOObib.html>

than 5500 references. One of the objectives of the project is to propose advanced search mechanisms for intensification and diversification. These mechanisms have been designed in an adaptive manner, since their effectiveness is related to the landscape of the MOP and to the instance solved.

In order to assess the performances of the proposed mechanisms, we always proceed in two steps: first, we carry out experiments on academic problems, for which some best known results exist; second, we use real industrial problems to cope with large and complex MOPs. The lack of references in terms of optimal or best known Pareto set is a major problem. Therefore, the obtained results in this project and the test data sets will be available at the URL <http://dolphin.lille.inria.fr/> at 'benchmark'.

### 3.2.1. Cooperation of metaheuristics

In order to benefit from the various advantages of the different metaheuristics, an interesting idea is to combine them. Indeed, the hybridization of metaheuristics allows the cooperation of methods having complementary behaviors. The efficiency and the robustness of such methods depend on the balance between the exploration of the whole search space and the exploitation of interesting areas.

Hybrid metaheuristics have received considerable interest these last years in the field of combinatorial optimization. A wide variety of hybrid approaches have been proposed in the literature and give very good results on numerous single objective optimization problems, which are either academic (traveling salesman problem, quadratic assignment problem, scheduling problem, etc) or real-world problems. This efficiency is generally due to the combinations of single-solution based methods (iterative local search, simulated annealing, tabu search, etc) with population-based methods (genetic algorithms, ants search, scatter search, etc). A taxonomy of hybridization mechanisms may be found in [84]. It proposes to decompose these mechanisms into four classes:

- *LRH class - Low-level Relay Hybrid*: This class contains algorithms in which a given metaheuristic is embedded into a single-solution metaheuristic. Few examples from the literature belong to this class.
- *LTH class - Low-level Teamwork Hybrid*: In this class, a metaheuristic is embedded into a population-based metaheuristic in order to exploit strengths of single-solution and population-based metaheuristics.
- *HRH class - High-level Relay Hybrid*: Here, self contained metaheuristics are executed in a sequence. For instance, a population-based metaheuristic is executed to locate interesting regions and then a local search is performed to exploit these regions.
- *HTH class - High-level Teamwork Hybrid*: This scheme involves several self-contained algorithms performing a search in parallel and cooperating. An example will be the island model, based on GAs, where the population is partitioned into small subpopulations and a GA is executed per subpopulation. Some individuals can migrate between subpopulations.

Let us notice, that if hybrid methods have been studied in the mono-criterion case, their application in the multi-objective context is not yet widely spread. The objective of the DOLPHIN project is to integrate specificities of multi-objective optimization into the definition of hybrid models.

### 3.2.2. Cooperation between metaheuristics and exact methods

Until now only few exact methods have been proposed to solve multi-objective problems. They are based either on a Branch-and-bound approach, on the algorithm  $A^{\star}$ , or on dynamic programming. However, these methods are limited to two objectives and, most of the time, cannot be used on a complete large scale problem. Therefore, sub search spaces have to be defined in order to use exact methods. Hence, in the same manner as hybridization of metaheuristics, the cooperation of metaheuristics and exact methods is also a main issue in this project. Indeed, it allows us to use the exploration capacity of metaheuristics, as well as the intensification ability of exact methods, which are able to find optimal solutions in a restricted search space. Sub search spaces have to be defined along the search. Such strategies can be found in the literature, but they are only applied to mono-objective academic problems.

We have extended the previous taxonomy for hybrid metaheuristics to the cooperation between exact methods and metaheuristics. Using this taxonomy, we are investigating cooperative multi-objective methods. In this context, several types of cooperations may be considered, according to the way the metaheuristic and the exact method cooperate. For instance, a metaheuristic can use an exact method for intensification or an exact method can use a metaheuristic to reduce the search space.

Moreover, a part of the DOLPHIN project deals with studying exact methods in the multi-objective context in order: i) to be able to solve small size problems and to validate proposed heuristic approaches; ii) to have more efficient/dedicated exact methods that can be hybridized with metaheuristics. In this context, the use of parallelism will push back limits of exact methods, which will be able to explore larger size search spaces [68].

### 3.2.3. Goals

Based on the previous works on multi-objective optimization, it appears that to improve metaheuristics, it becomes essential to integrate knowledge about the problem structure. This knowledge can be gained during the search. This would allow us to adapt operators which may be specific for multi-objective optimization or not. The goal here is to design auto-adaptive methods that are able to react to the problem structure. Moreover, regarding the hybridization and the cooperation aspects, the objectives of the DOLPHIN project are to deepen these studies as follows:

- *Design of metaheuristics for the multi-objective optimization:* To improve metaheuristics, it becomes essential to integrate knowledge about the problem structure, which we may get during the execution. This would allow us to adapt operators that may be specific for multi-objective optimization or not. The goal here is to design auto-adaptive methods that are able to react to the problem structure.
- *Design of cooperative metaheuristics:* Previous studies show the interest of hybridization for a global optimization and the importance of problem structure study for the design of efficient methods. It is now necessary to generalize hybridization of metaheuristics and to propose adaptive hybrid models that may evolve during the search while selecting the appropriate metaheuristic. Multi-objective aspects have to be introduced in order to cope with the specificities of multi-objective optimization.
- *Design of cooperative schemes between exact methods and metaheuristics:* Once the study on possible cooperation schemes is achieved, we will have to test and compare them in the multi-objective context.
- *Design and conception of parallel metaheuristics:* Our previous works on parallel metaheuristics allow us to speed up the resolution of large scale problems. It could be also interesting to study the robustness of the different parallel models (in particular in the multi-objective case) and to propose rules that determine, given a specific problem, which kind of parallelism to use. Of course these goals are not disjointed and it will be interesting to simultaneously use hybrid metaheuristics and exact methods. Moreover, those advanced mechanisms may require the use of parallel and distributed computing in order to easily make cooperating methods evolve simultaneously and to speed up the resolution of large scale problems.
- *Validation:* In order to validate the obtained results we always proceed in two phases: validation on academic problems, for which some best known results exist and use on real problems (industrial) to cope with problem size constraints.

Moreover, those advanced mechanisms are to be used in order to integrate the distributed multi-objective aspects in the ParadisEO platform (see the paragraph on software platform).

## 3.3. Parallel multi-objective optimization: models and software frameworks

Parallel and distributed computing may be considered as a tool to speedup the search to solve large MOPs and to improve the robustness of a given method. Moreover, the joint use of parallelism and cooperation allows improvements on the quality of the obtained Pareto sets. Following this objective, we will design and implement parallel models for metaheuristics (evolutionary algorithms, tabu search approach) and exact methods (branch-and-bound algorithm, branch-and-cut algorithm) to solve different large MOPs.



One of the goal of the DOLPHIN project is to integrate the developed parallel models into software frameworks. Several frameworks for parallel distributed metaheuristics have been proposed in the literature. Most of them focus only either on evolutionary algorithms or on local search methods. Only few frameworks are dedicated to the design of both families of methods. On the other hand, existing optimization frameworks either do not provide parallelism at all or just supply at most one parallel model. In this project, a new framework for parallel hybrid metaheuristics is proposed, named *Parallel and Distributed Evolving Objects (ParadisEO)* based on EO. The framework provides in a transparent way the hybridization mechanisms presented in the previous section, and the parallel models described in the next section. Concerning the developed parallel exact methods for MOPs, we will integrate them into well-known frameworks such as COIN.

### 3.3.1. Parallel models

According to the family of addressed metaheuristics, we may distinguish two categories of parallel models: parallel models that manage a single solution, and parallel models that handle a population of solutions. The major single solution-based parallel models are the following: the *parallel neighborhood exploration model* and the *multi-start model*.

- *The parallel neighborhood exploration model* is basically a "low level" model that splits the neighborhood into partitions that are explored and evaluated in parallel. This model is particularly interesting when the evaluation of each solution is costly and/or when the size of the neighborhood is large. It has been successfully applied to the mobile network design problem (see Application section).
- *The multi-start model* consists in executing in parallel several local searches (that may be heterogeneous), without any information exchange. This model raises particularly the following question: is it equivalent to execute  $k$  local searches during a time  $t$  than executing a single local search during  $k \times t$ ? To answer this question we tested a multi-start Tabu search on the quadratic assignment problem. The experiments have shown that the answer is often landscape-dependent. For example, the multi-start model may be well-suited for landscapes with multiple basins.

Parallel models that handle a population of solutions are mainly: the *island model*, the *central model* and the *distributed evaluation of a single solution*. Let us notice that the last model may also be used with single-solution metaheuristics.

- In the *island model*, the population is split into several sub-populations distributed among different processors. Each processor is responsible of the evolution of one sub-population. It executes all the steps of the metaheuristic from the selection to the replacement. After a given number of generations (synchronous communication), or when a convergence threshold is reached (asynchronous communication), the migration process is activated. Then, exchanges of solutions between sub-populations are realized, and received solutions are integrated into the local sub-population.
- *The central (Master/Worker) model* allows us to keep the sequentiality of the original algorithm. The master centralizes the population and manages the selection and the replacement steps. It sends sub-populations to the workers that execute the recombination and evaluation steps. The latter returns back newly evaluated solutions to the master. This approach is efficient when the generation and evaluation of new solutions is costly.
- *The distributed evaluation model* consists in a parallel evaluation of each solution. This model has to be used when, for example, the evaluation of a solution requires access to very large databases (data mining applications) that may be distributed over several processors. It may also be useful in a multi-objective context, where several objectives have to be computed simultaneously for a single solution.

As these models have now been identified, our objective is to study them in the multi-objective context in order to use them advisedly. Moreover, these models may be merged to combine different levels of parallelism and to obtain more efficient methods [73], [83].

### 3.3.2. Goals

Our objectives focus on these issues are the following:

- *Design of parallel models for metaheuristics and exact methods for MOPs:* We will develop parallel cooperative metaheuristics (evolutionary algorithms and local search algorithms such as the Tabu search) for solving different large MOPs. Moreover, we are designing a new exact method, named PPM (Parallel Partition Method), based on branch and bound and branch and cut algorithms. Finally, some parallel cooperation schemes between metaheuristics and exact algorithms have to be used to solve MOPs in an efficient manner.
- *Integration of the parallel models into software frameworks:* The parallel models for metaheuristics will be integrated in the ParadisEO software framework. The proposed multi-objective exact methods must be first integrated into standard frameworks for exact methods such as COIN and BOB++. A *coupling* with ParadisEO is then needed to provide hybridization between metaheuristics and exact methods.
- *Efficient deployment of the parallel models on different parallel and distributed architecture including GRIDs:* The designed algorithms and frameworks will be efficiently deployed on non-dedicated networks of workstations, dedicated cluster of workstations and SMP (Symmetric Multi-processors) machines. For GRID computing platforms, peer to peer (P2P) middlewares (XtremWeb-Condor) will be used to implement our frameworks. For this purpose, the different optimization algorithms may be re-visited for their efficient deployment.

## GEOSTAT Project-Team

### 3. Scientific Foundations

#### 3.1. Dynamics of complex systems

GEOSTAT is studying complex signals under the point of view of *nonlinear* methods, in the sense of *nonlinear physics* i.e. the methodologies developed to study complex systems, with a strong emphasis on multiresolution analysis. Linear methods in signal processing refer to the standard point of view under which operators are expressed by simple convolutions with impulse responses. Linear methods in signal processing are widely used, from least-square deconvolution methods in adaptive optics to source-filter models in speech processing. Linear methods do not unlock the multiscale structures and cascading variables of primary importance as previewed by the physics of the phenomena. This is the reason why new approaches, such as DFA (Detrended Fluctuation Analysis), Time-frequency analysis, variations on curvelets [38] etc. have appeared during the last decades. One important result obtained in GEOSTAT is the effective use of multiresolution analysis associated to optimal inference along the scales of a complex system. The multiresolution analysis is performed on dimensionless quantities given by the *singularity exponents* which encode properly the geometrical structures associated to multiscale organization. This is applied successfully in the derivation of high resolution ocean dynamics, or the high resolution mapping of gaseous exchanges between the ocean and the atmosphere; the latter is of primary importance for a quantitative evaluation of global warming. Understanding the dynamics of complex systems is recognized as a new discipline, which makes use of theoretical and methodological foundations coming from nonlinear physics, the study of dynamical systems and many aspects of computer science. One of the challenges is related to the question of *emergence* in complex systems: large-scale effects measurable macroscopically from a system made of huge numbers of interactive agents [29], [26], [43], [34]. Some quantities related to nonlinearity, such as Lyapunov exponents, Kolmogorov-Sinai entropy etc. can be computed at least in the phase space [27]. Consequently, knowledge from acquisitions of complex systems (which include *complex signals*) could be obtained from information about the phase space. A result from F. Takens [39] about strange attractors in turbulence has motivated the determination of discrete dynamical systems associated to time series [31], and consequently the theoretical determination of nonlinear characteristics associated to complex acquisitions. Emergence phenomena can also be traced inside complex signals themselves, by trying to localize information content geometrically. Fundamentally, in the nonlinear analysis of complex signals there are broadly two approaches: characterization by attractors (embedding and bifurcation) and time-frequency, multiscale/multiresolution approaches. Time-frequency analysis [28] and multiscale/multiresolution are the subjects of intense research and are profoundly reshaping the analysis of complex signals by nonlinear approaches [25], [30]. In real situations, the phase space associated to the acquisition of a complex phenomenon is unknown. It is however possible to relate, inside the signal's domain, local predictability to local reconstruction and deduce from that singularity exponents (SEs) [8] [5]. The SEs are defined at any point in the signal's domain, they relate, but are different, to other kinds of exponents used in the nonlinear analysis of complex signals. We are working on their relation with:

- properties in universality classes,
- the geometric localization of multiscale properties in complex signals,
- cascading characteristics of physical variables,
- optimal wavelets and inference in multiresolution analysis.

The alternative approach taken in GEOSTAT is microscopical, or geometrical: the multiscale structures which have their "fingerprint" in complex signals are being isolated in a single realization of the complex system, i.e. using the data of the signal itself, as opposed to the consideration of grand ensembles or a wide set of realizations. This is much harder than the ergodic approaches, but it is possible because a reconstruction formula such as the one derived in [40] is local and reconstruction in the signal's domain is related to predictability.

Nonlinear signal processing is making use of quantities related to predictability. For instance the first Lyapunov exponent  $\lambda_1$  is related, from Osedelec's theorem, to the limiting behaviour of the response, after a time  $t$ , to perturbation in the phase space  $\log R_\tau(t)$ :

$$\lambda_1 = \lim_{t \rightarrow \infty} \frac{1}{t} \langle \log R_\tau(t) \rangle \quad (44)$$

with  $\langle \cdot \rangle$  being time average and  $R_\tau$  the response to a perturbation [27]. More refined information is provided by the Kolmogorov-Sinai entropy:

$$h_{KS} = \lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow \infty} \frac{1}{t} \log N_{\text{eff}}(\varepsilon, t) \quad (45)$$

( $N_{\text{eff}}(\varepsilon, t)$  is related to events which appear with very high probability in long time). In GEOSTAT our aim is to relate these classical quantities (among others) to the behaviour of SEs, which are defined by a limiting behaviour

$$\mu(\mathcal{B}_r(\mathbf{x})) = \alpha(\mathbf{x}) r^{d+h(\mathbf{x})} + o(r^{d+h(\mathbf{x})}) \quad (r \rightarrow 0) \quad (46)$$

( $d$ : dimension of the signal's domain,  $\mu$ : multiscale measure, typically whose density is the gradient's norm,  $\mathcal{B}_r(\mathbf{x})$ : ball of radius  $r$  centered at  $\mathbf{x}$ ). For precise computation, SEs can be smoothly interpolated by projecting wavelets:

$$\mathcal{J}_\Psi \mu(\mathbf{x}, r) = \int_{\mathbb{R}^d} d\mu(\mathbf{x}') \frac{1}{r^d} \Psi\left(\frac{\mathbf{x} - \mathbf{x}'}{r}\right) \quad (47)$$

( $\Psi$ : mother wavelet, admissible or not). SEs are related to the framework of reconstructible systems, and consequently to predictability. They unlock the geometric localization of multiscale structures in a complex signal:

$$\mathcal{F}_h = \{\mathbf{x} \in \Omega \mid h(\mathbf{x}) = h\}, \quad (48)$$

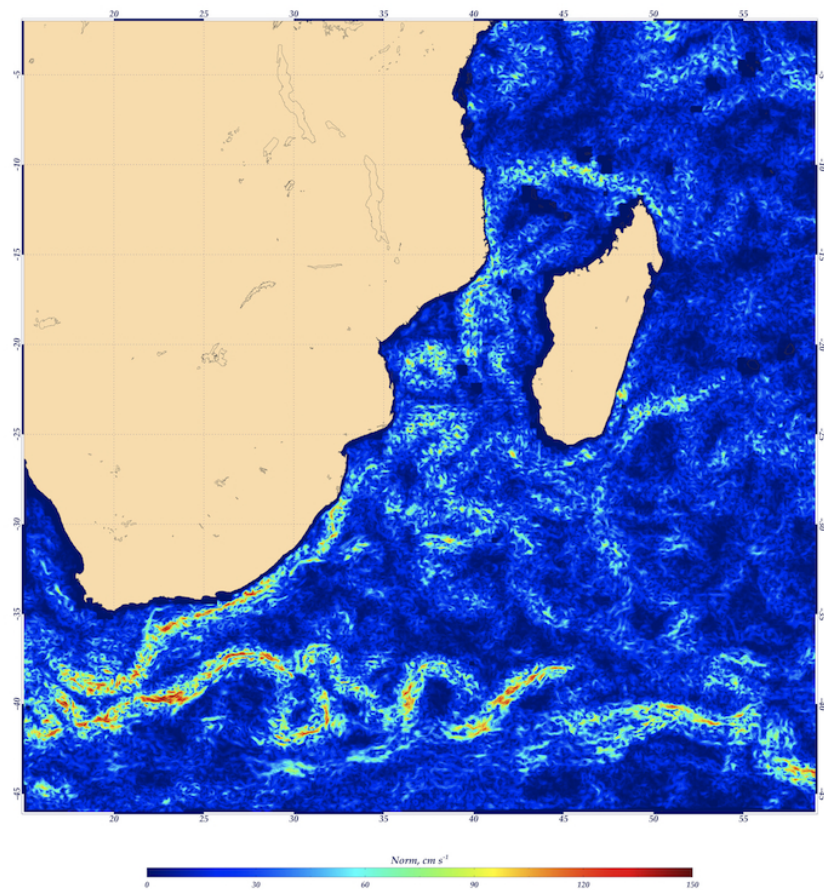
( $\Omega$ : signal's domain) and are consequently in relation with *optimal wavelets*:

$$\mathcal{J}_\psi[\mathbf{s}](\mathbf{x}, \mathbf{r}_1) = \zeta_{r_1/r_2}(\mathbf{x}) \mathcal{J}_\psi[\mathbf{s}](\mathbf{x}, \mathbf{r}_2) \quad (49)$$

( $\mathbf{r}_1 < \mathbf{r}_2$ : two scales of observation,  $\zeta$ : injection variable between the scales,  $\psi$ : optimal wavelet) and their **multiresolution analysis**. They are related to persistence along the scales and lead to multiresolution analysis whose coefficients verify

$$\alpha_s = \eta_1 \alpha_f + \eta_2 \quad (50)$$

with  $\alpha_s$  and  $\alpha_f$  referring to child and parent coefficients,  $\eta_1$  and  $\eta_2$  are random variables independent of  $\alpha_s$  and  $\alpha_f$  and also independent of each other.



*Figure 1. Visualization of the motion field computed at high spatial resolution (pixel size: 4kms) over a wide area around South Africa. The ocean dynamics is computed by propagating low resolution information coming from altimetry data (pixel size: 24 kms) along approximated optimal multiresolution analysis computed over the singularity exponents of Sea Surface Temperature data obtained from MODIS AQUA and OSTIA. Common work between GEOSTAT and DYNBIO (LEGOS, CNRS UMR 55 66, Toulouse).*

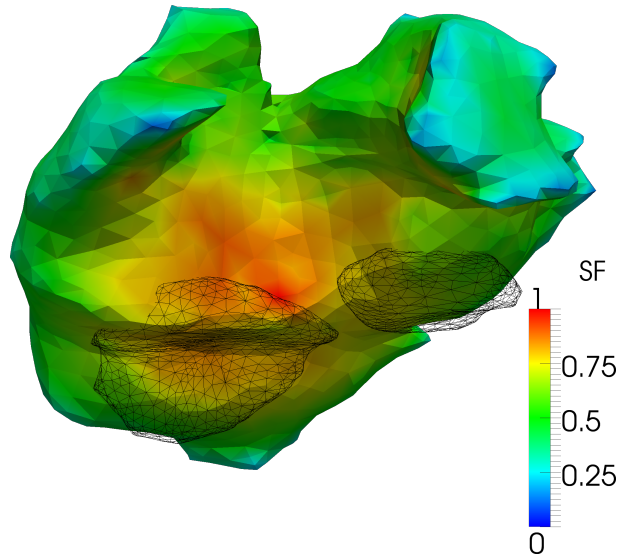


Figure 2. Result of the computation of a normalized source field over the 3D epicardial surface of the atria, from electric potential data acquired on a regular grid of electrodes placed on a patient's chest. There is a strong correlation between the red parts of the source field and the locations inside the heart where fibrillation occurs. Inputted data courtesy of IHU LIRYC.

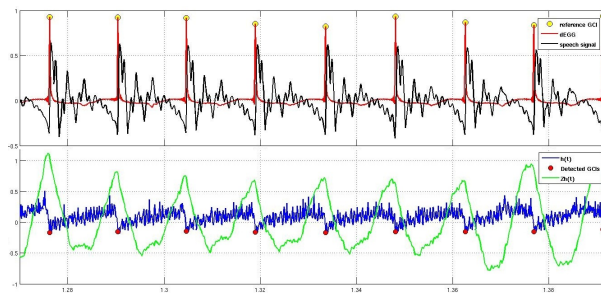


Figure 3. Top: A segment of a voiced speech signal (in black) along with the differentiated EGG (dEGG) recording (in red). Local maxima of dEGG shows the reference GCIs (yellow circles). Bottom: The singularity exponents (in blue) along with an auxiliary functional (in green) defined as  $Zh(t) = \sum_{u=t-T_L}^{t-\delta t} h(u) - \sum_{u=t}^{t+T_L} h(u)$  ( $h(t)$ : singularity exponent at  $t$ ). In each positive half-period of  $Zh(t)$ , the minimum of singularity exponents is taken as the GCI (red circle).

In a first example we give some insight about the collaboration with LEGOS Dynbio team <sup>1</sup> about high-resolution ocean dynamics from microcanonical formulations in nonlinear complex signal analysis. LPEs relate to the geometric structures linked with the cascading properties of indefinitely divisible variables in turbulent flows. Cascading properties can be represented by optimal wavelets (OWs); this opens new and fascinating directions of research for the determination of ocean motion field at high spatial resolution. OWs in a microcanonical sense pave the way for the determination of the energy injection mechanisms between the scales. From this results a new method for the complete evaluation of oceanic motion field is introduced; it consists in propagating along the scales the norm and the orientation of ocean dynamics deduced at low spatial resolution (geostrophic from altimetry and a part of ageostrophic from wind stress products). Using this approach, there is no need to use several temporal occurrences. Instead, the proper determination of the turbulent cascading and energy injection mechanisms in oceanographic signals allows the determination of oceanic motion field at the SST or Ocean colour spatial resolution (pixel size: 4 kms). We use the Regional Ocean Modelling System (ROMS) to validate the results on simulated data and compare the motion fields obtained with other techniques. See figure 1 .

In a second example, we show in figure 2 the highly promising results obtained in the application of nonlinear signal processing and multiscale techniques to the localization of heart fibrillation phenomenon acquired from a real patient and mapped over a reconstructed 3D surface of the heart. The notion of *source field*, defined in GEOSTAT from the computation of derivative measures related to the singularity exponents allows the localization of arhythmic phenomena inside the heart [6].

Our last example is about speech. In speech analysis, we use the concept of the Most Singular Manifold (MSM) to localize critical events in domain of this signal. We show that in case of voiced speech signals, the MSM coincides with the instants of significant excitation of the vocal tract system. It is known that these major excitations occur when the glottis is closed, and hence, they are called the Glottal Closure Instants (GCI). We use the MSM to develop a reliable and noise robust GCI detection algorithm and we evaluate our algorithm using contemporaneous Electro-Glotto-Graph (EGG) recordings. See figure 3 .

---

<sup>1</sup><http://www.legos.obs-mip.fr/recherches/equipes/dynbio>.

## MISTIS Project-Team

### 3. Scientific Foundations

#### 3.1. Mixture models

**Participants:** Angelika Studeny, Thomas Vincent, Christine Bakhous, Lotfi Chaari, Senan James Doyle, Jean-Baptiste Durand, Florence Forbes, Stéphane Girard, Marie-José Martinez, Darren Wraith.

mixture of distributions, EM algorithm, missing data, conditional independence, statistical pattern recognition, clustering, unsupervised and partially supervised learning

In a first approach, we consider statistical parametric models,  $\theta$  being the parameter, possibly multi-dimensional, usually unknown and to be estimated. We consider cases where the data naturally divides into observed data  $y = y_1, \dots, y_n$  and unobserved or missing data  $z = z_1, \dots, z_n$ . The missing data  $z_i$  represents for instance the memberships of one of a set of  $K$  alternative categories. The distribution of an observed  $y_i$  can be written as a finite mixture of distributions,

$$f(y_i | \theta) = \sum_{k=1}^K P(z_i = k | \theta) f(y_i | z_i, \theta). \quad (51)$$

These models are interesting in that they may point out hidden variable responsible for most of the observed variability and so that the observed variables are *conditionally* independent. Their estimation is often difficult due to the missing data. The Expectation-Maximization (EM) algorithm is a general and now standard approach to maximization of the likelihood in missing data problems. It provides parameter estimation but also values for missing data.

Mixture models correspond to independent  $z_i$ 's. They are increasingly used in statistical pattern recognition. They enable a formal (model-based) approach to (unsupervised) clustering.

#### 3.2. Markov models

**Participants:** Angelika Studeny, Thomas Vincent, Christine Bakhous, Lotfi Chaari, Senan James Doyle, Jean-Baptiste Durand, Florence Forbes, Darren Wraith.

graphical models, Markov properties, conditional independence, hidden Markov trees, clustering, statistical learning, missing data, mixture of distributions, EM algorithm, stochastic algorithms, selection and combination of models, statistical pattern recognition, image analysis, hidden Markov field, Bayesian inference

Graphical modelling provides a diagrammatic representation of the logical structure of a joint probability distribution, in the form of a network or graph depicting the local relations among variables. The graph can have directed or undirected links or edges between the nodes, which represent the individual variables. Associated with the graph are various Markov properties that specify how the graph encodes conditional independence assumptions.

It is the conditional independence assumptions that give graphical models their fundamental modular structure, enabling computation of globally interesting quantities from local specifications. In this way graphical models form an essential basis for our methodologies based on structures.



The graphs can be either directed, e.g. Bayesian Networks, or undirected, e.g. Markov Random Fields. The specificity of Markovian models is that the dependencies between the nodes are limited to the nearest neighbor nodes. The neighborhood definition can vary and be adapted to the problem of interest. When parts of the variables (nodes) are not observed or missing, we refer to these models as Hidden Markov Models (HMM). Hidden Markov chains or hidden Markov fields correspond to cases where the  $z_i$ 's in (1) are distributed according to a Markov chain or a Markov field. They are a natural extension of mixture models. They are widely used in signal processing (speech recognition, genome sequence analysis) and in image processing (remote sensing, MRI, etc.). Such models are very flexible in practice and can naturally account for the phenomena to be studied.

Hidden Markov models are very useful in modelling spatial dependencies but these dependencies and the possible existence of hidden variables are also responsible for a typically large amount of computation. It follows that the statistical analysis may not be straightforward. Typical issues are related to the neighborhood structure to be chosen when not dictated by the context and the possible high dimensionality of the observations. This also requires a good understanding of the role of each parameter and methods to tune them depending on the goal in mind. Regarding estimation algorithms, they correspond to an energy minimization problem which is NP-hard and usually performed through approximation. We focus on a certain type of methods based on the mean field principle and propose effective algorithms which show good performance in practice and for which we also study theoretical properties. We also propose some tools for model selection. Eventually we investigate ways to extend the standard Hidden Markov Field model to increase its modelling power.

### 3.3. Functional Inference, semi- and non-parametric methods

**Participants:** El-Hadji Deme, Jonathan El-Methni, Ludovic Leau-Mercier, Stéphane Girard, Gildas Mazo, Kai Qin, Huu Giao Nguyen, Farida Enikeeva, Seydou-Nourou Sylla.

dimension reduction, extreme value analysis, functional estimation.

We also consider methods which do not assume a parametric model. The approaches are non-parametric in the sense that they do not require the assumption of a prior model on the unknown quantities. This property is important since, for image applications for instance, it is very difficult to introduce sufficiently general parametric models because of the wide variety of image contents. Projection methods are then a way to decompose the unknown quantity on a set of functions (e.g. wavelets). Kernel methods which rely on smoothing the data using a set of kernels (usually probability distributions) are other examples. Relationships exist between these methods and learning techniques using Support Vector Machine (SVM) as this appears in the context of *level-sets estimation* (see section 3.3.2). Such non-parametric methods have become the cornerstone when dealing with functional data [59]. This is the case, for instance, when observations are curves. They enable us to model the data without a discretization step. More generally, these techniques are of great use for *dimension reduction* purposes (section 3.3.3). They enable reduction of the dimension of the functional or multivariate data without assumptions on the observations distribution. Semi-parametric methods refer to methods that include both parametric and non-parametric aspects. Examples include the Sliced Inverse Regression (SIR) method [68] which combines non-parametric regression techniques with parametric dimension reduction aspects. This is also the case in *extreme value analysis* [58], which is based on the modelling of distribution tails (see section 3.3.1). It differs from traditional statistics which focuses on the central part of distributions, i.e. on the most probable events. Extreme value theory shows that distribution tails can be modelled by both a functional part and a real parameter, the extreme value index.

#### 3.3.1. Modelling extremal events

Extreme value theory is a branch of statistics dealing with the extreme deviations from the bulk of probability distributions. More specifically, it focuses on the limiting distributions for the minimum or the maximum of a large collection of random observations from the same arbitrary distribution. Let  $X_{1,n} \leq \dots \leq X_{n,n}$  denote  $n$  ordered observations from a random variable  $X$  representing some quantity of interest. A  $p_n$ -quantile of  $X$  is the value  $x_{p_n}$  such that the probability that  $X$  is greater than  $x_{p_n}$  is  $p_n$ , i.e.  $P(X > x_{p_n}) = p_n$ . When  $p_n < 1/n$ , such a quantile is said to be extreme since it is usually greater than the maximum observation  $X_{n,n}$  (see Figure 1).

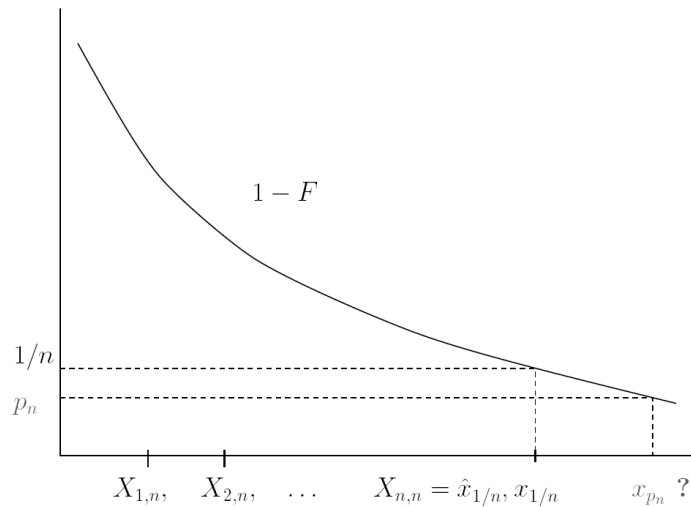


Figure 1. The curve represents the survival function  $x \rightarrow P(X > x)$ . The  $1/n$ -quantile is estimated by the maximum observation so that  $\hat{x}_{1/n} = X_{n,n}$ . As illustrated in the figure, to estimate  $p_n$ -quantiles with  $p_n < 1/n$ , it is necessary to extrapolate beyond the maximum observation.

To estimate such quantiles therefore requires dedicated methods to extrapolate information beyond the observed values of  $X$ . Those methods are based on Extreme value theory. This kind of issue appeared in hydrology. One objective was to assess risk for highly unusual events, such as 100-year floods, starting from flows measured over 50 years. To this end, semi-parametric models of the tail are considered:

$$P(X > x) = x^{-1/\theta} \ell(x), \quad x > x_0 > 0, \quad (52)$$

where both the extreme-value index  $\theta > 0$  and the function  $\ell(x)$  are unknown. The function  $\ell$  is a slowly varying function *i.e.* such that

$$\frac{\ell(tx)}{\ell(x)} \rightarrow 1 \quad \text{as } x \rightarrow \infty \quad (53)$$

for all  $t > 0$ . The function  $\ell(x)$  acts as a nuisance parameter which yields a bias in the classical extreme-value estimators developed so far. Such models are often referred to as heavy-tail models since the probability of extreme events decreases at a polynomial rate to zero. It may be necessary to refine the model (2,3) by specifying a precise rate of convergence in (3). To this end, a second order condition is introduced involving an additional parameter  $\rho \leq 0$ . The larger  $\rho$  is, the slower the convergence in (3) and the more difficult the estimation of extreme quantiles.

More generally, the problems that we address are part of the risk management theory. For instance, in reliability, the distributions of interest are included in a semi-parametric family whose tails are decreasing exponentially fast. These so-called Weibull-tail distributions [9] are defined by their survival distribution function:

$$P(X > x) = \exp \{-x^\theta \ell(x)\}, \quad x > x_0 > 0. \quad (54)$$

Gaussian, gamma, exponential and Weibull distributions, among others, are included in this family. An important part of our work consists in establishing links between models (2) and (4) in order to propose new estimation methods. We also consider the case where the observations were recorded with a covariate information. In this case, the extreme-value index and the  $p_n$ -quantile are functions of the covariate. We propose estimators of these functions by using moving window approaches, nearest neighbor methods, or kernel estimators.

### 3.3.2. Level sets estimation

Level sets estimation is a recurrent problem in statistics which is linked to outlier detection. In biology, one is interested in estimating reference curves, that is to say curves which bound 90% (for example) of the population. Points outside this bound are considered as outliers compared to the reference population. Level sets estimation can be looked at as a conditional quantile estimation problem which benefits from a non-parametric statistical framework. In particular, boundary estimation, arising in image segmentation as well as in supervised learning, is interpreted as an extreme level set estimation problem. Level sets estimation can also be formulated as a linear programming problem. In this context, estimates are sparse since they involve only a small fraction of the dataset, called the set of support vectors.

### 3.3.3. Dimension reduction

Our work on high dimensional data requires that we face the curse of dimensionality phenomenon. Indeed, the modelling of high dimensional data requires complex models and thus the estimation of high number of parameters compared to the sample size. In this framework, dimension reduction methods aim at replacing the original variables by a small number of linear combinations with as small as a possible loss of information. Principal Component Analysis (PCA) is the most widely used method to reduce dimension in data. However, standard linear PCA can be quite inefficient on image data where even simple image distortions can lead to highly non-linear data. Two directions are investigated. First, non-linear PCAs can be proposed, leading to semi-parametric dimension reduction methods [62]. Another field of investigation is to take into account the application goal in the dimension reduction step. One of our approaches is therefore to develop new Gaussian models of high dimensional data for parametric inference [53]. Such models can then be used in a Mixtures or Markov framework for classification purposes. Another approach consists in combining dimension reduction, regularization techniques, and regression techniques to improve the Sliced Inverse Regression method [68].

## **MODAL Project-Team**

### **3. Scientific Foundations**

#### **3.1. Generative model design**

The first objective of MODAL consists in designing, analyzing, estimating and evaluating new generative parametric models for multivariate and/or heterogeneous data. It corresponds typically to continuous and categorical data but it includes also other widespread ones like ordinal, functional, ranks,... Designed models have to take into account potential correlations between variables while being (1) justifiable and realistic, (2) meaningful and parsimoniously parameterized, (3) of low computational complexity. The main purpose is to identify a few theoretical and general principles for model generation, loosely dependent on the variable nature. In this context, we propose two concurrent approaches which could be general enough for dealing with correlation between many types of homogeneous or heterogeneous variables:

- Designs general models by combining two extreme models (full dependent and full independent) which are well-defined for most of variables;
- Uses kernels as a general way for dealing with multivariate and heterogeneous variables.

#### **3.2. Data visualization**

The second objective of MODAL is to propose meaningful and quite accurate low dimensional visualizations of data typically in two-dimensional (2D) space, less frequently in one-dimensional (1D) or three-dimensional (3D) spaces, by using the generative models designed in the first objective. We propose also to visualize simultaneously the data and the model. All visualizations will depend on the aim at hand (typically clustering, classification or density estimation). The main originality of this objective lies in the use of models for visualization, strategy from which we expect to have a better control on the subjectivity necessarily induced by any graphical display. In addition, the proposed approach has to be general enough to be independent on the variable nature. Note that the visualization objective is consistent with the dissemination of our methodologies through specific softwares. Indeed, displaying data is an important step in the data analysis process.

## REALOPT Project-Team

### 3. Scientific Foundations

#### 3.1. Introduction

*Combinatorial optimization* is the field of discrete optimization problems. In many applications, the most important decisions (control variables) are binary (on/off decisions) or integer (indivisible quantities). Extra variables can represent continuous adjustments or amounts. This results in models known as *mixed integer programs* (MIP), where the relationships between variables and input parameters are expressed as linear constraints and the goal is defined as a linear objective function. MIPs are notoriously difficult to solve: good quality estimations of the optimal value (bounds) are required to prune enumeration-based global-optimization algorithms whose complexity is exponential. In the standard approach to solving an MIP is so-called *branch-and-bound algorithm* : (i) one solves the linear programming (LP) relaxation using the simplex method; (ii) if the LP solution is not integer, one adds a disjunctive constraint on a fractional component (rounding it up or down) that defines two sub-problems; (iii) one applies this procedure recursively, thus defining a binary enumeration tree that can be pruned by comparing the local LP bound to the best known integer solution. Commercial MIP solvers are essentially based on branch-and-bound (such IBM Ilog-CPLEX or FICO/Dash-Optimization's Xpress-mp). They have made tremendous progress over the last decade (with a speedup by a factor of 60). But extending their capabilities remains a continuous challenge; given the combinatorial explosion inherent to enumerative solution techniques, they remain quickly overwhelmed beyond a certain problem size or complexity.

Progress can be expected from the development of tighter formulations. Central to our field is the characterization of polyhedra defining or approximating the solution set and combinatorial algorithms to identify "efficiently" a minimum cost solution or separate an unfeasible point. With properly chosen formulations, exact optimization tools can be competitive with other methods (such as meta-heuristics) in constructing good approximate solutions within limited computational time, and of course has the important advantage of being able to provide a performance guarantee through the relaxation bounds. Decomposition techniques are implicitly leading to better problem formulation as well, while constraint propagation are tools from artificial intelligence to further improve formulation through intensive preprocessing. A new trend is the study of non-linear models (non linearities are inherent in some engineering, economic and scientific applications) where solution techniques build on the best MIP approaches while demanding much more than simple extensions. Robust optimization is another area where recent progress have been made: the aim is to produce optimized solutions that remain of good quality even if the problem data has stochastic variations. In all cases, the study of specific models and challenging industrial applications is quite relevant because developments made into a specific context can become generic tools over time and see their way into commercial software.

Our project brings together researchers with expertise mathematical programming (polyhedral approaches, Dantzig-Wolfe decomposition, non-linear integer programming, stochastic programming, and dynamic programming), graph theory (characterization of graph properties, combinatorial algorithms) and constraint programming in the aim of producing better quality formulations and developing new methods to exploit these formulations. These new results are then applied to find high quality solutions for practical combinatorial problems such as routing, network design, planning, scheduling, cutting and packing problems.

#### 3.2. Polyhedral approaches for MIP

Adding valid inequalities to the polyhedral description of an MIP allows one to improve the resulting LP bound and hence to better prune the enumeration tree. In a cutting plane procedure, one attempt to identify valid inequalities that are violated by the LP solution of the current formulation and adds them to the formulation. This can be done at each node of the branch-and-bound tree giving rise to a so-called *branch-and-cut algorithm* [59]. The goal is to reduce the resolution of an integer program to that of a linear

program by deriving a linear description of the convex hull of the feasible solutions. Polyhedral theory tells us that if  $X$  is a mixed integer program:  $X = P \cap \mathbb{Z}^n \times \mathbb{R}^p$  where  $P = \{x \in \mathbb{R}^{n+p} : Ax \leq b\}$  with matrix  $(A, b) \in \mathbb{Q}^{m \times (n+p+1)}$ , then  $\text{conv}(X)$  is a polyhedron that can be described in terms of linear constraints, i.e. it writes as  $\text{conv}(X) = \{x \in \mathbb{R}^{n+p} : Cx \leq d\}$  for some matrix  $(C, d) \in \mathbb{Q}^{m' \times (n+p+1)}$  although the dimension  $m'$  is typically quite large. A fundamental result in this field is the equivalence of complexity between solving the combinatorial optimization problem  $\min\{cx : x \in X\}$  and solving the *separation problem* over the associated polyhedron  $\text{conv}(X)$ : if  $\tilde{x} \notin \text{conv}(X)$ , find a linear inequality  $\pi x \geq \pi_0$  satisfied by all points in  $\text{conv}(X)$  but violated by  $\tilde{x}$ . Hence, for NP-hard problems, one can not hope to get a compact description of  $\text{conv}(X)$  nor a polynomial time exact separation routine. Polyhedral studies focus on identifying some of the inequalities that are involved in the polyhedral description of  $\text{conv}(X)$  and derive efficient *separation procedures* (cutting plane generation). Only a subset of the inequalities  $Cx \leq d$  can offer a good approximation, that combined with a branch-and-bound enumeration techniques permits to solve the problem. Using *cutting plane algorithm* at each node of the branch-and-bound tree, gives rise to the algorithm called *branch-and-cut*.

### 3.3. Decomposition and reformulation approaches

An hierarchical approach to tackle complex combinatorial problems consists in considering separately different substructures (subproblems). If one is able to implement relatively efficient optimization on the substructures, this can be exploited to reformulate the global problem as a selection of specific subproblem solutions that together form a global solution. If the subproblems correspond to subset of constraints in the MIP formulation, this leads to Dantzig-Wolfe decomposition. If it corresponds to isolating a subset of decision variables, this leads to Bender's decomposition. Both lead to extended formulations of the problem with either a huge number of variables or constraints. Dantzig-Wolfe approach requires specific algorithmic approaches to generate subproblem solutions and associated global decision variables dynamically in the course of the optimization. This procedure is known as *column generation*, while its combination with branch-and-bound enumeration is called, *branch-and-price*. Alternatively, in Bender's approach, when dealing with exponentially many constraints in the reformulation, *cutting plane procedures* defined in the previous section reveal to be powerful. When optimization on a substructure is (relatively) easy, there often exists a tight reformulation of this substructure typically in an extended variable space. This gives rise powerful reformulation of the global problem, although it might be impractical given its size (typically pseudo-polynomial). It can be possible to project (part of) the extended formulation in a smaller dimensional space if not the original variable space to bring polyhedral insight (cuts derived through polyhedral studies can often be recovered through such projections).

### 3.4. Constraint Programming (CP)

Constraint Programming focuses on iteratively reducing the variable domains (sets of feasible values) by applying logical and problem-specific operators. The latter propagates on selected variables the restrictions that are implied by the other variable domains through the relations between variables that are defined by the constraints of the problem. Combined with enumeration, it gives rise to exact optimization algorithms. A CP approach is particularly effective for tightly constrained problems, feasibility problems and min-max problems (minimizing the maximum of several variable values). Mixed Integer Programming (MIP), on the other hand, is effective for loosely constrained problems and for problems with an objective function defined as the weighted sum of variables. Many problems belong to the intersection of these two classes. For example, some scheduling and timetabling problems are tightly constrained and have a sum-type objective. For such problems, it is reasonable to use algorithms that exploit complementary strengths of Constraint Programming and Mixed Integer Programming.

### 3.5. Mixed Integer NonLinear Programming (MINLP)

Many engineering, management, and scientific applications involve not only discrete decisions, but also nonlinear relationships that significantly affect the feasibility and optimality of solutions. MINLP problems

combine the difficulties of MIP with the challenges of handling nonlinear functions. MINLP is one of the most flexible modeling paradigms available. However, solving such models is much more challenging: available softwares are not nearly as effective as standard softwares for linear MIP. The most powerful algorithms combine sophisticated methods that maintain outer linear programming approximation or convex relaxations with branch-and-bound enumeration; hence, the role of strong convex reformulations is crucial. The development of results for structured sub-models are essential building blocks. Preprocessing and bound reduction (domain reduction logic similar to that used in CP) are quite important too. Finally, decomposition methods also permit to develop tight outer approximations.

### **3.6. Polyhedral Combinatorics and Graph Theory**

Many fundamental combinatorial optimization problems can be modeled as the search for a specific structure in a graph. For example, ensuring connectivity in a network amounts to building a *tree* that spans all the nodes. Inquiring about its resistance to failure amounts to searching for a minimum cardinality *cut* that partitions the graph. Selecting disjoint pairs of objects is represented by a so-called *matching*. Disjunctive choices can be modeled by edges in a so-called *conflict graph* where one searches for *stable sets* – a set of nodes that are not incident to one another. Polyhedral combinatorics is the study of combinatorial algorithms involving polyhedral considerations. Not only it leads to efficient algorithms, but also, conversely, efficient algorithms often imply polyhedral characterizations and related min-max relations. Developments of polyhedral properties of a fundamental problem will typically provide us with more interesting inequalities well suited for a branch-and-cut algorithm to more general problems. Furthermore, one can use the fundamental problems as new building bricks to decompose the more general problem at hand. For problem that let themselves easily be formulated in a graph setting, the graph theory and in particular graph decomposition theorem might help.

## SELECT Project-Team

### 3. Scientific Foundations

#### 3.1. General presentation

We learned from the applications we treated that some assumptions which are currently used in asymptotic theory for model selection are often irrelevant in practice. For instance, it is not realistic to assume that the target belongs to the family of models in competition. Moreover, in many situations, it is useful to make the size of the model depend on the sample size which make the asymptotic analysis breakdown. An important aim of SELECT is to propose model selection criteria which take these practical constraints into account.

#### 3.2. A non asymptotic view for model selection

An important purpose of SELECT is to build and analyze penalized log-likelihood model selection criteria that are efficient when the number of models in competition grows to infinity with the number of observations. Concentration inequalities are a key tool for that purpose and lead to data-driven penalty choice strategies. A major issue of SELECT consists of deepening the analysis of data-driven penalties both from the theoretical and the practical side. There is no universal way of calibrating penalties but there are several different general ideas that we want to develop, including heuristics derived from the Gaussian theory, special strategies for variable selection and using resampling methods.

#### 3.3. Taking into account the modeling purpose in model selection

Choosing a model is not only difficult theoretically. From a practical point of view, it is important to design model selection criteria that accommodate situations in which the data probability distribution  $P$  is unknown and which take the model user's purpose into account. Most standard model selection criteria assume that  $P$  belongs to one of a set of models, without considering the purpose of the model. By also considering the model user's purpose, we avoid or overcome certain theoretical difficulties and can produce flexible model selection criteria with data-driven penalties. The latter is useful in supervised Classification and hidden-structure models.

#### 3.4. Bayesian model selection

The Bayesian approach to statistical problems is fundamentally probabilistic. A joint probability distribution is used to describe the relationships among all the unknowns and the data. Inference is then based on the posterior distribution i.e. the conditional probability distribution of the parameters given the observed data. Exploiting the internal consistency of the probability framework, the posterior distribution extracts the relevant information in the data and provides a complete and coherent summary of post-data uncertainty. Using the posterior to solve specific inference and decision problems is then straightforward, at least in principle.



## SEQUEL Project-Team

### 3. Scientific Foundations

#### 3.1. Introduction

SEQUEL is primarily grounded on two domains:

- the problem of decision under uncertainty,
- statistical analysis and statistical learning, which provide the general concepts and tools to solve this problem.

To help the reader who is unfamiliar with these questions, we briefly present key ideas below.

#### 3.2. Decision-making Under Uncertainty

The phrase “Decision under uncertainty” refers to the problem of taking decisions when we do not have a full knowledge neither of the situation, nor of the consequences of the decisions, as well as when the consequences of decision are non deterministic.

We introduce two specific sub-domains, namely the Markov decision processes which models sequential decision problems, and bandit problems.

##### 3.2.1. Reinforcement Learning

Sequential decision processes occupy the heart of the SEQUEL project; a detailed presentation of this problem may be found in Puterman’s book [70].

A Markov Decision Process (MDP) is defined as the tuple  $(\mathcal{X}, \mathcal{A}, P, r)$  where  $\mathcal{X}$  is the state space,  $\mathcal{A}$  is the action space,  $P$  is the probabilistic transition kernel, and  $r : \mathcal{X} \times \mathcal{A} \times \mathcal{X} \rightarrow \mathbb{R}$  is the reward function. For the sake of simplicity, we assume in this introduction that the state and action spaces are finite. If the current state (at time  $t$ ) is  $x \in \mathcal{X}$  and the chosen action is  $a \in \mathcal{A}$ , then the Markov assumption means that the transition probability to a new state  $x' \in \mathcal{X}$  (at time  $t + 1$ ) only depends on  $(x, a)$ . We write  $p(x'|x, a)$  the corresponding transition probability. During a transition  $(x, a) \rightarrow x'$ , a reward  $r(x, a, x')$  is incurred.

In the MDP  $(\mathcal{X}, \mathcal{A}, P, r)$ , each initial state  $x_0$  and action sequence  $a_0, a_1, \dots$  gives rise to a sequence of states  $x_1, x_2, \dots$ , satisfying  $\mathbb{P}(x_{t+1} = x' | x_t = x, a_t = a) = p(x'|x, a)$ , and rewards  $r_1, r_2, \dots$  defined by  $r_t = r(x_t, a_t, x_{t+1})$ .

The history of the process up to time  $t$  is defined to be  $H_t = (x_0, a_0, \dots, x_{t-1}, a_{t-1}, x_t)$ . A policy  $\pi$  is a sequence of functions  $\pi_0, \pi_1, \dots$ , where  $\pi_t$  maps the space of possible histories at time  $t$  to the space of probability distributions over the space of actions  $\mathcal{A}$ . To follow a policy means that, in each time step, we assume that the process history up to time  $t$  is  $x_0, a_0, \dots, x_t$  and the probability of selecting an action  $a$  is equal to  $\pi_t(x_0, a_0, \dots, x_t)(a)$ . A policy is called stationary (or Markovian) if  $\pi_t$  depends only on the last visited state. In other words, a policy  $\pi = (\pi_0, \pi_1, \dots)$  is called stationary if  $\pi_t(x_0, a_0, \dots, x_t) = \pi_0(x_t)$  holds for all  $t \geq 0$ . A policy is called deterministic if the probability distribution prescribed by the policy for any history is concentrated on a single action. Otherwise it is called a stochastic policy.

We move from an MD process to an MD problem by formulating the goal of the agent, that is what the sought policy  $\pi$  has to optimize? It is very often formulated as maximizing (or minimizing), in expectation, some functional of the sequence of future rewards. For example, an usual functional is the infinite-time horizon sum of discounted rewards. For a given (stationary) policy  $\pi$ , we define the value function  $V^\pi(x)$  of that policy  $\pi$  at a state  $x \in \mathcal{X}$  as the expected sum of discounted future rewards given that we state from the initial state  $x$  and follow the policy  $\pi$ :

---

<sup>1</sup>Note that for simplicity, we considered the case of a deterministic reward function, but in many applications, the reward  $r_t$  itself is a random variable.

$$V^\pi(x) = \mathbb{E} \left[ \sum_{t=0}^{\infty} \gamma^t r_t | x_0 = x, \pi \right], \quad (55)$$

where  $\mathbb{E}$  is the expectation operator and  $\gamma \in (0, 1)$  is the discount factor. This value function  $V^\pi$  gives an evaluation of the performance of a given policy  $\pi$ . Other functionals of the sequence of future rewards may be considered, such as the undiscounted reward (see the stochastic shortest path problems [66]) and average reward settings. Note also that, here, we considered the problem of maximizing a reward functional, but a formulation in terms of minimizing some cost or risk functional would be equivalent.

In order to maximize a given functional in a sequential framework, one usually applies Dynamic Programming (DP) [64], which introduces the optimal value function  $V^*(x)$ , defined as the optimal expected sum of rewards when the agent starts from a state  $x$ . We have  $V^*(x) = \sup_{\pi} V^\pi(x)$ . Now, let us give two definitions about policies:

- We say that a policy  $\pi$  is optimal, if it attains the optimal values  $V^*(x)$  for any state  $x \in \mathcal{X}$ , i.e., if  $V^\pi(x) = V^*(x)$  for all  $x \in \mathcal{X}$ . Under mild conditions, deterministic stationary optimal policies exist [65]. Such an optimal policy is written  $\pi^*$ .
- We say that a (deterministic stationary) policy  $\pi$  is greedy with respect to (w.r.t.) some function  $V$  (defined on  $\mathcal{X}$ ) if, for all  $x \in \mathcal{X}$ ,

$$\pi(x) \in \arg \max_{a \in \mathcal{A}} \sum_{x' \in \mathcal{X}} p(x'|x, a) [r(x, a, x') + \gamma V(x')].$$

where  $\arg \max_{a \in \mathcal{A}} f(a)$  is the set of  $a \in \mathcal{A}$  that maximizes  $f(a)$ . For any function  $V$ , such a greedy policy always exists because  $\mathcal{A}$  is finite.

The goal of Reinforcement Learning (RL), as well as that of dynamic programming, is to design an optimal policy (or a good approximation of it).

The well-known Dynamic Programming equation (also called the Bellman equation) provides a relation between the optimal value function at a state  $x$  and the optimal value function at the successors states  $x'$  when choosing an optimal action: for all  $x \in \mathcal{X}$ ,

$$V^*(x) = \max_{a \in \mathcal{A}} \sum_{x' \in \mathcal{X}} p(x'|x, a) [r(x, a, x') + \gamma V^*(x')]. \quad (56)$$

The benefit of introducing this concept of optimal value function relies on the property that, from the optimal value function  $V^*$ , it is easy to derive an optimal behavior by choosing the actions according to a policy greedy w.r.t.  $V^*$ . Indeed, we have the property that a policy greedy w.r.t. the optimal value function is an optimal policy:

$$\pi^*(x) \in \arg \max_{a \in \mathcal{A}} \sum_{x' \in \mathcal{X}} p(x'|x, a) [r(x, a, x') + \gamma V^*(x')]. \quad (57)$$

In short, we would like to mention that most of the reinforcement learning methods developed so far are built on one (or both) of the two following approaches ([76]):

- Bellman's dynamic programming approach, based on the introduction of the value function. It consists in learning a "good" approximation of the optimal value function, and then using it to derive a greedy policy w.r.t. this approximation. The hope (well justified in several cases) is that the performance  $V^\pi$  of the policy  $\pi$  greedy w.r.t. an approximation  $V$  of  $V^*$  will be close to optimality. This approximation issue of the optimal value function is one of the major challenge inherent to the reinforcement learning problem. **Approximate dynamic programming** addresses

the problem of estimating performance bounds (e.g. the loss in performance  $\|V^* - V^\pi\|$  resulting from using a policy  $\pi$ -greedy w.r.t. some approximation  $V$ - instead of an optimal policy) in terms of the approximation error  $\|V^* - V\|$  of the optimal value function  $V^*$  by  $V$ . Approximation theory and Statistical Learning theory provide us with bounds in terms of the number of sample data used to represent the functions, and the capacity and approximation power of the considered function spaces.

- Pontryagin’s maximum principle approach, based on sensitivity analysis of the performance measure w.r.t. some control parameters. This approach, also called **direct policy search** in the Reinforcement Learning community aims at directly finding a good feedback control law in a parameterized policy space without trying to approximate the value function. The method consists in estimating the so-called **policy gradient**, i.e. the sensitivity of the performance measure (the value function) w.r.t. some parameters of the current policy. The idea being that an optimal control problem is replaced by a parametric optimization problem in the space of parameterized policies. As such, deriving a policy gradient estimate would lead to performing a stochastic gradient method in order to search for a local optimal parametric policy.

Finally, many extensions of the Markov decision processes exist, among which the Partially Observable MDPs (POMDPs) is the case where the current state does not contain all the necessary information required to decide for sure of the best action.

### 3.2.2. Multi-arm Bandit Theory

Bandit problems illustrate the fundamental difficulty of decision making in the face of uncertainty: A decision maker must choose between what seems to be the best choice (“exploit”), or to test (“explore”) some alternative, hoping to discover a choice that beats the current best choice.

The classical example of a bandit problem is deciding what treatment to give each patient in a clinical trial when the effectiveness of the treatments are initially unknown and the patients arrive sequentially. These bandit problems became popular with the seminal paper [71], after which they have found applications in diverse fields, such as control, economics, statistics, or learning theory.

Formally, a  $K$ -armed bandit problem ( $K \geq 2$ ) is specified by  $K$  real-valued distributions. In each time step a decision maker can select one of the distributions to obtain a sample from it. The samples obtained are considered as rewards. The distributions are initially unknown to the decision maker, whose goal is to maximize the sum of the rewards received, or equivalently, to minimize the regret which is defined as the loss compared to the total payoff that can be achieved given full knowledge of the problem, i.e., when the arm giving the highest expected reward is pulled all the time.

The name “bandit” comes from imagining a gambler playing with  $K$  slot machines. The gambler can pull the arm of any of the machines, which produces a random payoff as a result: When arm  $k$  is pulled, the random payoff is drawn from the distribution associated to  $k$ . Since the payoff distributions are initially unknown, the gambler must use exploratory actions to learn the utility of the individual arms. However, exploration has to be carefully controlled since excessive exploration may lead to unnecessary losses. Hence, to play well, the gambler must carefully balance exploration and exploitation. Auer *et al.* [63] introduced the algorithm UCB (Upper Confidence Bounds) that follows what is now called the “optimism in the face of uncertainty principle”. Their algorithm works by computing upper confidence bounds for all the arms and then choosing the arm with the highest such bound. They proved that the expected regret of their algorithm increases at most at a logarithmic rate with the number of trials, and that the algorithm achieves the smallest possible regret up to some sub-logarithmic factor (for the considered family of distributions).

## 3.3. Statistical analysis of time series

Many of the problems of machine learning can be seen as extensions of classical problems of mathematical statistics to their (extremely) non-parametric and model-free cases. Other machine learning problems are founded on such statistical problems. Statistical problems of sequential learning are mainly those that are concerned with the analysis of time series. These problems are as follows.

### 3.3.1. Prediction of Sequences of Structured and Unstructured Data

Given a series of observations  $x_1, \dots, x_n$  it is required to give forecasts concerning the distribution of the distribution of the future observations  $x_{n+1}, x_{n+2}, \dots$ ; in the simplest case, that of the next outcome  $x_{n+1}$ . Then  $x_{n+1}$  is revealed and the process continues. Different goals can be formulated in this setting. One can either make some assumptions on the probability measure that generates the sequence  $x_1, \dots, x_n, \dots$ , such as that the outcomes are independent and identically distributed (i.i.d.), or that the sequence is a Markov chain, that it is a stationary process, etc. More generally, one can assume that the data is generated by a probability measure that belongs to a certain set  $\mathcal{C}$ . In these cases the goal is to have the discrepancy between the predicted and the “true” probabilities to go to zero, if possible, with guarantees on the speed of convergence.

Alternatively, rather than making some assumptions on the data, one can change the goal: the predicted probabilities should be asymptotically as good as those given by the best reference predictor from a certain pre-defined set.

Another dimension of complexity in this problem concerns the nature of observations  $x_i$ . In the simplest case, they come from a finite space, but already basic applications often require real-valued observations. Moreover, function or even graph-valued observations often arise in practice, in particular in applications concerning Web data. In these settings estimating even simple characteristics of probability distributions of the future outcomes becomes non-trivial, and new learning algorithms for solving these problems are in order.

### 3.3.2. Hypothesis testing

Given a series of observations of  $x_1, \dots, x_n, \dots$  generated by some unknown probability measure  $\mu$ , the problem is to test a certain given hypothesis  $H_0$  about  $\mu$ , versus a given alternative hypothesis  $H_1$ . There are many different examples of this problem. Perhaps the simplest one is testing a simple hypothesis “ $\mu$  is Bernoulli i.i.d. measure with probability of 0 equals  $1/2$ ” versus “ $\mu$  is Bernoulli i.i.d. with the parameter different from  $1/2$ ”. More interesting cases include the problems of model verification: for example, testing that  $\mu$  is a Markov chain, versus that it is a stationary ergodic process but not a Markov chain. In the case when we have not one but several series of observations, we may wish to test the hypothesis that they are independent, or that they are generated by the same distribution. Applications of these problems to a more general class of machine learning tasks include the problem of feature selection, the problem of testing that a certain behaviour (such as pulling a certain arm of a bandit, or using a certain policy) is better (in terms of achieving some goal, or collecting some rewards) than another behaviour, or than a class of other behaviours.

The problem of hypothesis testing can also be studied in its general formulations: given two (abstract) hypothesis  $H_0$  and  $H_1$  about the unknown measure that generates the data, find out whether it is possible to test  $H_0$  against  $H_1$  (with confidence), and if yes then how can one do it.

### 3.3.3. Change Point Analysis

A stochastic process is generating the data. At some point, the process distribution changes. In the “offline” situation, the statistician observes the resulting sequence of outcomes and has to estimate the point or the points at which the change(s) occurred. In online setting, the goal is to detect the change as quickly as possible.

These are the classical problems in mathematical statistics, and probably among the last remaining statistical problems not adequately addressed by machine learning methods. The reason for the latter is perhaps in that the problem is rather challenging. Thus, most methods available so far are parametric methods concerning piecewise constant distributions, and the change in distribution is associated with the change in the mean. However, many applications, including DNA analysis, the analysis of (user) behaviour data, etc., fail to comply with this kind of assumptions. Thus, our goal here is to provide completely non-parametric methods allowing for any kind of changes in the time-series distribution.

### 3.3.4. Clustering Time Series, Online and Offline

The problem of clustering, while being a classical problem of mathematical statistics, belongs to the realm of unsupervised learning. For time series, this problem can be formulated as follows: given several samples  $x^1 = (x_{n_1}^1, \dots, x_{n_1}^1), \dots, x^N = (x_{n_N}^1, \dots, x_{n_N}^N)$ , we wish to group similar objects together. While this is of

course not a precise formulation, it can be made precise if we assume that the samples were generated by  $k$  different distributions.

The online version of the problem allows for the number of observed time series to grow with time, in general, in an arbitrary manner.

### 3.3.5. Online Semi-Supervised Learning

Semi-supervised learning (SSL) is a field of machine learning that studies learning from both labeled and unlabeled examples. This learning paradigm is extremely useful for solving real-world problems, where data is often abundant but the resources to label them are limited.

Furthermore, *online* SSL is suitable for adaptive machine learning systems. In the classification case, learning is viewed as a repeated game against a potentially adversarial nature. At each step  $t$  of this game, we observe an example  $\mathbf{x}_t$ , and then predict its label  $\hat{y}_t$ .

The challenge of the game is that we only exceptionally observe the true label  $y_t$ . In the extreme case, which we also study, only a handful of labeled examples are provided in advance and set the initial bias of the system while unlabeled examples are gathered online and update the bias continuously. Thus, if we want to adapt to changes in the environment, we have to rely on indirect forms of feedback, such as the structure of data.

## 3.4. Statistical Learning and Bayesian Analysis

Before detailing some issues in these fields, let us remind the definition of a few terms.

**Machine learning** refers to a system capable of the autonomous acquisition and integration of knowledge. This capacity to learn from experience, analytical observation, and other means, results in a system that can continuously self-improve and thereby offer increased efficiency and effectiveness.

**Statistical learning** is an approach to machine intelligence that is based on statistical modeling of data. With a statistical model in hand, one applies probability theory and decision theory to get an algorithm. This is opposed to using training data merely to select among different algorithms or using heuristics/“common sense” to design an algorithm.

**Bayesian Analysis** applies to data that could be seen as observations in the more general meaning of the term. These data may not only come from classical sensors but also from any *device* recording information. From an operational point of view, like for statistical learning, uncertainty about the data is modeled by a probability measure thus defining the so-called likelihood functions. This last one depend upon parameters defining the state of the world we focus on for decision purposes. Within the Bayesian framework the uncertainty about these parameters is also modeled by probability measures, the priors that are subjective probabilities. Using probability theory and decision theory, one then defines new algorithms to estimate the parameters of interest and/or associated decisions. According to the International Society for Bayesian Analysis (source: <http://bayesian.org>), and from a more general point of view, this overall process could be summarize as follows: one assesses the current state of knowledge regarding the issue of interest, gather new data to address remaining questions, and then update and refine their understanding to incorporate both new and old data. Bayesian inference provides a logical, quantitative framework for this process based on probability theory.

**Kernel method.** Generally speaking, a kernel function is a function that maps a couple of points to a real value. Typically, this value is a measure of dissimilarity between the two points. Assuming a few properties on it, the kernel function implicitly defines a dot product in some function space. This very nice formal property as well as a bunch of others have ensured a strong appeal for these methods in the last 10 years in the field of function approximation. Many classical algorithms have been “kernelized”, that is, restated in a much more general way than their original formulation. Kernels also implicitly induce the representation of data in a certain “suitable” space where the problem to solve (classification, regression, ...) is expected to be simpler (non-linearity turns to linearity).

The fundamental tools used in SEQUEL come from the field of statistical learning [68]. We briefly present the most important for us to date, namely, kernel-based non parametric function approximation, and non parametric Bayesian models.

### 3.4.1. Non-parametric methods for Function Approximation

In statistics in general, and applied mathematics, the approximation of a multi-dimensional real function given some samples is a well-known problem (known as either regression, or interpolation, or function approximation, ...). Regressing a function from data is a key ingredient of our research, or to the least, a basic component of most of our algorithms. In the context of sequential learning, we have to regress a function while data samples are being obtained one at a time, while keeping the constraint to be able to predict points at any step along the acquisition process. In sequential decision problems, we typically have to learn a value function, or a policy.

Many methods have been proposed for this purpose. We are looking for suitable ones to cope with the problems we wish to solve. In reinforcement learning, the value function may have areas where the gradient is large; these are areas where the approximation is difficult, while these are also the areas where the accuracy of the approximation should be maximal to obtain a good policy (and where, otherwise, a bad choice of action may imply catastrophic consequences).

We particularly favor non parametric methods since they make quite a few assumptions about the function to learn. In particular, we have strong interests in  $l_1$ -regularization, and the (kernelized-)LARS algorithm.  $l_1$ -regularization yields sparse solutions, and the LARS approach produces the whole regularization path very efficiently, which helps solving the regularization parameter tuning problem.

### 3.4.2. Nonparametric Bayesian Estimation

Numerous problems may be solved efficiently by a Bayesian approach. The use of Monte-Carlo methods allows us to handle non-linear, as well as non-Gaussian, problems. In their standard form, they require the formulation of probability densities in a parametric form. For instance, it is a common usage to use Gaussian likelihood, because it is handy. However, in some applications such as Bayesian filtering, or blind deconvolution, the choice of a parametric form of the density of the noise is often arbitrary. If this choice is wrong, it may also have dramatic consequences on the estimation quality. To overcome this shortcoming, one possible approach is to consider that this density must also be estimated from data. A general Bayesian approach then consists in defining a probabilistic space associated with the possible outcomes of the *object* to be estimated. Applied to density estimation, it means that we need to define a probability measure on the probability density of the noise : such a measure is called a *random measure*. The classical Bayesian inference procedures can then been used. This approach being by nature non parametric, the associated frame is called *Non Parametric Bayesian*.

In particular, mixtures of Dirichlet processes [67] provide a very powerful formalism. Dirichlet Processes are a possible random measure and Mixtures of Dirichlet Processes are an extension of well-known finite mixture models. Given a mixture density  $f(x|\theta)$ , and  $G(d\theta) = \sum_{k=1}^{\infty} \omega_k \delta_{U_k}(d\theta)$ , a Dirichlet process, we define a mixture of Dirichlet processes as:

$$F(x) = \int_{\Theta} f(x|\theta)G(d\theta) = \sum_{k=1}^{\infty} \omega_k f(x|U_k) \quad (58)$$

where  $F(x)$  is the density to be estimated. The class of densities that may be written as a mixture of Dirichlet processes is very wide, so that they really fit a very large number of applications.

Given a set of observations, the estimation of the parameters of a mixture of Dirichlet processes is performed by way of a Monte Carlo Markov Chain (MCMC) algorithm. Dirichlet Process Mixture are also widely used in clustering problems. Once the parameters of a mixture are estimated, they can be interpreted as the parameters of a specific cluster defining a class as well. Dirichlet processes are well known within the machine learning community and their potential in statistical signal processing still need to be developed.

### 3.4.3. Random Finite Sets for multisensor multitarget tracking

In the general multi-sensor multi-target Bayesian framework, an unknown (and possibly varying) number of targets whose states  $x_1, \dots, x_n$  are observed by several sensors which produce a collection of measurements  $z_1, \dots, z_m$  at every time step  $k$ . Well-known models to this problem are track-based models, such as the joint probability data association (JPDA), or joint multi-target probabilities, such as the joint multi-target probability density. Common difficulties in multi-target tracking arise from the fact that the system state and the collection of measures from sensors are unordered and their size evolve randomly through time. Vector-based algorithms must therefore account for state coordinates exchanges and missing data within an unknown time interval. Although this approach is very popular and has resulted in many algorithms in the past, it may not be the optimal way to tackle the problem, since the state and the data are in fact *sets* and not vectors.

The random finite set theory provides a powerful framework to deal with these issues. Mahler's work on finite sets statistics (FISST) provides a mathematical framework to build multi-object densities and derive the Bayesian rules for state prediction and state estimation. Randomness on object number and their states are encapsulated into random finite sets (RFS), namely multi-target(state) sets  $X = \{x_1, \dots, x_n\}$  and multi-sensor (measurement) set  $Zk = \{z_1, \dots, z_m\}$ . The objective is then to propagate the multitarget probability density  $f_{k|k}(X|Z(k))$  by using the Bayesian set equations at every time step  $k$ :

$$\begin{aligned} f_{k+1|k}(X|Z^{(k)}) &= \int f_{k+1|k}(X|W)f_{k|k}(W|Z^{(k)})\delta W \\ f_{k+1|k+1}(X|Z^{(k+1)}) &= \frac{f_{k+1}(Z_{k+1}|X)f_{k+1|k}(X|Z^{(k)})}{\int f_{k+1}(Z_{k+1}|W)f_{k+1|k}(W|Z^{(k)})\delta W} \end{aligned} \quad (59)$$

where:

- $X = \{x_1, \dots, x_n\}$  is a multi-target state, i.e. a finite set of elements  $x_i$  defined on the single-target space  $\mathcal{X}$ ; <sup>2</sup>
- $Z_{k+1} = \{z_1, \dots, z_m\}$  is the current multi-sensor observation, i.e. a collection of measures  $z_i$  produced at time  $k + 1$  by all the sensors;
- $Z^{(k)} = \bigcup_{t \leq k} Z_t$  is the collection of observations up to time  $k$ ;
- $f_{k|k}(W|Z^{(k)})$  is the current multi-target posterior density in state  $W$ ;
- $f_{k+1|k}(X|W)$  is the current multi-target Markov transition density, from state  $W$  to state  $X$ ;
- $f_{k+1}(Z|X)$  is the current multi-sensor/multi-target likelihood function.

Although equations (5) may seem similar to the classical single-sensor/single-target Bayesian equations, they are generally intractable because of the presence of the *set integrals*. For, a RFS  $\Xi$  is characterized by the family of its Janossy densities  $j_{\Xi,1}(x_1), j_{\Xi,2}(x_1, x_2) \dots$  and not just by one density as it is the case with vectors. Mahler then introduced the PHD, defined on single-target state space. The PHD is the quantity whose integral on any region  $S$  is the expected number of targets inside  $S$ . Mahler proved that the PHD is the first-moment density of the multi-target probability density. Although defined on single-state space  $X$ , the PHD encapsulates information on both target number and states.

<sup>2</sup>The state  $x_i$  of a target is usually composed of its position, its velocity, etc.

## **SIERRA Project-Team**

### **3. Scientific Foundations**

#### **3.1. Supervised Learning**

This part of our research focuses on methods where, given a set of examples of input/output pairs, the goal is to predict the output for a new input, with research on kernel methods, calibration methods, and multi-task learning.

#### **3.2. Unsupervised Learning**

We focus here on methods where no output is given and the goal is to find structure of certain known types (e.g., discrete or low-dimensional) in the data, with a focus on matrix factorization, statistical tests, dimension reduction, and semi-supervised learning.

#### **3.3. Parsimony**

The concept of parsimony is central to many areas of science. In the context of statistical machine learning, this takes the form of variable or feature selection. The team focuses primarily on structured sparsity, with theoretical and algorithmic contributions (this is the main topic of the ERC starting investigator grant awarded to F. Bach).

#### **3.4. Optimization**

Optimization in all its forms is central to machine learning, as many of its theoretical frameworks are based at least in part on empirical risk minimization. The team focuses primarily on convex and bandit optimization, with a particular focus on large-scale optimization.



## TAO Project-Team

### 3. Scientific Foundations

#### 3.1. Introduction

This section describes TAO main research directions at the crossroad of Machine Learning and Evolutionary Computation. Since 2008, TAO has been structured in several special interest groups (SIGs) to enable the agile investigation of long-term or emerging theoretical or applicative issues. The comparatively small size of TAO SIGs enables in-depth and lively discussions; the fact that all TAO members belong to several SIGs, on the basis of their personal interests, enforces the strong and informal collaboration of the groups, and the fast information dissemination.

The first two SIGs consolidate the key TAO scientific pillars, while others evolve and adapt to new topics.

The first one, OPT-SIG, addresses stochastic continuous optimization, taking advantage of the fact that TAO is acknowledged the best French research group and one of the top international groups in evolutionary computation from a theoretical and algorithmic standpoint. A main priority on the OPT-SIG research agenda is to provide theoretical and algorithmic guarantees for the current world best continuous stochastic optimizer, CMA-ES, ranging from convergence analysis (Youhei Akimoto's and Verena Heidrich-Meister's post-docs) to a rigorous benchmarking methodology. Incidentally, this benchmark platform has been acknowledged since 2008 as "the" international continuous optimization benchmark, and its extension is at the core of the ANR project NumBBO (starting end 2012). Another priority is to address the current limitations of CMA-ES in terms of high-dimensional or expensive optimization (respectively Ouassim Ait El Hara's and Ilya Loshchilov's PhDs). Mouadh Yagoubi and Zyed Bouzarkouna (resp. CIFRE PSA and IFP-EN) PhD's have continued the EC tradition of industrial breakthrough applications, on which the EC fame solidly relies.

The second SIG, UCT-SIG, benefits from the MoGo expertise and its past and present world records in the domain of computer-Go, establishing the international visibility of TAO in sequential decision making. Since 2010, UCT-SIG resolutely moves to address the problems of energy management from a fundamental and applied perspective. On the one hand, energy management offers a host of challenging issues, ranging from long-horizon policy optimization to the combinatorial nature of the search space, from the modeling of prior knowledge to non-stationary environment to name a few. On the other hand, the energy management issue can hardly be tackled in a pure academic perspective: tight collaborations with industrial partners are needed to access the true operational constraints. Such international and national collaborations have been started by Olivier Teytaud during his one-year stay in Taiwan, and witnessed by the FP7 STREP Citines, the pending ADEME Post contract, and the Ilab with Artelys.

A third SIG, DIS-SIG, is devoted to the modeling and optimization of (large scale) distributed systems. DIS-SIG pursues and extends the goals of the former *Autonomic Computing* SIG, initiated by Cécile Germain and investigating the use of statistical machine learning for large scale computational architectures, from data acquisition (the Grid Observatory in the European Grid Initiative) to grid management and fault detection. More generally, how to model and manage network-based activities has been acknowledged a key topic *per se* in the last months, including: i) the modeling of multi-agent systems and the exploitation of simulation results in the SimTools RNSC network frame; ii) the management of the core communication topology for distributed SAT solving, in the *Microsoft-TAO project* framework. Further extensions are planned in the context of the TIMCO FUI project (started end 2012); the challenge is not only to port ML algorithms on massively distributed architectures, but to see how these architectures can inspire new ML criteria and methodologies.

A fourth SIG, CRI-SIG, focuses on the design of learning and optimization criteria. It elaborates on the lessons learned from the former *Complex Systems* SIG, showing that the key issue in challenging applications is to design the objective itself. Such targeted criteria are pervasive in the study and building of autonomous cognitive systems, ranging from intrinsic rewards in robotics to the notion of saliency in vision and image understanding. The desired criteria can also result from fundamental requirements, such as scale invariance in a statistical physics perspective, and guide the algorithmic design. Additionally, the criteria can also be domain-driven and reflect the expert priors concerning the structure of the sought solution (e.g. spatio-temporal consistency); the challenge is to formulate such criteria in a mixed convex/non differentiable objective function, amenable to tractable optimization.

The activity of the former *Crossing the chasm* SIG gradually decreased after the completion of the 2 PhD theses funded by the Microsoft/Inria joint lab (Adapt project) and devoted to hyper-parameter tuning. Indeed hyper-parameter tuning is still present in TAO, chiefly for continuous optimization (OPT-SIG, section 3.3) and AI planning (CRI-SIG, section 3.4).

## 3.2. Optimal Decision Making under Uncertainty

**Participants:** Olivier Teytaud [correspondent], Jean-Joseph Christophe, Adrien Couëtoux, Hassen Doghmen, Jérémie Decock, Nicolas Galichet, Manuel Loth, Marc Schoenauer, Michèle Sebag.

The UCT-SIG works on sequential optimization problems, where a decision has to be made at each time step along a finite time horizon, and the underlying problem involves uncertainties along an either adversarial or stochastic setting.

Application domains include energy management at various time scales and more generally planning, on the one hand, and games (Go, MineSweeper, NoGo [12]) on the other hand.

The main advances done this year include:

- In the domain of computer Go some new performances have been realized [16] and survey papers have been published in Communications of the ACM [10] and as a chapter [65].
- The extension of Upper Confidence Trees to continuous or large domains (states and/or actions) and to domains with high expertise or strong structure has been done [37], [31], [38].
- The extension of Upper Confidence Trees to multi-objective settings has been done, improving on scalarization-based multi-objective reinforcement learning [56].
- Anytime algorithms for discrete time control (note that the classical stochastic dynamic programming is by no means anytime) have been developed and integrated in the Metis software (section 5.1), based on the above cited results.
- Hybrid approaches combining upper confidence trees and e.g. direct policy search or domain specific approaches to yield robust performance w.r.t. long-term effects and take advantage of the combinatorial structure of the domain have been designed, specifically including problem-specific expertise in the playout phase in the domain of job-shop scheduling [53] or MineSweeper [54].
- Optimization algorithms for direct policy search have been designed [55].
- Within the European STREP Mash project, our favorite tools (in particular Monte-Carlo Tree Search) have been extended to difficult settings with no possibility to “undo” a decision [63]. The notion of “risk of exploration” has been investigated [59].
- An experimental analysis of bandit algorithms for small budget cases [36] got the excellent paper award at TAAI 2012.
- In collaboration with Christian Shulte (KTH, Stockholm), one of the main contributors to the well-known general-purpose CP solver *GECODE* (<http://www.gecode.org/>), and within the Microsoft-Inria joint lab- Adapt project, ideas from UCT have been integrated in GECODE and applied to the job-shop scheduling problem with good first results [52].
- The optimization of low-discrepancy sequences has been done, improving on the best results so far [7]; note that low-discrepancy sequences have been exploited in quite a few of our past works.

### 3.3. Continuous Optimization

**Participants:** Ouassim Ait ElHara, Yohei Akimoto, Anne Auger, Zyed Bouzarkouna, Alexandre Chotard, Nikolaus Hansen, Ilya Loshchilov, Verena Heidrich-Meisner, Yann Ollivier, Marc Schoenauer, Michèle Sebag, Olivier Teytaud, Mouadh Yagoubi.

Our main expertise in continuous optimization is on stochastic search algorithms. We address theory, algorithm design and applications. The methods we investigate are adaptive techniques able to learn iteratively parameters of the distribution used to sample solutions. The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is nowadays one of the most powerful methods for derivative-free continuous optimization. We work on different variants of the CMA-ES to improve it in various contexts as described below. We have proven the convergence of simplified variants of the CMA-ES algorithm using the theory of stochastic approximation providing the first proofs of convergence on composite of twice continuously differentiable functions and used Markov chain analysis for analyzing the step-size adaptation rule of the CMA-ES algorithm.

New algorithms based on surrogates and for constrained optimization. A new variant of CMA-ES to address constraint optimization has been designed [22]. In the Zyed Bouzarkouna's PhD, defended in April 2012, a CIFRE PhD in cooperation with IFP-EN (Institut Français du Pétrole - Energies Nouvelles), new variants of CMA-ES coupled with local-meta-models for expensive optimization have been proposed. They have been applied to well placement problem in oil industry [2]. In the context of his PhD thesis (to be defended in January 2013), Ilya Loshchilov has proposed different surrogate variants of CMA-ES based on ranking-SVM that preserve the invariance to monotonic transformation of the CMA-ES algorithm [51]. He has also explored new restart mechanisms for CMA-ES [47].

Benchmarking. We have continued our effort for improving standards in benchmarking and pursued the development of the COCO (COmparing Continuous Optimizers) platform. We have organized the ACM GECCO 2012 workshop on Black-Box-Optimization Benchmarking<sup>2</sup> and benchmarked different variants of the CMA-ES algorithms [27] [30] [29] [28] [48] [49] [50]. Our new starting ANR project NumBBO, centered on the COCO platform, aims at extending it for large-scale, expensive, constrained and multi-objective optimization.

Theoretical proofs of convergence. We have defined and analyzed a variant of CMA-ES being able to prove that its covariance matrix converges to the inverse Hessian on convex-quadratic functions [18]. We have analyzed the convergence of continuous time trajectories associated to step-size adaptive Evolution Strategies on monotonic  $C^2$ -composite Functions and proved the local convergence towards local minima [19]. This work has been the starting point for analyzing convergence of step-size adaptive ESs using the framework of stochastic approximation (work about to be submitted). In the context of his thesis, Alexandre Chotard has analyzed the step-size adaptation algorithm of CMA-ES on linear function using the theory of Markov Chains [35].

Multi-objective optimization. Mouadh Yagoubi has completed his PhD [3], a CIFRE cooperation with PSA (Peugeot-Citroen automotive industry). His work addressed the multi-disciplinary multi-objective optimization of a diesel engine, that led to propose some asynchronous parallelization of expensive objective functions (more than 2 days per evaluation for the complete 3D model) [57].

### 3.4. Designing criteria

**Participants:** Jamal Atif, Nicolas Bredèche, Cyril Furtlehner, Yoann Isaac, Victorin Martin, Jean-Marc Montanier, Hélène Paugam-Moisy, Marc Schoenauer, Michèle Sebag.

<sup>2</sup>see <http://coco.gforge.inria.fr/doku.php?id=bbob-2012>

This new SIG, rooted on the claim that *What matters is the criterion*, aims at defining new learning or optimization objectives reflecting fundamental properties of the model, the problem or the expert prior knowledge.

A statistical physics perspective. In the context of the ANR project TRAVESTI (<http://travesti.gforge.inria.fr>) which ended this year, we have worked to address more specifically the inverse pairwise Markov random field (MRF) model. On one side we have formalized how the Ising model [64] can be used to perform travel time inference in this context (submitted to AISTATS). Extending the ordinary linear response theory in the vicinity of the so called “Bethe reference point” instead of the interaction-free reference point [67], we were able to provide explicit and tractable formulas to the Plefka’s expansion and the related natural gradient at that point. These can be used into various possible algorithms for generating approximate solutions to the inverse Ising problem which we do investigate now. In parallel we have proposed also in [67], a method based on the “iterative proportional scaling” to learn both the factor graph and the couplings by selecting links one by one (submitted to AISTATS). In the Gaussian MRF case this can be implemented efficiently due to local transformations of the precision matrix after adding one link, which is unfortunately not possible in the Ising MRF. It is competitive with  $L_0$  based approach in terms of precision and computational cost, while incidentally the  $L_1$  based method potentially cheaper is not working well for this problem. The flexibility of the method offers in addition the possibility to combine it with spectral constraints like walk-summability with belief propagation or/and graph structure constrain to enforce compatibility with BP. With no additional computational cost we get a complete set of good trade-offs between likelihood and compatibility with BP. Concerning the analysis of the belief propagation we establish in [60] some sufficient condition for encoding a set of local marginals into a stable belief propagation fixed point. Following some work of last year concerning the modeling of congestion at the microscopic level we have finalized in [9] the analysis of a new family of queuing processes where the service rate is coupled stochastically to the number of clients leading a large deviation formulation of the fundamental diagram of traffic flow.

Multi-objective AI Planning. Within the ANR project DESCARWIN (<http://descarwin.lri.fr>), Mostepha-Redouane Kouadjia worked on the **multi-objective approach** to AI Planning using the Evolutionary Planner *Divide-and-Evolve*, that evolves a sequential decomposition of the problem at hand: each sub-problem is then solved in turn by some embedded classical planner [72]. Even though the embedded planner is single-objective, DaE can nevertheless handle multi-objective problems: Current work includes the implementation of the multi-objective version of DaE, the definition of some benchmark suite, and some first numerical experiments, comparing in particular the results of a full Pareto approach to those of the classical aggregation method. These works resulted in 3 conference papers recently accepted, introducing a tunable benchmark test suite [45], demonstrating that the best quality measure for parameter tuning in this multi-objective framework is the hypervolume, even in the case of the aggregation approach [46], and comparing the evolutionary multi-objective approach with the aggregation method, the only method known to the AI Planning community [44]. The parameter-tuning algorithm designed for DAE, called Learn-and-Optimize, was published in the selection of papers from *Evolution Artificielle* conference [26]. Though originally designed for Evolutionary AI Planning, the method is applicable to domains where instances sharing similar characteristics w.r.t parameter tuning can be grouped in domains.

Image understanding. Sequential image understanding refers to the decision making paradigm where objects in an image are successively segmented/recognized following a predefined strategy. Such an approach generally raises some issues about the “best” segmentation sequence to follow and/or how to avoid error propagation. Within the new sequential recognition framework proposed in [8], these issues are addressed as the objects to segment/recognize are represented by a model describing the spatial relations between objects. The process is guided by a criterion derived from visual attention, specifically a saliency map, used to optimize the segmentation sequence. Spatial knowledge is also used to ensure the consistency of the results and to allow backtracking on the segmentation order if needed. The proposed approach was applied for the segmentation of internal brain structures

in magnetic resonance images. The results show the relevance of the optimization criteria and the relevance of the backtracking procedure to guarantee good and consistent results. In [70] we propose a method for simultaneously segmenting and recognizing objects in images, based on a structural representation of the scene and on a constraint propagation method. Within the ANR project LOGIMA, our goal is to address sequential object recognition as an abduction process [69]. Similar principles are at the core of Yoann Isaac's PhD (Digiteo Unsupervised Brain project), in collaboration with CEA LIST. The dictionary-learning approach used to decompose the EEG signal is required to comply with the structure of the data (e.g. spatio-temporal continuity; submitted).

**Robotic value systems.** Within the European SYMBRION IP, a key milestone toward autonomous cognitive agents has been to provide robots with internal or external rewards, yielding an interesting or competent behavior. Firstly, an objective-free setting referred to as open-ended evolution has been investigated [17], [5], where the criterion to be optimized is left implicit in the reproduction process. Secondly, preference-based reinforcement learning has been investigated in Riad Akrou's PhD, where the robot demonstrations are assessed by the expert and these assessments are used to learn a model of the expert's expectations. In [21], this work has been extended and combined with active learning to yield state-of-the-art performances with few binary feedbacks from the expert. The hormone-based neural net controller first proposed by T. Schmickl et al. has been thoroughly analyzed and simplified in collaboration with Artificial Life Laboratory from Graz [34].

### 3.5. Distributed systems

**Participants:** Cécile Germain-Renaud [correspondent], Philippe Caillou, Dawei Feng, Nadjib Lazaar, Michèle Sebag.

The DIS-SIG explores the issues related to modeling and optimizing distributed systems, ranging from very large scale computational grids to multi-agent systems and distributed constraint solvers.

**Coping with non-stationarity.** Most existing work on modeling the dynamics of grid behavior assumes a steady-state system and concludes to some form of long-range dependence (slowly decaying correlation) in the associated time-series. But the physical (economic and sociological) processes governing the grid behavior dispel the stationarity hypothesis. When the behavior can be modeled as a time series, an appealing class of models is a sequence of stationary processes separated by break points. The optimisation problem for structural break detection is difficult, because of high dimensionality and a complex objective function. Then, evolutionary algorithms are a method of choice. [15] revisits the optimisation strategy in the light of the general advances in evolutionary computation and the specific opportunities for a separable representation. The single-level optimisation problem is decoupled into a bilevel optimisation. The upper level is the problem of finding the optimal number and location of the break points. The lower level optimises the autoregressive models given the number and locations of break points. At the upper level, our optimisation strategy exploits the state-of-the-art CMA-ES (Covariance Matrix Adaptation - Evolutionary Strategy) instead of the relatively straightforward Genetic Algorithm proposed in the classic AutoPARM fitting procedure for non-stationary time series. The associated representation addresses an important shortcoming of : the distance in the chromosomes space better maps to the distance in the model space. Furthermore, the representation becomes scalable. More precisely, it scales linearly with the length of the data set, independently of the cost of the objective function.

**Fault management.** Isolating users from the inevitable faults in large distributed systems is critical to Quality of Experience. Thus a significant part of the software infrastructure of large scale distributed systems collects information that will be exploited to discover if, where, and when the system is faulty. In the context of end-to-end probing as the class of monitoring techniques, minimizing the number of probes for a given discovery performance target is critical. While detection and diagnosis have the obvious advantage of providing an explanation of the failure, by exhibiting culprits, they strongly rely on a priori knowledge that is not available for massively distributed

systems. Thus [39], [62] formulates the problem of probe selection for fault prediction based on end-to-end probing as a Collaborative Prediction (CP) problem, based on the reasonable assumption of an underlying factorial model. On an extensive experimental dataset from the EGI grid, the combination of the Maximum Margin Matrix Factorization approach to CP and Active Learning shows excellent performance, reducing the number of probes typically by 80% to 90%.

**Multi-agent and games.** The main research focus concerning multi-agent systems was on the observation and automatic description of multi-agent based simulations. Whereas usual parameter space exploration systems observe several experiments, the increasing complexity of simulations makes it harder to understand and describe what happens during a single experiment. It is simple to define global indicators to have an overview of the simulation or to follow individual agents, but the most interesting phenomena often occur at an intermediate level, where *groups of agents* are found. The group level is also suited to analyse and display the dynamics of the model. The online and agent-oriented analysis was achieved and its statistical soundness was assessed [6]. In collaboration with the CEA LIST laboratory, we developed a generic tool (SimAnalyser), which can be used online (with NetLogo) or offline (with Logs) to identify, describe, follow [33] and reproduce [58] clusters of agents in a simulation. To select the most interesting clusters and descriptive variables, new activity indicators reflecting the simulation dynamics have been designed [13].

**Parallel SAT Solving.** Recent Parallel SAT solvers use the so-called Conflict-Directed Clause Learning to exchange clauses between the different cores. However, when the number of cores increases, systematic clause sharing leads to communication saturation. Nadjib Lazaar's post-doc, funded by the Microsoft-Inria joint lab, investigated how the communication topology can be optimized online using a Multi-Armed Bandit setting [68], with an improvement of circa 10% (in number of problems solved) and over 50% (in computational time) over ManySAT 2.0, on the 2012 SAT and UNSAT problem suite.

## ALEA Project-Team

### 3. Scientific Foundations

#### 3.1. Scientific Foundations

This idea of analyzing nature systems and transferring the underlying principles into stochastic algorithms and technical implementations is one of the central component of the ALEA team project. Adapting nature mechanisms and biological capabilities clearly provides a better understanding of the real processes, and it also improves the performance and the power of engineers devices. Our project is centered on both the understanding of biological processes in terms of mathematical, physical and chemical models, and on the other hand, on the use of these biology inspired stochastic algorithms to solve complex engineering problems.

There is a huge series of virtual interfaces, robotic devices, numerical schemes and stochastic algorithms which were invented mimicking biological processes or simulating natural mechanisms. The terminology "*mimicking or simulating*" doesn't really mean to find an exact copy of natural processes, but *to elaborate the mathematical principles so that they can be abstracted from the original biological or physical model*. In our context, the whole series of evolutionary type principles discussed in previous sections can be abstracted into only three different and natural classes of stochastic algorithms, depending on the nature of the biology-inspired interaction mechanism used in the stochastic evolution model. These three stochastic search models are listed below :

1) *Branching and interacting particle systems ( birth and death chains, spatial branching processes, mean-field interaction between generations):*

The first generation of adaptive branching-selection algorithms is very often built on the same genetic type paradigm: When exploring a state space with many particles, we duplicate better fitted individuals at the expense of light particles with poor fitness die. From a computational point of view, we generate a large number of random problem solvers. Each one is then rated according to a fitness or performance function defined by the developer. Mimicking natural selection, an evolutionary algorithm selects the best solvers in each generation and breeds them.

2) *Reinforced random walks and self-interacting chains (reinforced learning strategies, interaction processes with respect to the occupation measure of the past visited sites):*

This type of reinforcement is observed frequently in nature and society, where "beneficial" interactions with the past history tend to be repeated. A new class of historical mean field type interpretation models of reinforced processes were developed by the team project leader in a pair of articles [56], [55]. Self interaction gives the opportunity to build new stochastic search algorithms with the ability to, in a sense, re-initialized their exploration from the past, re-starting from some better fitted initial value already met in the past [57], [58].

3) *Random tree based stochastic exploration models (coalescent and genealogical tree search explorations techniques on path space):*

The last generation of stochastic random tree models is concerned with biology-inspired algorithms on paths and excursions spaces. These genealogical adaptive search algorithms coincide with genetic type particle models in excursion spaces. They have been applied with success in generating the excursion distributions of Markov processes evolving in critical and rare event regimes, as well as in path estimation and related smoothing problems arising in advanced signal processing (cf. [53] and references therein). We underline the fact that the complete mathematical analysis of these random tree models, including their long time behavior, their propagations of chaos properties, as well as their combinatorial structures are far from being completed. This class of genealogical tree based models has been introduced in [54] for solving smoothing problems and more generally Feynman-Kac semigroups on path spaces, see also [52], [53], and references therein.

## ASPI Project-Team

### 3. Scientific Foundations

#### 3.1. Interacting Monte Carlo methods and particle approximation of Feynman–Kac distributions

Monte Carlo methods are numerical methods that are widely used in situations where (i) a stochastic (usually Markovian) model is given for some underlying process, and (ii) some quantity of interest should be evaluated, that can be expressed in terms of the expected value of a functional of the process trajectory, which includes as an important special case the probability that a given event has occurred. Numerous examples can be found, e.g. in financial engineering (pricing of options and derivative securities) [50], in performance evaluation of communication networks (probability of buffer overflow), in statistics of hidden Markov models (state estimation, evaluation of contrast and score functions), etc. Very often in practice, no analytical expression is available for the quantity of interest, but it is possible to simulate trajectories of the underlying process. The idea behind Monte Carlo methods is to generate independent trajectories of this process or of an alternate instrumental process, and to build an approximation (estimator) of the quantity of interest in terms of the weighted empirical probability distribution associated with the resulting independent sample. By the law of large numbers, the above estimator converges as the size  $N$  of the sample goes to infinity, with rate  $1/\sqrt{N}$  and the asymptotic variance can be estimated using an appropriate central limit theorem. To reduce the variance of the estimator, many variance reduction techniques have been proposed. Still, running independent Monte Carlo simulations can lead to very poor results, because trajectories are generated *blindly*, and only afterwards are the corresponding weights evaluated. Some of the weights can happen to be negligible, in which case the corresponding trajectories are not going to contribute to the estimator, i.e. computing power has been wasted.

A recent and major breakthrough, has been the introduction of interacting Monte Carlo methods, also known as sequential Monte Carlo (SMC) methods, in which a whole (possibly weighted) sample, called *system of particles*, is propagated in time, where the particles

- *explore* the state space under the effect of a *mutation* mechanism which mimics the evolution of the underlying process,
- and are *replicated* or *terminated*, under the effect of a *selection* mechanism which automatically concentrates the particles, i.e. the available computing power, into regions of interest of the state space.

In full generality, the underlying process is a discrete–time Markov chain, whose state space can be finite, continuous, hybrid (continuous / discrete), graphical, constrained, time varying, pathwise, etc.,

the only condition being that it can easily be *simulated*.

In the special case of particle filtering, originally developed within the tracking community, the algorithms yield a numerical approximation of the optimal Bayesian filter, i.e. of the conditional probability distribution of the hidden state given the past observations, as a (possibly weighted) empirical probability distribution of the system of particles. In its simplest version, introduced in several different scientific communities under the name of *bootstrap filter* [52], *Monte Carlo filter* [57] or *condensation* (conditional density propagation) algorithm [54], and which historically has been the first algorithm to include a redistribution step, the selection mechanism is governed by the likelihood function: at each time step, a particle is more likely to survive and to replicate at the next generation if it is consistent with the current observation. The algorithms also provide as a by–product a numerical approximation of the likelihood function, and of many other contrast functions for parameter estimation in hidden Markov models, such as the prediction error or the conditional least–squares criterion.



Particle methods are currently being used in many scientific and engineering areas

positioning, navigation, and tracking [53], [47], visual tracking [54], mobile robotics [48], [71], ubiquitous computing and ambient intelligence, sensor networks, risk evaluation and simulation of rare events [51], genetics, molecular simulation [49], etc.

Other examples of the many applications of particle filtering can be found in the contributed volume [34] and in the special issue of *IEEE Transactions on Signal Processing* devoted to *Monte Carlo Methods for Statistical Signal Processing* in February 2002, where the tutorial paper [35] can be found, and in the textbook [67] devoted to applications in target tracking. Applications of sequential Monte Carlo methods to other areas, beyond signal and image processing, e.g. to genetics, can be found in [66]. A recent overview can also be found in [37].

Particle methods are very easy to implement, since it is sufficient in principle to simulate independent trajectories of the underlying process. The whole problematic is multidisciplinary, not only because of the already mentioned diversity of the scientific and engineering areas in which particle methods are used, but also because of the diversity of the scientific communities which have contributed to establish the foundations of the field

target tracking, interacting particle systems, empirical processes, genetic algorithms (GA), hidden Markov models and nonlinear filtering, Bayesian statistics, Markov chain Monte Carlo (MCMC) methods.

These algorithms can be interpreted as numerical approximation schemes for Feynman–Kac distributions, a pathwise generalization of Gibbs–Boltzmann distributions, in terms of the weighted empirical probability distribution associated with a system of particles. This abstract point of view [42], [40], has proved to be extremely fruitful in providing a very general framework to the design and analysis of numerical approximation schemes, based on systems of branching and / or interacting particles, for nonlinear dynamical systems with values in the space of probability distributions, associated with Feynman–Kac distributions. Many asymptotic results have been proved as the number  $N$  of particles (sample size) goes to infinity, using techniques coming from applied probability (interacting particle systems, empirical processes [74]), see e.g. the survey article [42] or the recent textbook [40], and references therein

convergence in  $L^p$ , convergence as empirical processes indexed by classes of functions, uniform convergence in time, see also [63], [64], central limit theorem, see also [59], propagation of chaos, large deviations principle, etc.

The objective here is to systematically study the impact of the many algorithmic variants on the convergence results.

## 3.2. Statistics of HMM

Hidden Markov models (HMM) form a special case of partially observed stochastic dynamical systems, in which the state of a Markov process (in discrete or continuous time, with finite or continuous state space) should be estimated from noisy observations. The conditional probability distribution of the hidden state given past observations is a well-known example of a normalized (nonlinear) Feynman–Kac distribution, see 3.1. These models are very flexible, because of the introduction of latent variables (non observed) which allows to model complex time dependent structures, to take constraints into account, etc. In addition, the underlying Markovian structure makes it possible to use numerical algorithms (particle filtering, Markov chain Monte Carlo methods (MCMC), etc.) which are computationally intensive but whose complexity is rather small. Hidden Markov models are widely used in various applied areas, such as speech recognition, alignment of biological sequences, tracking in complex environment, modeling and control of networks, digital communications, etc.

Beyond the recursive estimation of a hidden state from noisy observations, the problem arises of statistical inference of HMM with general state space [38], including estimation of model parameters, early monitoring and diagnosis of small changes in model parameters, etc.

**Large time asymptotics** A fruitful approach is the asymptotic study, when the observation time increases to infinity, of an extended Markov chain, whose state includes (i) the hidden state, (ii) the observation, (iii) the prediction filter (i.e. the conditional probability distribution of the hidden state given observations at all previous time instants), and possibly (iv) the derivative of the prediction filter with respect to the parameter. Indeed, it is easy to express the log-likelihood function, the conditional least-squares criterion, and many other classical contrast processes, as well as their derivatives with respect to the parameter, as additive functionals of the extended Markov chain.

The following general approach has been proposed

- first, prove an exponential stability property (i.e. an exponential forgetting property of the initial condition) of the prediction filter and its derivative, for a misspecified model,
- from this, deduce a geometric ergodicity property and the existence of a unique invariant probability distribution for the extended Markov chain, hence a law of large numbers and a central limit theorem for a large class of contrast processes and their derivatives, and a local asymptotic normality property,
- finally, obtain the consistency (i.e. the convergence to the set of minima of the associated contrast function), and the asymptotic normality of a large class of minimum contrast estimators.

This programme has been completed in the case of a finite state space [7], and has been generalized [43] under an uniform minoration assumption for the Markov transition kernel, which typically does only hold when the state space is compact. Clearly, the whole approach relies on the existence of an exponential stability property of the prediction filter, and the main challenge currently is to get rid of this uniform minoration assumption for the Markov transition kernel [41], [64], so as to be able to consider more interesting situations, where the state space is noncompact.

**Small noise asymptotics** Another asymptotic approach can also be used, where it is rather easy to obtain interesting explicit results, in terms close to the language of nonlinear deterministic control theory [58]. Taking the simple example where the hidden state is the solution to an ordinary differential equation, or a nonlinear state model, and where the observations are subject to additive Gaussian white noise, this approach consists in assuming that covariances matrices of the state noise and of the observation noise go simultaneously to zero. If it is reasonable in many applications to consider that noise covariances are small, this asymptotic approach is less natural than the large time asymptotics, where it is enough (provided a suitable ergodicity assumption holds) to accumulate observations and to see the expected limit laws (law of large numbers, central limit theorem, etc.). In opposition, the expressions obtained in the limit (Kullback-Leibler divergence, Fisher information matrix, asymptotic covariance matrix, etc.) take here a much more explicit form than in the large time asymptotics.

The following results have been obtained using this approach

- the consistency of the maximum likelihood estimator (i.e. the convergence to the set  $M$  of global minima of the Kullback-Leibler divergence), has been obtained using large deviations techniques, with an analytical approach [55],
- if the abovementioned set  $M$  does not reduce to the true parameter value, i.e. if the model is not identifiable, it is still possible to describe precisely the asymptotic behavior of the estimators [56]: in the simple case where the state equation is a noise-free ordinary differential equation and using a Bayesian framework, it has been shown that (i) if the rank  $r$  of the Fisher information matrix  $I$  is constant in a neighborhood of the set  $M$ , then this set is a differentiable submanifold of codimension  $r$ , (ii) the posterior probability distribution of the parameter converges to a random probability distribution in the limit, supported by the manifold  $M$ , absolutely continuous w.r.t. the Lebesgue measure on  $M$ , with an explicit expression for the density, and (iii) the posterior probability distribution of the suitably normalized difference between the parameter and its projection on the manifold  $M$ , converges to a mixture of Gaussian probability distributions on the normal spaces to the manifold  $M$ , which generalized the usual asymptotic normality property,

- it has been shown [65] that (i) the parameter dependent probability distributions of the observations are locally asymptotically normal (LAN) [61], from which the asymptotic normality of the maximum likelihood estimator follows, with an explicit expression for the asymptotic covariance matrix, i.e. for the Fisher information matrix  $I$ , in terms of the Kalman filter associated with the linear tangent linear Gaussian model, and (ii) the score function (i.e. the derivative of the log-likelihood function w.r.t. the parameter), evaluated at the true value of the parameter and suitably normalized, converges to a Gaussian r.v. with zero mean and covariance matrix  $I$ .

### 3.3. Multilevel splitting for rare event simulation

See 4.2, and 5.1, 5.6, 5.10 and 5.11.

The estimation of the small probability of a rare but critical event, is a crucial issue in industrial areas such as nuclear power plants, food industry, telecommunication networks, finance and insurance industry, air traffic management, etc.

In such complex systems, analytical methods cannot be used, and naive Monte Carlo methods are clearly inefficient to estimate accurately very small probabilities. Besides importance sampling, an alternate widespread technique consists in multilevel splitting [60], where trajectories going towards the critical set are given offsprings, thus increasing the number of trajectories that eventually reach the critical set. As shown in [5], the Feynman–Kac formalism of 3.1 is well suited for the design and analysis of splitting algorithms for rare event simulation.

**Propagation of uncertainty** Multilevel splitting can be used in static situations. Here, the objective is to learn the probability distribution of an output random variable  $Y = F(X)$ , where the function  $F$  is only defined pointwise for instance by a computer programme, and where the probability distribution of the input random variable  $X$  is known and easy to simulate from. More specifically, the objective could be to compute the probability of the output random variable exceeding a threshold, or more generally to evaluate the cumulative distribution function of the output random variable for different output values. This problem is characterized by the lack of an analytical expression for the function, the computational cost of a single pointwise evaluation of the function, which means that the number of calls to the function should be limited as much as possible, and finally the complexity and / or unavailability of the source code of the computer programme, which makes any modification very difficult or even impossible, for instance to change the model as in importance sampling methods.

The key issue is to learn as fast as possible regions of the input space which contribute most to the computation of the target quantity. The proposed splitting methods consists in (i) introducing a sequence of intermediate regions in the input space, implicitly defined by exceeding an increasing sequence of thresholds or levels, (ii) counting the fraction of samples that reach a level given that the previous level has been reached already, and (iii) improving the diversity of the selected samples, usually using an artificial Markovian dynamics. In this way, the algorithm learns

- the transition probability between successive levels, hence the probability of reaching each intermediate level,
- and the probability distribution of the input random variable, conditioned on the output variable reaching each intermediate level.

A further remark, is that this conditional probability distribution is precisely the optimal (zero variance) importance distribution needed to compute the probability of reaching the considered intermediate level.

**Rare event simulation** To be specific, consider a complex dynamical system modelled as a Markov process, whose state can possibly contain continuous components and finite components (mode, regime, etc.), and the objective is to compute the probability, hopefully very small, that a critical region of the state space is reached by the Markov process before a final time  $T$ , which can be deterministic and fixed, or random (for instance the time of return to a recurrent set, corresponding to a nominal behaviour).

The proposed splitting method consists in (i) introducing a decreasing sequence of intermediate, more and more critical, regions in the state space, (ii) counting the fraction of trajectories that reach an intermediate region before time  $T$ , given that the previous intermediate region has been reached before time  $T$ , and (iii) regenerating the population at each stage, through redistribution. In addition to the non-intrusive behaviour of the method, the splitting methods make it possible to learn the probability distribution of typical critical trajectories, which reach the critical region before final time  $T$ , an important feature that methods based on importance sampling usually miss. Many variants have been proposed, whether

- the branching rate (number of offsprings allocated to a successful trajectory) is fixed, which allows for depth-first exploration of the branching tree, but raises the issue of controlling the population size,
- the population size is fixed, which requires a breadth-first exploration of the branching tree, with random (multinomial) or deterministic allocation of offsprings, etc.

Just as in the static case, the algorithm learns

- the transition probability between successive levels, hence the probability of reaching each intermediate level,
- and the entrance probability distribution of the Markov process in each intermediate region.

Contributions have been given to

- minimizing the asymptotic variance, obtained through a central limit theorem, with respect to the shape of the intermediate regions (selection of the importance function), to the thresholds (levels), to the population size, etc.
- controlling the probability of extinction (when not even one trajectory reaches the next intermediate level),
- designing and studying variants suited for hybrid state space (resampling per mode, marginalization, mode aggregation),

and in the static case, to

- minimizing the asymptotic variance, obtained through a central limit theorem, with respect to intermediate levels, to the Metropolis kernel introduced in the mutation step, etc.

A related issue is global optimization. Indeed, the difficult problem of finding the set  $M$  of global minima of a real-valued function  $V$  can be replaced by the apparently simpler problem of sampling a population from a probability distribution depending on a small parameter, and asymptotically supported by the set  $M$  as the small parameter goes to zero. The usual approach here is to use the cross-entropy method [68], [39], which relies on learning the optimal importance distribution within a prescribed parametric family. On the other hand, multilevel splitting methods could provide an alternate nonparametric approach to this problem.

### 3.4. Nearest neighbor estimates

This additional topic was not present in the initial list of objectives, and has emerged only recently.

In pattern recognition and statistical learning, also known as machine learning, nearest neighbor (NN) algorithms are amongst the simplest but also very powerful algorithms available. Basically, given a training set of data, i.e. an  $N$ -sample of i.i.d. object-feature pairs, with real-valued features, the question is how to generalize, that is how to guess the feature associated with any new object. To achieve this, one chooses some integer  $k$  smaller than  $N$ , and takes the mean-value of the  $k$  features associated with the  $k$  objects that are nearest to the new object, for some given metric.

In general, there is no way to guess exactly the value of the feature associated with the new object, and the minimal error that can be done is that of the Bayes estimator, which cannot be computed by lack of knowledge of the distribution of the object–feature pair, but the Bayes estimator can be useful to characterize the strength of the method. So the best that can be expected is that the NN estimator converges, say when the sample size  $N$  grows, to the Bayes estimator. This is what has been proved in great generality by Stone [69] for the mean square convergence, provided that the object is a finite–dimensional random variable, the feature is a square–integrable random variable, and the ratio  $k/N$  goes to 0. Nearest neighbor estimator is not the only local averaging estimator with this property, but it is arguably the simplest.

The asymptotic behavior when the sample size grows is well understood in finite dimension, but the situation is radically different in general infinite dimensional spaces, when the objects to be classified are functions, images, etc.

**Nearest neighbor classification in infinite dimension** In finite dimension, the  $k$ –nearest neighbor classifier is universally consistent, i.e. its probability of error converges to the Bayes risk as  $N$  goes to infinity, whatever the joint probability distribution of the pair, provided that the ratio  $k/N$  goes to zero. Unfortunately, this result is no longer valid in general metric spaces, and the objective is to find out reasonable sufficient conditions for the weak consistency to hold. Even in finite dimension, there are exotic distances such that the nearest neighbor does not even get closer (in the sense of the distance) to the point of interest, and the state space needs to be complete for the metric, which is the first condition. Some regularity on the regression function is required next. Clearly, continuity is too strong because it is not required in finite dimension, and a weaker form of regularity is assumed. The following consistency result has been obtained: if the metric space is separable and if some Besicovich condition holds, then the nearest neighbor classifier is weakly consistent. Note that the Besicovich condition is always fulfilled in finite dimensional vector spaces (this result is called the Besicovich theorem), and that a counterexample [3] can be given in an infinite dimensional space with a Gaussian measure (in this case, the nearest neighbor classifier is clearly nonconsistent). Finally, a simple example has been found which verifies the Besicovich condition with a noncontinuous regression function.

**Rates of convergence of the functional  $k$ –nearest neighbor estimator** Motivated by a broad range of potential applications, such as regression on curves, rates of convergence of the  $k$ –nearest neighbor estimator of the regression function, based on  $N$  independent copies of the object–feature pair, have been investigated when the object is in a suitable ball in some functional space. Using compact embedding theory, explicit and general finite sample bounds can be obtained for the expected squared difference between the  $k$ –nearest neighbor estimator and the Bayes regression function, in a very general setting. The results have also been particularized to classical function spaces such as Sobolev spaces, Besov spaces and reproducing kernel Hilbert spaces. The rates obtained are genuine nonparametric convergence rates, and up to our knowledge the first of their kind for  $k$ –nearest neighbor regression.

This emerging topic has produced several theoretical advances [1], [2] in collaboration with Gérard Biau (université Pierre et Marie Curie, ENS Paris and EPI CLASSIC, Inria Paris—Rocquencourt), and a possible target application domain has been identified in the statistical analysis of recommendation systems, that would be a source of interesting problems.

## CQFD Project-Team

### 3. Scientific Foundations

#### 3.1. Introduction

The scientific objectives of the team are to provide mathematical tools for modeling and optimization of complex systems. These systems require mathematical representations which are in essence dynamic, multi-model and stochastic. This increasing complexity poses genuine scientific challenges in the domain of modeling and optimization. More precisely, our research activities are focused on stochastic optimization and (parametric, semi-parametric, multidimensional) statistics which are complementary and interlinked topics. It is essential to develop simultaneously statistical methods for the estimation and control methods for the optimization of the models.

#### 3.2. Main research topics

- Stochastic modeling: Markov chain, Piecewise Deterministic Markov Processes (PDMP), Markov Decision Processes (MDP).

The mathematical representation of complex systems is a preliminary step to our final goal corresponding to the optimization of its performance. For example, in order to optimize the predictive maintenance of a system, it is necessary to choose the adequate model for its representation. The step of modeling is crucial before any estimation or computation of quantities related to its optimization. For this we have to represent all the different regimes of the system and the behavior of the physical variables under each of these regimes. Moreover, we must also select the dynamic variables which have a potential effect on the physical variable and the quantities of interest. The team CQFD works on the theory of Piecewise Deterministic Markov Processes (PDMP's) and on Markov Decision Processes (MDP's). These two classes of systems form general families of controlled stochastic processes suitable for the modeling of sequential decision-making problems in the continuous-time (PDMPs) and discrete-time (MDP's) context. They appear in many fields such as engineering, computer science, economics, operations research and constitute powerful class of processes for the modeling of complex system.

- Estimation methods: estimation for PDMP; estimation in non- and semi parametric regression modeling.

To the best of our knowledge, there does not exist any general theory for the problems of estimating parameters of PDMPs although there already exist a large number of tools for sub-classes of PDMPs such as point processes and marked point processes. However, to fill the gap between these specific models and the general class of PDMPs, new theoretical and mathematical developments will be on the agenda of the whole team. In the framework of non-parametric regression or quantile regression, we focus on kernel estimators or kernel local linear estimators for complete data or censored data. New strategies for estimating semi-parametric models via recursive estimation procedures have also received an increasing interest recently. The advantage of the recursive estimation approach is to take into account the successive arrivals of the information and to refine, step after step, the implemented estimation algorithms. These recursive methods do require restarting calculation of parameter estimation from scratch when new data are added to the base. The idea is to use only the previous estimations and the new data to refresh the estimation. The gain in time could be very interesting and there are many applications of such approaches.

- Dimension reduction: dimension-reduction via SIR and related methods, dimension-reduction via multidimensional and classification methods.

Most of the dimension reduction approaches seek for lower dimensional subspaces minimizing the loss of some statistical information. This can be achieved in modeling framework or in exploratory data analysis context.

In modeling framework we focus our attention on semi-parametric models in order to conjugate the advantages of parametric and nonparametric modeling. On the one hand, the parametric part of the model allows a suitable interpretation for the user. On the other hand, the functional part of the model offers a lot of flexibility. In this project, we are especially interested in the semi-parametric regression model  $Y = f(X'\theta) + \varepsilon$ , the unknown parameter  $\theta$  belongs to  $\mathbb{R}^p$  for a single index model, or is such that  $\theta = [\theta_1, \dots, \theta_d]$  (where each  $\theta_k$  belongs to  $\mathbb{R}^p$  and  $d \leq p$  for a multiple indices model), the noise  $\varepsilon$  is a random error with unknown distribution, and the link function  $f$  is an unknown real valued function. Another way to see this model is the following: the variables  $X$  and  $Y$  are independent given  $X'\theta$ . In our semi-parametric framework, the main objectives are to estimate the parametric part  $\theta$  as well as the nonparametric part which can be the link function  $f$ , the conditional distribution function of  $Y$  given  $X$  or the conditional quantile  $q_\alpha$ . In order to estimate the dimension reduction parameter  $\theta$  we focus on the Sliced Inverse Regression (SIR) method which has been introduced by Li [88] and Duan and Li [76]

Methods of dimension reduction are also important tools in the field of data analysis, data mining and machine learning. They provide a way to understand and visualize the structure of complex data sets. Traditional methods among others are principal component analysis for quantitative variables or multiple component analysis for qualitative variables. New techniques have also been proposed to address these challenging tasks involving many irrelevant and redundant variables and often comparably few observation units. In this context, we focus on the problem of synthetic variables construction, whose goals include increasing the predictor performance and building more compact variables subsets. Clustering of variables is used for feature construction. The idea is to replace a group of "similar" variables by a cluster centroid, which becomes a feature. The most popular algorithms include K-means and hierarchical clustering. For a review, see, e.g., the textbook of Duda [77]

- Stochastic optimal control: optimal stopping, impulse control, continuous control, linear programming, singular perturbation, martingale problem.

The first objective is to focus on the development of computational methods.

- In the continuous-time context, stochastic control theory has from the numerical point of view, been mainly concerned with Stochastic Differential Equations (SDEs in short). From the practical and theoretical point of view, the numerical developments for this class of processes are extensive and largely complete. It capitalizes on the connection between SDEs and second order partial differential equations (PDEs in short) and the fact that the properties of the latter equations are very well understood. It is, however, hard to deny that the development of computational methods for the control of PDMPs has received little attention. One of the main reasons is that the role played by the familiar PDEs in the diffusion models is here played by certain systems of integro-differential equations for which there is not (and cannot be) a unified theory such as for PDEs as emphasized by M.H.A. Davis in his book. To the best knowledge of the team, there is only one attempt to tackle this difficult problem by O.L.V. Costa and M.H.A. Davis. The originality of our project consists in studying this unexplored area. It is very important to stress the fact that these numerical developments will give rise to a lot of theoretical issues such as type of approximations, convergence results, rates of convergence,....
- Theory for MDP's has reached a rather high degree of maturity, although the classical tools such as value iteration, policy iteration and linear programming, and their various extensions, are not applicable in practice. We believe that the theoretical progress of MDP's must be in parallel with the corresponding numerical developments. Therefore, solving

MDP's numerically is an awkward and important problem both from the theoretical and practical point of view. In order to meet this challenge, the fields of neural networks, neuro-dynamic programming and approximate dynamic programming became recently an active area of research. Such methods found their roots in heuristic approaches, but theoretical results for convergence results are mainly obtained in the context of finite MDP's. Hence, an ambitious challenge is to investigate such numerical problems but for models with general state and action spaces. Our motivation is to develop theoretically consistent computational approaches for approximating optimal value functions and finding optimal policies.

Analysis of various problems arising in MDPs leads to a large variety of interesting mathematical problems. The second objective of the team is to study some theoretical aspects related to MDPs such as convex analytical methods and singular perturbation.



## I4S Team

### 3. Scientific Foundations

#### 3.1. Introduction

In this section, the main features for the key monitoring issues, namely identification, detection, and diagnostics, are provided, and a particular instantiation relevant for vibration monitoring is described.

It should be stressed that the foundations for identification, detection, and diagnostics, are fairly general, if not generic. Handling high order linear dynamical systems, in connection with finite elements models, which call for using subspace-based methods, is specific to vibration-based SHM. Actually, one particular feature of model-based sensor information data processing as exercised in I4S, is the combined use of black-box or semi-physical models together with physical ones. Black-box and semi-physical models are, for example, eigenstructure parameterizations of linear MIMO systems, of interest for modal analysis and vibration-based SHM. Such models are intended to be identifiable. However, due to the large model orders that need to be considered, the issue of model order selection is really a challenge. Traditional advanced techniques from statistics such as the various forms of Akaike criteria (AIC, BIC, MDL, ...) do not work at all. This gives rise to new research activities specific to handling high order models.

Our approach to monitoring assumes that a model of the monitored system is available. This is a reasonable assumption, especially within the SHM areas. The main feature of our monitoring method is its intrinsic ability to the early warning of small deviations of a system with respect to a reference (safe) behavior under usual operating conditions, namely without any artificial excitation or other external action. Such a normal behavior is summarized in a reference parameter vector  $\theta_0$ , for example a collection of modes and mode-shapes.

#### 3.2. Identification

The behavior of the monitored continuous system is assumed to be described by a parametric model  $\{\mathbf{P}_\theta, \theta \in \Theta\}$ , where the distribution of the observations  $(Z_0, \dots, Z_N)$  is characterized by the parameter vector  $\theta \in \Theta$ . An *estimating function*, for example of the form :

$$\mathcal{K}_N(\theta) = 1/N \sum_{k=0}^N K(\theta, Z_k)$$

is such that  $\mathbf{E}_\theta[\mathcal{K}_N(\theta)] = 0$  for all  $\theta \in \Theta$ . In many situations,  $\mathcal{K}$  is the gradient of a function to be minimized : squared prediction error, log-likelihood (up to a sign), .... For performing model identification on the basis of observations  $(Z_0, \dots, Z_N)$ , an estimate of the unknown parameter is then [34] :

$$\hat{\theta}_N = \arg \{ \theta \in \Theta : \mathcal{K}_N(\theta) = 0 \}$$

Assuming that  $\theta^*$  is the true parameter value, and that  $\mathbf{E}_{\theta^*}[\mathcal{K}_N(\theta)] = 0$  if and only if  $\theta = \theta^*$  with  $\theta^*$  fixed (identifiability condition), then  $\hat{\theta}_N$  converges towards  $\theta^*$ . Thanks to the central limit theorem, the vector  $\mathcal{K}_N(\theta^*)$  is asymptotically Gaussian with zero mean, with covariance matrix  $\Sigma$  which can be either computed or estimated. If, additionally, the matrix  $\mathcal{J}_N = -\mathbf{E}_{\theta^*}[\mathcal{K}'_N(\theta^*)]$  is invertible, then using a Taylor expansion and the constraint  $\mathcal{K}_N(\hat{\theta}_N) = 0$ , the asymptotic normality of the estimate is obtained :

$$\sqrt{N}(\hat{\theta}_N - \theta^*) \approx \mathcal{J}_N^{-1} \sqrt{N} \mathcal{K}_N(\theta^*)$$

In many applications, such an approach must be improved in the following directions :

- *Recursive estimation*: the ability to compute  $\widehat{\theta}_{N+1}$  simply from  $\widehat{\theta}_N$ ;
- *Adaptive estimation*: the ability to *track* the true parameter  $\theta^*$  when it is time-varying.

### 3.3. Detection

Our approach to on-board detection is based on the so-called asymptotic statistical local approach, which we have extended and adapted [5], [4], [2]. It is worth noticing that these investigations of ours have been initially motivated by a vibration monitoring application example. It should also be stressed that, as opposite to many monitoring approaches, our method does not require repeated identification for each newly collected data sample.

For achieving the early detection of small deviations with respect to the normal behavior, our approach generates, on the basis of the reference parameter vector  $\theta_0$  and a new data record, indicators which automatically perform :

- The early detection of a slight mismatch between the model and the data;
- A preliminary diagnostics and localization of the deviation(s);
- The tradeoff between the magnitude of the detected changes and the uncertainty resulting from the estimation error in the reference model and the measurement noise level.

These indicators are computationally cheap, and thus can be embedded. This is of particular interest in some applications, such as flutter monitoring, as explained in module 4.4 .

As in most fault detection approaches, the key issue is to design a *residual*, which is ideally close to zero under normal operation, and has low sensitivity to noises and other nuisance perturbations, but high sensitivity to small deviations, before they develop into events to be avoided (damages, faults, ...). The originality of our approach is to :

- *Design* the residual basically as a *parameter estimating function*,
- *Evaluate* the residual thanks to a kind of central limit theorem, stating that the residual is asymptotically Gaussian and reflects the presence of a deviation in the parameter vector through a change in its own mean vector, which switches from zero in the reference situation to a non-zero value.

This is actually a strong result, which transforms any detection problem concerning a parameterized stochastic *process* into the problem of monitoring the mean of a Gaussian *vector*.

The behavior of the monitored system is again assumed to be described by a parametric model  $\{\mathbf{P}_\theta, \theta \in \Theta\}$ , and the safe behavior of the process is assumed to correspond to the parameter value  $\theta_0$ . This parameter often results from a preliminary identification based on reference data, as in module 3.2 .

Given a new  $N$ -size sample of sensors data, the following question is addressed : *Does the new sample still correspond to the nominal model  $\mathbf{P}_{\theta_0}$  ?* One manner to address this generally difficult question is the following. The asymptotic local approach consists in deciding between the nominal hypothesis and a *close* alternative hypothesis, namely :

$$\text{(Safe) } \mathbf{H}_0 : \theta = \theta_0 \quad \text{and} \quad \text{(Damaged) } \mathbf{H}_1 : \theta = \theta_0 + \eta/\sqrt{N} \quad (60)$$

where  $\eta$  is an unknown but fixed change vector. A residual is generated under the form :

$$\zeta_N = 1/\sqrt{N} \sum_{k=0}^N K(\theta_0, Z_k) = \sqrt{N} \mathcal{K}_N(\theta_0) . \quad (61)$$

If the matrix  $\mathcal{J}_N = -\mathbf{E}_{\theta_0}[\mathcal{K}'_N(\theta_0)]$  converges towards a limit  $\mathcal{J}$ , then the central limit theorem shows [31] that the residual is asymptotically Gaussian :

$$\zeta_N \xrightarrow{N \rightarrow \infty} \begin{cases} \mathcal{N}(0, \Sigma) & \text{under } \mathbf{P}_{\theta_0} , \\ \mathcal{N}(\mathcal{J}\eta, \Sigma) & \text{under } \mathbf{P}_{\theta_0 + \eta/\sqrt{N}} , \end{cases} \quad (62)$$

where the asymptotic covariance matrix  $\Sigma$  can be estimated, and manifests the deviation in the parameter vector by a change in its own mean value. Then, deciding between  $\eta = 0$  and  $\eta \neq 0$  amounts to compute the following  $\chi^2$ -test, provided that  $\mathcal{J}$  is full rank and  $\Sigma$  is invertible :

$$\chi^2 = \bar{\zeta}^T \mathbf{F}^{-1} \bar{\zeta} \geq \lambda . \quad (63)$$

where

$$\bar{\zeta} \triangleq \mathcal{J}^T \Sigma^{-1} \zeta_N \quad \text{and} \quad \mathbf{F} \triangleq \mathcal{J}^T \Sigma^{-1} \mathcal{J} \quad (64)$$

With this approach, it is possible to decide, with a quantifiable error level, if a residual value is significantly different from zero, for assessing whether a fault/damage has occurred. It should be stressed that the residual and the sensitivity and covariance matrices  $\mathcal{J}$  and  $\Sigma$  can be evaluated (or estimated) for the nominal model. In particular, it is *not* necessary to re-identify the model, and the sensitivity and covariance matrices can be pre-computed off-line.

### 3.4. Diagnostics

A further monitoring step, often called *fault isolation*, consists in determining which (subsets of) components of the parameter vector  $\theta$  have been affected by the change. Solutions for that are now described. How this relates to diagnostics is addressed afterwards.

#### 3.4.1. Isolation.

The question: *which (subsets of) components of  $\theta$  have changed ?*, can be addressed using either nuisance parameters elimination methods or a multiple hypotheses testing approach [29]. Here we only sketch two intuitively simple statistical nuisance elimination techniques, which proceed by projection and rejection, respectively.

The fault vector  $\eta$  is partitioned into an informative part and a nuisance part, and the sensitivity matrix  $\mathcal{J}$ , the Fisher information matrix  $\mathbf{F} = \mathcal{J}^T \Sigma^{-1} \mathcal{J}$  and the normalized residual  $\bar{\zeta} = \mathcal{J}^T \Sigma^{-1} \zeta_N$  are partitioned accordingly

$$\eta = \begin{pmatrix} \eta_a \\ \eta_b \end{pmatrix} , \quad \mathcal{J} = \begin{pmatrix} \mathcal{J}_a & \mathcal{J}_b \end{pmatrix} , \quad \mathbf{F} = \begin{pmatrix} \mathbf{F}_{aa} & \mathbf{F}_{ab} \\ \mathbf{F}_{ba} & \mathbf{F}_{bb} \end{pmatrix} , \quad \bar{\zeta} = \begin{pmatrix} \bar{\zeta}_a \\ \bar{\zeta}_b \end{pmatrix} .$$

A rather intuitive statistical solution to the isolation problem, which can be called *sensitivity* approach, consists in projecting the deviations in  $\eta$  onto the subspace generated by the components  $\eta_a$  to be isolated, and deciding between  $\eta_a = \eta_b = 0$  and  $\eta_a \neq 0, \eta_b = 0$ . This results in the following test statistics :

$$t_a = \bar{\zeta}_a^T \mathbf{F}_{aa}^{-1} \bar{\zeta}_a , \quad (65)$$

where  $\bar{\zeta}_a$  is the partial residual (score). If  $t_a \geq t_b$ , the component responsible for the fault is considered to be  $a$  rather than  $b$ .

Another statistical solution to the problem of isolating  $\eta_a$  consists in viewing parameter  $\eta_b$  as a nuisance, and using an existing method for inferring part of the parameters while ignoring and being robust to the complementary part. This method is called *min-max approach*. It consists in replacing the nuisance parameter component  $\eta_b$  by its least favorable value, for deciding between  $\eta_a = 0$  and  $\eta_a \neq 0$ , with  $\eta_b$  unknown. This results in the following test statistics :

$$t_a^* = \bar{\zeta}_a^{*T} \mathbf{F}_a^{*-1} \bar{\zeta}_a^* , \quad (66)$$

where  $\bar{\zeta}_a^* \triangleq \bar{\zeta}_a - \mathbf{F}_{ab} \mathbf{F}_{bb}^{-1} \bar{\zeta}_b$  is the effective residual (score) resulting from the regression of the informative partial score  $\bar{\zeta}_a$  over the nuisance partial score  $\bar{\zeta}_b$ , and where the Schur complement  $\mathbf{F}_a^* = \mathbf{F}_{aa} - \mathbf{F}_{ab} \mathbf{F}_{bb}^{-1} \mathbf{F}_{ba}$  is the associated Fisher information matrix. If  $t_a^* \geq t_b^*$ , the component responsible for the fault is considered to be  $a$  rather than  $b$ .

The properties and relationships of these two types of tests are investigated in [28].

### 3.4.2. Diagnostics.

In most SHM applications, a complex physical system, characterized by a generally non identifiable parameter vector  $\Phi$  has to be monitored using a simple (black-box) model characterized by an identifiable parameter vector  $\theta$ . A typical example is the vibration monitoring problem in module 4.2 , for which complex finite elements models are often available but not identifiable, whereas the small number of existing sensors calls for identifying only simplified input-output (black-box) representations. In such a situation, two different diagnosis problems may arise, namely diagnosis in terms of the black-box parameter  $\theta$  and diagnosis in terms of the parameter vector  $\Phi$  of the underlying physical model.

The isolation methods sketched above are possible solutions to the former. Our approach to the latter diagnosis problem is basically a detection approach again, and not a (generally ill-posed) inverse problem estimation approach [3]. The basic idea is to note that the physical sensitivity matrix writes  $\mathcal{J} \mathcal{J}_{\Phi\theta}$ , where  $\mathcal{J}_{\Phi\theta}$  is the Jacobian matrix at  $\Phi_0$  of the application  $\Phi \mapsto \theta(\Phi)$ , and to use the sensitivity test (6) for the components of the parameter vector  $\Phi$ . Typically this results in the following type of directional test :

$$\chi_{\Phi}^2 = \zeta^T \Sigma^{-1} \mathcal{J} \mathcal{J}_{\Phi\theta} (\mathcal{J}_{\Phi\theta}^T \mathcal{J}^T \Sigma^{-1} \mathcal{J} \mathcal{J}_{\Phi\theta})^{-1} \mathcal{J}_{\Phi\theta}^T \mathcal{J}^T \Sigma^{-1} \zeta \geq \lambda . \quad (67)$$

It should be clear that the selection of a particular parameterization  $\Phi$  for the physical model may have a non negligible influence on such type of tests, according to the numerical conditioning of the Jacobian matrices  $\mathcal{J}_{\Phi\theta}$ .

As a summary, the machinery in modules 3.2 , 3.3 and 3.4 provides us with a generic framework for designing monitoring algorithms for continuous structures, machines and processes. This approach assumes that a model of the monitored system is available. This is a reasonable assumption within the field of applications described in module 4.2 , since most mechanical processes rely on physical principles which write in terms of equations, providing us with models. These important *modeling* and *parameterization* issues are among the questions we intend to investigate within our research program.

The key issue to be addressed within each parametric model class is the residual generation, or equivalently the choice of the *parameter estimating function*.

## 3.5. Subspace-based identification and detection

For reasons closely related to the vibrations monitoring applications described in module 4.2 , we have been investigating subspace-based methods, for both the identification and the monitoring of the eigenstructure  $(\lambda, \phi_\lambda)$  of the state transition matrix  $F$  of a linear dynamical state-space system :

$$\begin{cases} X_{k+1} = F X_k + V_{k+1} \\ Y_k = H X_k \end{cases}, \quad (68)$$

namely the  $(\lambda, \varphi_\lambda)$  defined by :

$$\det (F - \lambda I) = 0, \quad (F - \lambda I) \phi_\lambda = 0, \quad \varphi_\lambda \triangleq H \phi_\lambda \quad (69)$$

The (canonical) parameter vector in that case is :

$$\theta \triangleq \begin{pmatrix} \Lambda \\ \text{vec} \Phi \end{pmatrix} \quad (70)$$

where  $\Lambda$  is the vector whose elements are the eigenvalues  $\lambda$ ,  $\Phi$  is the matrix whose columns are the  $\varphi_\lambda$ 's, and  $\text{vec}$  is the column stacking operator.

Subspace-based methods is the generic name for linear systems identification algorithms based on either time domain measurements or output covariance matrices, in which different subspaces of Gaussian random vectors play a key role [39]. A contribution of ours, minor but extremely fruitful, has been to write the output-only covariance-driven subspace identification method under a form that involves a parameter estimating function, from which we define a *residual adapted to vibration monitoring* [1]. This is explained next.

### 3.5.1. Covariance-driven subspace identification.

Let  $R_i \triangleq \mathbf{E} (Y_k Y_{k-i}^T)$  and:

$$\mathcal{H}_{p+1,q} \triangleq \begin{pmatrix} R_0 & R_1 & \vdots & R_{q-1} \\ R_1 & R_2 & \vdots & R_q \\ \vdots & \vdots & \vdots & \vdots \\ R_p & R_{p+1} & \vdots & R_{p+q-1} \end{pmatrix} \triangleq \text{Hank} (R_i) \quad (71)$$

be the output covariance and Hankel matrices, respectively; and:  $G \triangleq \mathbf{E} (X_k Y_k^T)$ . Direct computations of the  $R_i$ 's from the equations (9) lead to the well known key factorizations :

$$\begin{aligned} R_i &= H F^i G \\ \mathcal{H}_{p+1,q} &= \mathcal{O}_{p+1}(H, F) \mathcal{C}_q(F, G) \end{aligned} \quad (72)$$

where:

$$\mathcal{O}_{p+1}(H, F) \triangleq \begin{pmatrix} H \\ HF \\ \vdots \\ HF^p \end{pmatrix} \quad \text{and} \quad \mathcal{C}_q(F, G) \triangleq (G \quad FG \quad \dots \quad F^{q-1}G) \quad (73)$$

are the observability and controllability matrices, respectively. The observation matrix  $H$  is then found in the first block-row of the observability matrix  $\mathcal{O}$ . The state-transition matrix  $F$  is obtained from the shift invariance property of  $\mathcal{O}$ . The eigenstructure  $(\lambda, \phi_\lambda)$  then results from (10).

Since the actual model order is generally not known, this procedure is run with increasing model orders.

### 3.5.2. Model parameter characterization.

Choosing the eigenvectors of matrix  $F$  as a basis for the state space of model (9) yields the following representation of the observability matrix:

$$\mathcal{O}_{p+1}(\theta) = \begin{pmatrix} \Phi \\ \Phi \Delta \\ \vdots \\ \Phi \Delta^p \end{pmatrix} \quad (74)$$

where  $\Delta \triangleq \text{diag}(\Lambda)$ , and  $\Lambda$  and  $\Phi$  are as in (11). Whether a nominal parameter  $\theta_0$  fits a given output covariance sequence  $(R_j)_j$  is characterized by [1]:

$$\mathcal{O}_{p+1}(\theta_0) \text{ and } \mathcal{H}_{p+1,q} \text{ have the same left kernel space.} \quad (75)$$

This property can be checked as follows. From the nominal  $\theta_0$ , compute  $\mathcal{O}_{p+1}(\theta_0)$  using (15), and perform e.g. a singular value decomposition (SVD) of  $\mathcal{O}_{p+1}(\theta_0)$  for extracting a matrix  $U$  such that:

$$U^T U = I_s \text{ and } U^T \mathcal{O}_{p+1}(\theta_0) = 0 \quad (76)$$

Matrix  $U$  is not unique (two such matrices relate through a post-multiplication with an orthonormal matrix), but can be regarded as a function of  $\theta_0$ . Then the characterization writes:

$$U(\theta_0)^T \mathcal{H}_{p+1,q} = 0 \quad (77)$$

### 3.5.3. Residual associated with subspace identification.

Assume now that a reference  $\theta_0$  and a new sample  $Y_1, \dots, Y_N$  are available. For checking whether the data agree with  $\theta_0$ , the idea is to compute the empirical Hankel matrix  $\hat{\mathcal{H}}_{p+1,q}$ :

$$\hat{\mathcal{H}}_{p+1,q} \triangleq \text{Hank}(\hat{R}_i), \quad \hat{R}_i \triangleq 1/(N-i) \sum_{k=i+1}^N Y_k Y_{k-i}^T \quad (78)$$

and to define the residual vector:

$$\zeta_N(\theta_0) \triangleq \sqrt{N} \text{vec} \left( U(\theta_0)^T \hat{\mathcal{H}}_{p+1,q} \right) \quad (79)$$

Let  $\theta$  be the actual parameter value for the system which generated the new data sample, and  $\mathbf{E}_\theta$  be the expectation when the actual system parameter is  $\theta$ . From (18), we know that  $\zeta_N(\theta_0)$  has zero mean when no change occurs in  $\theta$ , and nonzero mean if a change occurs. Thus  $\zeta_N(\theta_0)$  plays the role of a residual.

It is our experience that this residual has highly interesting properties, both for damage detection [1] and localization [3], and for flutter monitoring [8].

#### 3.5.4. Other uses of the key factorizations.

Factorization (3.5.1) is the key for a characterization of the canonical parameter vector  $\theta$  in (11), and for deriving the residual. Factorization (13) is also the key for :

- Proving consistency and robustness results [6];
- Designing an extension of covariance-driven subspace identification algorithm adapted to the presence and fusion of non-simultaneously recorded multiple sensors setups [7];
- Proving the consistency and robustness of this extension [9];
- Designing various forms of *input-output* covariance-driven subspace identification algorithms adapted to the presence of both known inputs and unknown excitations [10].

**MATHRISK Team (section vide)**



## REGULARITY Project-Team

### 3. Scientific Foundations

#### 3.1. Theoretical aspects: probabilistic modeling of irregularity

The modeling of essentially irregular phenomena is an important challenge, with an emphasis on understanding the sources and functions of this irregularity. Probabilistic tools are well-adapted to this task, provided one can design stochastic models for which the regularity can be measured and controlled precisely. Two points deserve special attention:

- first, the study of regularity has to be *local*. Indeed, in most applications, one will want to act on a system based on local temporal or spatial information. For instance, detection of arrhythmias in ECG or of krachs in financial markets should be performed in “real time”, or, even better, ahead of time. In this sense, regularity is a *local* indicator of the *local* health of a system.
- Second, although we have used the term “irregularity” in a generic and somewhat vague sense, it seems obvious that, in real-world phenomena, regularity comes in many colors, and a rigorous analysis should distinguish between them. As an example, at least two kinds of irregularities are present in financial logs: the local “roughness” of the records, and the local density and height of jumps. These correspond to two different concepts of regularity (in technical terms, Hölder exponents and local index of stability), and they both contribute a different manner to financial risk.

In view of the above, the *Regularity* team focuses on the design of methods that:

1. define and study precisely various relevant measures of local regularity,
2. allow to build stochastic models versatile enough to mimic the rapid variations of the different kinds of regularities observed in real phenomena,
3. allow to estimate as precisely and rapidly as possible these regularities, so as to alert systems in charge of control.

Our aim is to address the three items above through the design of mathematical tools in the field of probability (and, to a lesser extent, statistics), and to apply these tools to uncertainty management as described in the following section. We note here that we do not intend to address the problem of controlling the phenomena based on regularity, that would naturally constitute an item 4 in the list above. Indeed, while we strongly believe that generic tools may be designed to measure and model regularity, and that these tools may be used to analyze real-world applications, in particular in the field of uncertainty management, it is clear that, when it comes to control, application-specific tools are required, that we do not wish to address.

The research topics of the *Regularity* team can be roughly divided into two strongly interacting axes, corresponding to two complementary ways of studying regularity:

1. developments of tools allowing to characterize, measure and estimate various notions of local regularity, with a particular emphasis on the stochastic frame,
2. definition and fine analysis of stochastic models for which some aspects of local regularity may be prescribed.

These two aspects are detailed in sections 3.2 and 3.3 below.

#### 3.2. Tools for characterizing and measuring regularity

##### Fractional Dimensions

Although the main focus of our team is on characterizing *local* regularity, on occasions, it is interesting to use a *global* index of regularity. Fractional dimensions provide such an index. In particular, the *regularization dimension*, that was defined in [30], is well adapted to the study stochastic processes, as its definition allows to build robust estimators in an easy way. Since its introduction, regularization dimension has been used by various teams worldwide in many different applications including the characterization of certain stochastic processes, statistical estimation, the study of mammographies or galactograms for breast carcinomas detection, ECG analysis for the study of ventricular arrhythmia, encephalitis diagnosis from EEG, human skin analysis, discrimination between the nature of radioactive contaminations, analysis of porous media textures, well-logs data analysis, agro-alimentary image analysis, road profile analysis, remote sensing, mechanical systems assessment, analysis of video games, ... (see <http://regularity.saclay.inria.fr/theory/localregularity/biblioregdim> for a list of works using the regularization dimension).

### Hölder exponents

The simplest and most popular measures of local regularity are the pointwise and local Hölder exponents. For a stochastic process  $\{X(t)\}_{t \in \mathbb{R}}$  whose trajectories are continuous and nowhere differentiable, these are defined, at a point  $t_0$ , as the random variables:

$$\alpha_X(t_0, \omega) = \sup \left\{ \alpha : \limsup_{\rho \rightarrow 0} \sup_{t, u \in B(t_0, \rho)} \frac{|X_t - X_u|}{\rho^\alpha} < \infty \right\}, \quad (80)$$

and

$$\tilde{\alpha}_X(t_0, \omega) = \sup \left\{ \alpha : \limsup_{\rho \rightarrow 0} \sup_{t, u \in B(t_0, \rho)} \frac{|X_t - X_u|}{\|t - u\|^\alpha} < \infty \right\}. \quad (81)$$

Although these quantities are in general random, we will omit as is customary the dependency in  $\omega$  and  $X$  and write  $\alpha(t_0)$  and  $\tilde{\alpha}(t_0)$  instead of  $\alpha_X(t_0, \omega)$  and  $\tilde{\alpha}_X(t_0, \omega)$ .

The random functions  $t \mapsto \alpha_X(t_0, \omega)$  and  $t \mapsto \tilde{\alpha}_X(t_0, \omega)$  are called respectively the pointwise and local Hölder functions of the process  $X$ .

The pointwise Hölder exponent is a very versatile tool, in the sense that the set of pointwise Hölder functions of continuous functions is quite large (it coincides with the set of lower limits of sequences of continuous functions [5]). In this sense, the pointwise exponent is often a more precise tool (*i.e.* it varies in a more rapid way) than the local one, since local Hölder functions are always lower semi-continuous. This is why, in particular, it is the exponent that is used as a basis ingredient in multifractal analysis (see section 3.2). For certain classes of stochastic processes, and most notably Gaussian processes, it has the remarkable property that, at each point, it assumes an almost sure value [16]. SRP, mBm, and processes of this kind (see sections 3.3 and 3.3) rely on the sole use of the pointwise Hölder exponent for prescribing the regularity.

However,  $\alpha_X$  obviously does not give a complete description of local regularity, even for continuous processes. It is for instance insensitive to “oscillations”, contrarily to the local exponent. A simple example in the deterministic frame is provided by the function  $x^\gamma \sin(x^{-\beta})$ , where  $\gamma, \beta$  are positive real numbers. This so-called “chirp function” exhibits two kinds of irregularities: the first one, due to the term  $x^\gamma$  is measured by the pointwise Hölder exponent. Indeed,  $\alpha(0) = \gamma$ . The second one is due to the wild oscillations around 0, to which  $\alpha$  is blind. In contrast, the local Hölder exponent at 0 is equal to  $\frac{\gamma}{1+\beta}$ , and is thus influenced by the oscillatory behaviour.

Another, related, drawback of the pointwise exponent is that it is not stable under integro-differentiation, which sometimes makes its use complicated in applications. Again, the local exponent provides here a useful complement to  $\alpha$ , since  $\tilde{\alpha}$  is stable under integro-differentiation.

Both exponents have proved useful in various applications, ranging from image denoising and segmentation to TCP traffic characterization. Applications require precise estimation of these exponents.

### Stochastic 2-microlocal analysis

Neither the pointwise nor the local exponents give a complete characterization of the local regularity, and, although their joint use somewhat improves the situation, it is far from yielding the complete picture.

A fuller description of local regularity is provided by the so-called *2-microlocal analysis*, introduced by J.M. Bony [50]. In this frame, regularity at each point is now specified by two indices, which makes the analysis and estimation tasks more difficult. More precisely, a function  $f$  is said to belong to the *2-microlocal space*  $C_{x_0}^{s,s'}$ , where  $s + s' > 0$ ,  $s' < 0$ , if and only if its  $m = [s + s']$ -th order derivative exists around  $x_0$ , and if there exists  $\delta > 0$ , a polynomial  $P$  with degree lower than  $[s] - m$ , and a constant  $C$ , such that

$$\left| \frac{\partial^m f(x) - P(x)}{|x-x_0|^{[s]-m}} - \frac{\partial^m f(y) - P(y)}{|y-x_0|^{[s]-m}} \right| \leq C|x-y|^{s+s'-m}(|x-y| + |x-x_0|)^{-s'-[s]+m}$$

for all  $x, y$  such that  $0 < |x-x_0| < \delta$ ,  $0 < |y-x_0| < \delta$ . This characterization was obtained in [23], [31]. See [64], [66] for other characterizations and results. These spaces are stable through integro-differentiation, i.e.  $f \in C_x^{s,s'}$  if and only if  $f' \in C_x^{s-1,s'}$ . Knowing to which space  $f$  belongs thus allows to predict the evolution of its regularity after derivation, a useful feature if one uses models based on some kind differential equations. A lot of work remains to be done in this area, in order to obtain more general characterizations, to develop robust estimation methods, and to extend the “2-microlocal formalism”: this is a tool allowing to detect which space a function belongs to, from the computation of the Legendre transform of an auxiliary function known as its *2-microlocal spectrum*. This spectrum provide a wealth of information on the local regularity.

In [16], we have laid some foundations for a stochastic version of 2-microlocal analysis. We believe this will provide a fine analysis of the local regularity of random processes in a direction different from the one detailed for instance in [72]. We have defined random versions of the 2-microlocal spaces, and given almost sure conditions for continuous processes to belong to such spaces. More precise results have also been obtained for Gaussian processes. A preliminary investigation of the 2-microlocal behaviour of Wiener integrals has been performed.

### Multifractal analysis of stochastic processes

A direct use of the local regularity is often fruitful in applications. This is for instance the case in RR analysis or terrain modeling. However, in some situations, it is interesting to supplement or replace it by a more global approach known as *multifractal analysis* (MA). The idea behind MA is to group together all points with same regularity (as measured by the pointwise Hölder exponent) and to measure the “size” of the sets thus obtained [27], [51], [60]. There are mainly two ways to do so, a geometrical and a statistical one.

In the geometrical approach, one defines the *Hausdorff multifractal spectrum* of a process or function  $X$  as the function:  $\alpha \mapsto f_h(\alpha) = \dim \{t : \alpha_X(t) = \alpha\}$ , where  $\dim E$  denotes the Hausdorff dimension of the set  $E$ . This gives a fine measure-theoretic information, but is often difficult to compute theoretically, and almost impossible to estimate on numerical data.

The statistical path to MA is based on the so-called *large deviation multifractal spectrum*:

$$f_g(\alpha) = \lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{\log N_n^\varepsilon(\alpha)}{\log n},$$

where:

$$N_n^\varepsilon(\alpha) = \#\{k : \alpha - \varepsilon \leq \alpha_n^k \leq \alpha + \varepsilon\},$$

and  $\alpha_n^k$  is the “coarse grained exponent” corresponding to the interval  $I_n^k = [\frac{k}{n}, \frac{k+1}{n}]$ , i.e.:

$$\alpha_n^k = \frac{\log |Y_n^k|}{-\log n}.$$

Here,  $Y_n^k$  is some quantity that measures the variation of  $X$  in the interval  $I_n^k$ , such as the increment, the oscillation or a wavelet coefficient.

The large deviation spectrum is typically easier to compute and to estimate than the Hausdorff one. In addition, it often gives more relevant information in applications.

Under very mild conditions (*e.g.* for instance, if the support of  $f_g$  is bounded, [40]) the concave envelope of  $f_g$  can be computed easily from an auxiliary function, called the *Legendre multifractal spectrum*. To do so, one basically interprets the spectrum  $f_g$  as a rate function in a large deviation principle (LDP): define, for  $q \in \mathbb{R}$ ,

$$S_n(q) = \sum_{k=0}^{n-1} |Y_n^k|^q, \quad (82)$$

with the convention  $0^q := 0$  for all  $q \in \mathbb{R}$ . Let:

$$\tau(q) = \liminf_{n \rightarrow \infty} \frac{\log S_n(q)}{-\log(n)}.$$

The Legendre multifractal spectrum of  $X$  is defined as the Legendre transform  $\tau^*$  of  $\tau$ :

$$f_l(\alpha) := \tau^*(\alpha) := \inf_{q \in \mathbb{R}} (q\alpha - \tau(q)).$$

To see the relation between  $f_g$  and  $f_l$ , define the sequence of random variables  $Z_n := \log |Y_n^k|$  where the randomness is through a choice of  $k$  uniformly in  $\{0, \dots, n-1\}$ . Consider the corresponding moment generating functions:

$$c_n(q) := -\frac{\log E_n[\exp(qZ_n)]}{\log(n)}$$

where  $E_n$  denotes expectation with respect to  $P_n$ , the uniform distribution on  $\{0, \dots, n-1\}$ . A version of Gärtner-Ellis theorem ensures that if  $\lim c_n(q)$  exists (in which case it equals  $1 + \tau(q)$ ), and is differentiable, then  $c^* = f_g - 1$ . In this case, one says that the *weak multifractal formalism* holds, *i.e.*  $f_g = f_l$ . In favorable cases, this also coincides with  $f_h$ , a situation referred to as the *strong multifractal formalism*.

Multifractal spectra subsume a lot of information about the distribution of the regularity, that has proved useful in various situations. A most notable example is the strong correlation reported recently in several works between the narrowing of the multifractal spectrum of ECG and certain pathologies of the heart [61], [63]. Let us also mention the multifractality of TCP traffic, that has been both observed experimentally and proved on simplified models of TCP [2], [47].

#### Another colour in local regularity: jumps

As noted above, apart from Hölder exponents and their generalizations, at least another type of irregularity may sometimes be observed on certain real phenomena: discontinuities, which occur for instance on financial logs and certain biomedical signals. In this frame, it is of interest to supplement Hölder exponents and their extensions with (at least) an additional index that measures the local intensity and size of jumps. This is a topic we intend to pursue in full generality in the near future. So far, we have developed an approach in the particular frame of *multistable processes*. We refer to section 3.3 for more details.

### 3.3. Stochastic models

The second axis in the theoretical developments of the *Regularity* team aims at defining and studying stochastic processes for which various aspects of the local regularity may be prescribed.

#### Multifractional Brownian motion

One of the simplest stochastic process for which some kind of control over the Hölder exponents is possible is probably fractional Brownian motion (fBm). This process was defined by Kolmogorov and further studied by Mandelbrot and Van Ness, followed by many authors. The so-called “moving average” definition of fBm reads as follows:

$$Y_t = \int_{-\infty}^0 \left[ (t-u)^{H-\frac{1}{2}} - (-u)^{H-\frac{1}{2}} \right] \cdot \mathbb{W}(du) + \int_0^t (t-u)^{H-\frac{1}{2}} \cdot \mathbb{W}(du),$$

where  $\mathbb{W}$  denotes the real white noise. The parameter  $H$  ranges in  $(0, 1)$ , and it governs the pointwise regularity: indeed, almost surely, at each point, both the local and pointwise Hölder exponents are equal to  $H$ .

Although varying  $H$  yields processes with different regularity, the fact that the exponents are constant along any single path is often a major drawback for the modeling of real world phenomena. For instance, fBm has often been used for the synthesis natural terrains. This is not satisfactory since it yields images lacking crucial features of real mountains, where some parts are smoother than others, due, for instance, to erosion.

It is possible to generalize fBm to obtain a Gaussian process for which the pointwise Hölder exponent may be tuned at each point: the *multifractional Brownian motion (mBm)* is such an extension, obtained by substituting the constant parameter  $H \in (0, 1)$  with a *regularity function*  $H : \mathbb{R}_+ \rightarrow (0, 1)$ .

mBm was introduced independently by two groups of authors: on the one hand, Peltier and Levy-Vehel [28] defined the mBm  $\{X_t; t \in \mathbb{R}_+\}$  from the moving average definition of the fractional Brownian motion, and set:

$$X_t = \int_{-\infty}^0 \left[ (t-u)^{H(t)-\frac{1}{2}} - (-u)^{H(t)-\frac{1}{2}} \right] \cdot \mathbb{W}(du) + \int_0^t (t-u)^{H(t)-\frac{1}{2}} \cdot \mathbb{W}(du),$$

On the other hand, Benassi, Jaffard and Roux [49] defined the mBm from the harmonizable representation of the fBm, *i.e.*:

$$X_t = \int_{\mathbb{R}} \frac{e^{it\xi} - 1}{|\xi|^{H(t)+\frac{1}{2}}} \cdot \widehat{\mathbb{W}}(d\xi),$$

where  $\widehat{\mathbb{W}}$  denotes the complex white noise.

The Hölder exponents of the mBm are prescribed almost surely: the pointwise Hölder exponent is  $\alpha_X(t) = H(t) \wedge \alpha_H(t)$  a.s., and the local Hölder exponent is  $\tilde{\alpha}_X(t) = H(t) \wedge \tilde{\alpha}_H(t)$  a.s. Consequently, the regularity of the sample paths of the mBm are determined by the function  $H$  or by its regularity. The multifractional Brownian motion is our prime example of a stochastic process with prescribed local regularity.

The fact that the local regularity of mBm may be tuned *via* a functional parameter has made it a useful model in various areas such as finance, biomedicine, geophysics, image analysis, .... A large number of studies have been devoted worldwide to its mathematical properties, including in particular its local time. In addition, there is now a rather strong body of work dealing the estimation of its functional parameter, *i.e.* its local regularity. See <http://regularity.saclay.inria.fr/theory/stochasticmodels/bibliombm> for a partial list of works, applied or theoretical, that deal with mBm.

#### Self-regulating processes

We have recently introduced another class of stochastic models, inspired by mBm, but where the local regularity, instead of being tuned “exogenously”, is a function of the amplitude. In other words, at each point  $t$ , the Hölder exponent of the process  $X$  verifies almost surely  $\alpha_X(t) = g(X(t))$ , where  $g$  is a fixed deterministic function verifying certain conditions. A process satisfying such an equation is generically termed a *self-regulating process* (SRP). The particular process obtained by adapting adequately mBm is called the self-regulating multifractional process [3]. Another instance is given by modifying the Lévy construction of Brownian motion [42]. The motivation for introducing self-regulating processes is based on the following general fact: in nature, the local regularity of a phenomenon is often related to its amplitude. An intuitive example is provided by natural terrains: in young mountains, regions at higher altitudes are typically more irregular than regions at lower altitudes. We have verified this fact experimentally on several digital elevation models [7]. Other natural phenomena displaying a relation between amplitude and exponent include temperatures records and RR intervals extracted from ECG [38].

To build the SRMP, one starts from a field of fractional Brownian motions  $B(t, H)$ , where  $(t, H)$  span  $[0, 1] \times [a, b]$  and  $0 < a < b < 1$ . For each fixed  $H$ ,  $B(t, H)$  is a fractional Brownian motion with exponent  $H$ . Denote:

$$\overline{X}_{\alpha'}^{\beta'} = \alpha' + (\beta' - \alpha') \frac{X - \min_K(X)}{\max_K(X) - \min_K(X)}$$

the affine rescaling between  $\alpha'$  and  $\beta'$  of an arbitrary continuous random field over a compact set  $K$ . One considers the following (stochastic) operator, defined almost surely:

$$\begin{aligned} \Lambda_{\alpha', \beta'} : \mathcal{C}([0, 1], [\alpha, \beta]) &\rightarrow \mathcal{C}([0, 1], [\alpha, \beta]) \\ Z(\cdot) &\mapsto \overline{B(\cdot, g(Z(\cdot)))}_{\alpha'}^{\beta'} \end{aligned}$$

where  $\alpha \leq \alpha' < \beta' \leq \beta$ ,  $\alpha$  and  $\beta$  are two real numbers, and  $\alpha', \beta'$  are random variables adequately chosen. One may show that this operator is contractive with respect to the sup-norm. Its unique fixed point is the SRMP. Additional arguments allow to prove that, indeed, the Hölder exponent at each point is almost surely  $g(t)$ .

An example of a two dimensional SRMP with function  $g(x) = 1 - x^2$  is displayed on figure 1 .

We believe that SRP open a whole new and very promising area of research.

### Multistable processes

Non-continuous phenomena are commonly encountered in real-world applications, *e.g.* financial records or EEG traces. For such processes, the information brought by the Hölder exponent must be supplemented by some measure of the density and size of jumps. Stochastic processes with jumps, and in particular Lévy processes, are currently an active area of research.

The simplest class of non-continuous Lévy processes is maybe the one of stable processes [74]. These are mainly characterized by a parameter  $\alpha \in (0, 2]$ , the *stability index* ( $\alpha = 2$  corresponds to the Gaussian case, that we do not consider here). This index measures in some precise sense the intensity of jumps. Paths of stable processes with  $\alpha$  close to 2 tend to display “small jumps”, while, when  $\alpha$  is near 0, their aspect is governed by large ones.

In line with our quest for the characterization and modeling of various notions of local regularity, we have defined *multistable processes*. These are processes which are “locally” stable, but where the stability index  $\alpha$  is now a function of time. This allows to model phenomena which, at times, are “almost continuous”, and at others display large discontinuities. Such a behaviour is for instance obvious on almost any sufficiently long financial record.

More formally, a multistable process is a process which is, at each time  $u$ , tangent to a stable process [59]. Recall that a process  $Y$  is said to be tangent at  $u$  to the process  $Y'_u$  if:

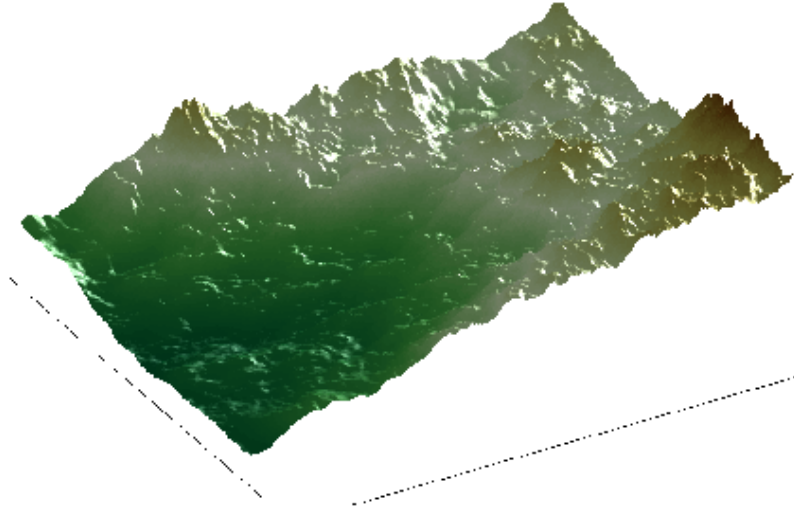


Figure 1. Self-regulating multifractional process with  $g(x) = 1 - x^2$

$$\lim_{r \rightarrow 0} \frac{Y(u + rt) - Y(u)}{r^h} = Y'_u(t), \quad (83)$$

where the limit is understood either in finite dimensional distributions or in the stronger sense of distributions. Note  $Y'_u$  may and in general will vary with  $u$ .

One approach to defining multistable processes is similar to the one developed for constructing mBm [28]: we consider fields of stochastic processes  $X(t, u)$ , where  $t$  is time and  $u$  is an independent parameter that controls the variation of  $\alpha$ . We then consider a “diagonal” process  $Y(t) = X(t, t)$ , which will be, under certain conditions, “tangent” at each point  $t$  to a process  $t \mapsto X(t, u)$ .

A particular class of multistable processes, termed “linear multistable multifractional motions” (lmmm) takes the following form [9], [8]. Let  $(E, \mathcal{E}, m)$  be a  $\sigma$ -finite measure space, and  $\Pi$  be a Poisson process on  $E \times \mathbb{R}$  with mean measure  $m \times \mathcal{L}$  ( $\mathcal{L}$  denotes the Lebesgue measure). An lmmm is defined as:

$$Y(t) = a(t) \sum_{(X,Y) \in \Pi} Y^{<-1/\alpha(t)>} \left( |t - X|^{h(t)-1/\alpha(t)} - |X|^{h(t)-1/\alpha(t)} \right) \quad (t \in \mathbb{R}). \quad (84)$$

where  $x^{<y>} := \text{sign}(x)|x|^y$ ,  $a : \mathbb{R} \rightarrow \mathbb{R}^+$  is a  $C^1$  function and  $\alpha : \mathbb{R} \rightarrow (0, 2)$  and  $h : \mathbb{R} \rightarrow (0, 1)$  are  $C^2$  functions.

In fact, lmmm are somewhat more general than said above: indeed, the couple  $(h, \alpha)$  allows to prescribe at each point, under certain conditions, both the pointwise Hölder exponent and the local intensity of jumps. In this sense, they generalize both the mBm and the linear multifractional stable motion [75]. From a broader perspective, such multistable multifractional processes are expected to provide relevant models for TCP traces, financial logs, EEG and other phenomena displaying time-varying regularity both in terms of Hölder exponents and discontinuity structure.

Figure 2 displays a graph of an lmm with linearly increasing  $\alpha$  and linearly decreasing  $H$ . One sees that the path has large jumps at the beginning, and almost no jumps at the end. Conversely, it is smooth (between jumps) at the beginning, but becomes jaggier and jaggier as time evolves.

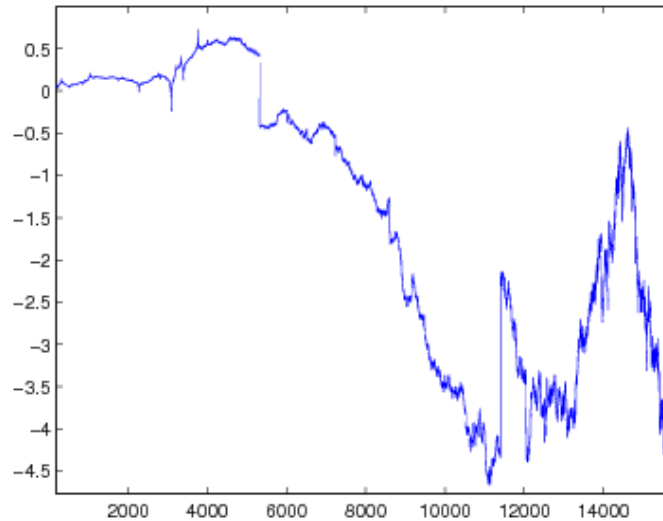


Figure 2. Linear multistable multifractional motion with linearly increasing  $\alpha$  and linearly decreasing  $H$

### Multiparameter processes

In order to use stochastic processes to represent the variability of multidimensional phenomena, it is necessary to define extensions for indices in  $\mathbb{R}^N$  ( $N \geq 2$ ) (see [67] for an introduction to the theory of multiparameter processes). Two different kinds of extensions of multifractional Brownian motion have already been considered: an isotropic extension using the Euclidean norm of  $\mathbb{R}^N$  and a tensor product of one-dimensional processes on each axis. We refer to [13] for a comprehensive survey.

These works have highlighted the difficulty of giving satisfactory definitions for increment stationarity, Hölder continuity and covariance structure which are not closely dependent on the structure of  $\mathbb{R}^N$ . For example, the Euclidean structure can be unadapted to represent natural phenomena.

A promising improvement in the definition of multiparameter extensions is the concept of *set-indexed processes*. A set-indexed process is a process whose indices are no longer “times” or “locations” but may be some compact connected subsets of a metric measure space. In the simplest case, this framework is a generalization of the classical multiparameter processes [62]: usual multiparameter processes are set-indexed processes where the indexing subsets are simply the rectangles  $[0, t]$ , with  $t \in \mathbb{R}_+^N$ .

Set-indexed processes allow for greater flexibility, and should in particular be useful for the modeling of censored data. This situation occurs frequently in biology and medicine, since, for instance, data may not be constantly monitored. Censored data also appear in natural terrain modeling when data are acquired from sensors in presence of hidden areas. In these contexts, set-indexed models should constitute a relevant frame.

A set-indexed extension of fBm is the first step toward the modeling of irregular phenomena within this more general frame. In [18], the so-called *set-indexed fractional Brownian motion (sifBm)* was defined as the mean-zero Gaussian process  $\{\mathbf{B}_U^H; U \in \mathcal{A}\}$  such that



$$\forall U, V \in \mathcal{A}; \quad E[\mathbf{B}_U^H \mathbf{B}_V^H] = \frac{1}{2} \left[ m(U)^{2H} + m(V)^{2H} - m(U \Delta V)^{2H} \right]$$

where  $\mathcal{A}$  is a collection of connected compact subsets of a measure metric space and  $0 < H \leq \frac{1}{2}$ .

This process appears to be the only set-indexed process whose projection on increasing paths is a one-parameter fractional Brownian motion [17]. The construction also provides a way to define fBm's extensions on non-euclidean spaces, *e.g.* indices can belong to the unit hyper-sphere of  $\mathbb{R}^N$ . The study of fractal properties needs specific definitions for increment stationarity and self-similarity of set-indexed processes [20]. We have proved that the sifBm is the only Gaussian set-indexed process satisfying these two (extended) properties.

In the specific case of the indexing collection  $\mathcal{A} = \{[0, t], t \in \mathbb{R}_+^N\} \cup \{\emptyset\}$ , the sifBm can be seen as a multiparameter extension of fBm which is called *multiparameter fractional Brownian motion (MpfBm)*. This process differs from the Lévy fractional Brownian motion and the fractional Brownian sheet, which are also multiparameter extensions of fBm (but do not derive from set-indexed processes). The local behaviour of the sample paths of the MpfBm has been studied in [12]. The self-similarity index  $H$  is proved to be the almost sure value of the local Hölder exponent at any point, and the Hausdorff dimension of the graph is determined in function of  $H$ .

The increment stationarity property for set-indexed processes, previously defined in the study of the sifBm, allows to consider set-indexed processes whose increments are independent and stationary. This generalizes the definition of Bass-Pyke and Adler-Feigin for Lévy processes indexed by subsets of  $\mathbb{R}^N$ , to a more general indexing collection. We have obtained a Lévy-Khintchine representation for these set-indexed Lévy processes and we also characterized this class of Markov processes.

## TOSCA Project-Team

### 3. Scientific Foundations

#### 3.1. Scientific Foundations

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

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

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