

RESEARCH CENTER Bordeaux - Sud-Ouest

FIELD

Activity Report 2017

Section New Results

Edition: 2018-02-19

LFANT Project-Team

6. New Results

6.1. Non commutative number theory

Participant: Jean Paul Cerri.

Pierre Lezowski has studied in [11], Euclidean properties of matrix algebras. He proved that if A is a commutative ring and if n > 1 is an integer, then $M_n(A)$ is right and left Euclidean if and only if A is a principal ideal ring. Moreover, under the hypothesis that the stathm takes integer values, he established that if A is an integral domain, then $M_n(A)$ is ω -stage right and left Euclidean if and only if A is a Bézout ring. He also proved, under the same hypothesis, that if A is a K-Hermite ring, then $M_n(A)$ is (4n - 3)-stage left and right Euclidean, that if A is an elementary divisor ring, then $M_n(A)$ is (2n - 1)-stage left and right Euclidean, and that if A is a principal ideal ring, then $M_n(A)$ is 2-stage right and left Euclidean. In each case, he obtained an explicit algorithm allowing to compute, among other things, right or left gcd in $M_n(A)$.

Jean-Paul Cerri and Pierre Lezowski have generalized in [18], Cerri's algorithm (for the computation of the upper part of the norm-Euclidean spectrum of a number field) to totally definite quaternion fields. This allowed them to establish the exact value of the norm-Euclidean minimum of many orders in totally definite quaternion fields over a quadratic number field. Before this work, nobody knew how to compute the exact value of such a minimum when the base number field has degree > 1. They also proved that the Euclidean minimum and the inhomogeneous minimum of orders in such quaternion fields are always equal and that moreover they are rational under the hypothesis that the base number field is not quadratic, which remains the only open case, as for real number fields.

In [12] Lezowski determines which cyclic field of degree d are norm-Euclidean for d = 5, 7, 19, 31, 37, 43, 47, 59, 67, 71, 73, 79, 97.

6.2. Cryptographic Protocols

Participant: Guilhem Castagnos.

In [15] G. Castagnos, L. Imbert, and F. Laguillaumie revisit a recent cryptographic primitive called *encryption switching protocols* (ESP). This primitive was introduced by Couteau, Peters and Pointcheval last year. It allows to switch ciphertexts between two encryption schemes. If such an ESP is built with two schemes that are respectively additively and multiplicatively homomorphic, it naturally gives rise to a secure 2-party computation protocol. It is thus perfectly suited for evaluating functions, such as multivariate polynomials, given as arithmetic circuits. Couteau et al. built an ESP to switch between Elgamal and Paillier encryptions which do not naturally fit well together. Consequently, they had to design a clever variant of Elgamal over $\mathbf{Z}/n\mathbf{Z}$ with a costly shared decryption.

In this work, Castagnos *et. al.* first present a conceptually simple generic construction for encryption switching protocols. Then, they give an efficient instantiation of our generic approach that uses two well-suited protocols, namely a variant of Elgamal in $\mathbf{Z}/p\mathbf{Z}$ and the Castagnos-Laguillaumie encryption which is additively homomorphic over $\mathbf{Z}/p\mathbf{Z}$. Among other advantages, this allows to perform all computations modulo a prime p instead of an RSA modulus. Overall, this solution leads to significant reductions in the number of rounds as well as the number of bits exchanged by the parties during the interactive protocols. They also show how to extend its security to the malicious setting.

This paper was presented at the CRYPTO Conference 2017, and is part of the ALAMBIC project.

6.3. Algorithmic number theory

Participant: Henri Cohen.

The book [17] by Henri Cohen on *Modular Forms: A Classical Approach* has been published. The theory of modular forms is a fundamental tool used in many areas of mathematics and physics. It is also a very concrete subject in itself and abounds with an amazing number of surprising identities. This comprehensive textbook, gives a complete picture of the classical aspects of the subject, with an emphasis on explicit formulas. Content include: elliptic functions and theta functions, the modular group, its subgroups, and general aspects of holomorphic and nonholomorphic modular forms, with an emphasis on explicit examples. The heart of the book is the classical theory developed by Hecke and continued up to the Atkin–Lehner–Li theory of newforms and including the theory of Eisenstein series, Rankin–Selberg theory, and a more general theory of theta series including the Weil representation. The final chapter also explores in some detail more general types of modular forms such as half-integral weight, Hilbert, Jacobi, Maass, and Siegel modular forms.

The article by Bill Allombert, Jean-Paul Allouche and Michel Mendès France on *Euler's divergent series* and an elementary model in Statistical Physics has been published in Statistical Physics Ars Mathematica Contemporanea. This article study the multiple integral of a multivariate exponential taken with respect either to the Lebesgue measure or to the discrete uniform Bernoulli measure. In the first case the integral is linked to Euler's everywhere divergent power series and its generalizations, while in the second case the integral is linked to a one-dimensional model of spin systems as encountered in physics.

Bill Allombert has worked with Nicolas Brisebarre and Alain Lasjaunias on *a two-valued sequence and related continued fractions in power series fields*. They explicitly describe a noteworthy transcendental continued fraction in the field of power series over \mathbb{Q} , having irrationality measure equal to 3. This continued fraction is a generating function of a particular sequence in the set $\{1, 2\}$.

In the Pari software, K. Belabas and H. Cohen have added an extensive new package mf for modular forms. This package allows to build spaces of classical modular form $M_k(\Gamma_0(N), \chi)$ where $2k \in \mathbb{Z}$ and perform standard tasks like finding bases, splitting the space using Hecke operators and the computation of eigenforms. It also solves important difficult problems: the computation of forms of weight 1, the realization of Shimura lifts as an explicit isomorphism between Kohnen's +-space $S_k^+(\Gamma 0(4N), \chi)$ and $S_{2k-1}(\Gamma_0(N), \chi^2)$ and the Fourier expansion of $f \mid_k \gamma$ for arbitrary f and arbitrary $\gamma \in \operatorname{GL}_2(\mathbb{Q})^+$, which includes as a special case the expansion of f at all cusps (where other modular form packages usualy deal with the expansion at infinity and the cusps reachable via Atkin-Lehner operators, e.g. all cusps in squarefree levels). The latter is especially important as it allows an explicit description of Atkin-Lehner operators, the evaluation of f arbitrary points in the upper-half plane, the computation of period polynomials and Pettersson products, etc.

6.4. Elliptic curve and Abelian varieties cryptology

Participant: Damien Robert.

In [21], E. Milio and D. Robert describe an algorithm to evaluate in quasi-linear time Hilbert modular functions in dimension 2, and also how to recover in time quasi-linear the period matrix from the value of the function. They apply this theory to the modular functions $j(\tau/\beta)$ and $\theta(\tau/\beta)$ where β is a totally real positive number of the quadratic real field corresponding to the Hilbert surface to construct modular polynomials parametrizing cyclic isogenies between principally polarised abelian varieties. This extends the construction of classical modular polynomials but allow to have much smaller polynomials, which allow to compute them up to norm $\ell = 91$ rather than $\ell = 7$ in dimension 2 for classical polynomials.

In [19], Dudeanu, Alina and Jetchev, Dimitar and Robert, Damien and Vuille, Marius describe an algorithm to compute cyclic isogenies from their kernels. This extends the work of [10] from isogenies with maximal isotropic kernels for the Weil pairing to cyclic isogenies, using real multiplication. Such isogenies are indispensable to fully explore the isogeny graph and will be able to speed up a lot of algorithms that needs isogenous curves, like the CRT method for class polynomials.

6.5. Arbitrary-precision ball arithmetic

Participant: Fredrik Johansson.

During the year, F. Johansson has released three new versions (2.10, 2.11 and 2.12) of the Arb software for arbitrary-precision ball arithmetic.

The paper [] describing Arb has been published in the IEEE Transactions on Computers and was selected as the best paper of this journal's Special Issue on Computer Arithmetic. As a result, a video presentation was featured on the journal's website and Johansson was invited to present the paper in a special session at the 24th IEEE Symposium on Computer Arithmetic (ARITH24) at Imperial College London, UK.

In [20], Johansson describes the first complete algorithm for computing the Lambert W function rigorously in complex ball arithmetic.

6.6. Python and Julia computer algebra packages

Participant: Fredrik Johansson.

F. Johansson together with C. Fieker, W. Hart and T. Hofmann of TU Kaiserslautern have developed Nemo and Hecke, two packages for computer algebra and algebraic number theory using the Julia programming language. The paper [16] describing Nemo and Hecke has been published in the proceedings of ISSAC, the main international computer algebra conference.

The paper [14] describing the SymPy package for computer algebra in Python has been published. SymPy is a highly collaborative international project and F. Johansson is one of the 27 coauthors of this paper. Johansson's main contributions to the software include developing the mpmath package used for arbitrary-precision numerical evaluation. In addition, Johansson has issued the stable version 1.0 release of mpmath.

CAGIRE Project-Team

7. New Results

7.1. Development of an accurate and stable finite volume scheme for simulating low Mach number flows with or without acoustic waves

Starting from the Roe scheme and various low Mach fixes for the barotropic Euler PDE's system and using a 2-scale asymptotic analysis of the (semi)-discrete system, a new Roe based scheme is derived whose set of dissipative terms is chosen in order to ensure both accuracy, stability and checkerboard free behaviour. A paper on this topic is being finalized.

7.2. Analysis of liquid sheet flowing under gravity

In the framework of an informal cooperation with Y. Le Guer and K. El Omari who are supervising the PhD thesis of A. Kacem, we have been involved in the experimental and the numerical study of liquid sheets falling under gravity [14]. The various flow regimes have been characterized in terms of the relevant dimensionless numbers (Weber, Reynolds, Ohnesorge) with a particular emphazis on the regimes leading to the appearance fo holes within the liquid sheet. A journal paper has been submitted mid-2017 and is presently under revision.

7.3. First order hyperbolic formulation of dissipative systems

In the framework of the leave of Vincent Perrier at National institute of Aerospace, a general framework for defining first order formulation of nonlinear dissipative systems equipped with an entropy has been developed. The numerical methods for discretizing this type of system are still in development.

7.4. Improvement of turbulent heat flux modelling for buoyant flows

Several modifications were introduced in the Elliptic Blending Differential Flux Model (EB-DFM) to account for the influence of wall blockage on the turbulent heat flux. These modifications are introduced in order to reproduce, in association with the most recent version of the EB-RSM, the full range of regimes, from forced to natural convection, without any case-specific modification. The interest of the new model is demonstrated using analytical arguments, a priori tests and computations in channel flows in the different convection regimes, as well as in a differentially heated cavity. This work is published in *Int. J. Heat Fluid Flow* [10].

7.5. Modelling of turbulent flows with strong variations of the physical properties

The effects of a strong transverse temperature gradient, very common in industrial applications, on a turbulent Poiseuille flow were studied numerically using RANS models, in order to determine the closure level necessary to reproduce the influence of variations of the physical properties, for a wide range of wall-temperature ratios. Eddy-viscosity models prove able to correctly reproduce the asymmetry of the flow and the tendency toward relaminarization close to the hot wall, which are mainly due to the strong variations of the molecular viscosity. Discrepancies in the predictions of the different closure levels only appear for the highest temperature ratios. A journal paper is under revision for publication in *Int. J. Heat Fluid Flow*.

CARDAMOM Project-Team

7. New Results

7.1. High order discretizations on unstructured meshes

- Participants: Héloise Beaugendre, Cécile Dobrzynski, Mario Ricchiuto, Quentin Viville
- Corresponding member: Héloise Beaugendre

A *p*-adaptive continuous residual distribution scheme has been proposed. Under certain conditions, primarily the expression of the total residual on a given element K into residuals on the sub-elements of K and the use of a suitable combination of quadrature formulas, it is possible to change locally the degree of the polynomial approximation of the solution. The discrete solution can then be considered non continuous across the interface of elements of different orders, while the numerical scheme still verifies the hypothesis of the discrete Lax–Wendroff theorem which ensures its convergence to a correct weak solution. The construction of our p-adaptive method has been done in the frame of a continuous residual distribution (RD) scheme. Different test cases for non-linear equations at different flow velocities demonstrate numerically the validity of the theoretical results.

As an evolution, a hp-adaptive RD scheme for the penalized Navier-Stokes equations has also been developed. The method combines hp-adaptation and penalization within a Residual Distribution scheme. The proposed method is an embedded boundary method that provides a simple and accurate treatment of the wall boundary conditions by the technique of penalization and anisotropic mesh adaptation. This method extends the IBM-LS-AUM method to higher order elements and is based on the construction of a p-adaptive RD scheme combined with an anisotropic mesh adaptation method. It has been applied to the resolution of the penalized Navier-Stokes equations. The robustness of the method is showed in practice with numerical experiments for different Mach regimes and Reynolds numbers in dimension two and three.

A novel formulation of residual distributions schemes as well as finite volume schemes on unstructured triangulations in curvilinear coordinates has been proposed within the PhD of L. Arpaia [31], generalising the work of [3]. The simulations reveal that the RD method proposed has very low dissipation when compared to the results existing in literature. This is very encouraging for applications in meteorology. A positivity preserving variant of the method has been used to compute large scale impact and inundation of the 2011 Tohoku tsunami [33], [32].

7.2. High order mesh generation and mesh adaptation

- Participants: Luca Arpaia, Cécile Dobrzynski, Marco Lorini, Mario Ricchiuto
- Corresponding member: Cécile Dobrzynski

This year several new algorithmic improvements have been obtained which will allow to enhance our meshing tools:

- We have enhanced our work on *r*-adaptation techniques for time dependent equations. These techniques are based on mesh deformations obtained by solving continuous differential equations for the local displacements. These equations are controlled by an error monitor. Several improvements have been made. We have proposed a new mixed model to compute the mesh deformations. This model is based on one hand on a Laplacian model and on the other hand on an Elasticity model. It takes advantages of the two approaches: a refined mesh where the solution varies a lot and a smooth gradation of the edges size elsewhere. We have applied this technic to 2d unsteady compressible simulations and we have preliminary results in three dimensions
- Additional work on *r*-adaptation has also involved a simple extension to spherical coordinates, allowing an efficient treatment of inundation caused by large scale tusnami waves [33], [32]

- A novel strategy to solve the finite volume discretization of the unsteady Euler equations within the ALE framework over tetrahedral adaptive grids have been proposed [11]. The volume changes due to local mesh adaptation are treated as continuous deformations of the finite volumes and they are taken into account by adding fictitious numerical fluxes to the governing equation. This peculiar interpretation enables to avoid any explicit interpolation of the solution between different grids and to compute grid velocities so that the GCL is automatically fulfilled also for connectivity changes. The solution on the new grid is obtained through standard ALE techniques, thus preserving the underlying scheme properties, such as conservativeness, stability and monotonicity. The adaptation procedure includes node insertion, node deletion, edge swapping and points relocation and it is exploited both to enhance grid quality after the boundary movement and to modify the grid spacing to increase solution accuracy. We have demonstrated the ability of the method on three-dimensional simulations of steady and unsteady flow fields.
- We extended our technique for generating high order curved meshes to immersed boundary problem. Based on a level-set function, we curved the mesh according to the 0-level-set. Preliminary results in 2d have been performed for compressible simulations
- Initial work on the use of fitting techniques to exactly compute moving shocks has been performed. The benefit of this approach in completely removing all numerical artefacts related to the capturing of the discontinuity, and in recovering the full order of accuracy have been shown for both straight and mildly curved discontinuities [42]

7.3. Uncertainty Quantification and robust design optimization

- Participants: Andrea Cortesi, Pietro Marco Congedo, Nassim Razaaly, Sanson Francois
- Corresponding member: Pietro Marco Congedo

Concerning Uncertainty Quantification techniques, we have worked in three main directions. First, we developed novel techniques for building efficient and low-cost surrogate. In [43], two main points are introduced. Firstly, a technique which couples Universal Kriging with sparse Polynomial Dimensional Decomposition (PDD) to build a metamodel with improved accuracy. The polynomials selected by the adaptive PDD representation are used as a sparse basis to build an Universal Kriging surrogate model. The second is a strategy, derived from anisotropic mesh adaptation, to adaptively add a fixed number of new training points to an existing Design of Experiments. Moreover, we have explored in [44] how active subspaces are used to find a low-dimensional dependence structures in the input-to-output map of the forward numerical solver. Then, surrogate models on the active variables are used to accelerate the forward uncertainty propagation by Monte Carlo sampling and the Markov Chain Monte Carlo sampling of the posterior distribution for Bayesian inversion. Then, the forward and backward methodologies are applied to the simulation of a hypersonic flow around a cylinder, in conditions for which experimental data are available, revealing new insights towards the potential exploitation of heat flux data for freestream rebuilding.

The second action has been oriented towards the development of efficient techniques for computing lowprobability estimations. In [50], we have proposed a novel algorithm permitting to both building an accurate metamodel and to provide a statistically consistent error. In fact, it relies on a novel metamodel building strategy, which aims to refine the limit-state region in all the branches "equally", even in the case of multiple failure regions, with a robust stopping building criterion. Additionally, another importance sampling technique is proposed, permitting to drastically reduce the computational cost when estimating some reference values, or when a very weak failure-probability event should be computed directly from the metamodel.

Third, we have worked on the propagation of uncertainties through systems of solvers [25]. A System of Solvers (SoS) is a set of interdependent solvers where an output of an upstream solver can be the input of downstream solvers. In this work, we restrict ourselves to directed SoS with one-way dependences between solvers. Performing Uncertainty Quantification (UQ) analysis in SoS is challenging because it typically encapsulates a large number of uncertain input parameters and classical UQ methods, such as spectral expansions and Gaussian process models, are affected by the curse of dimensionality. In this work, we develop

an original mathematical framework, based on Gaussian Process (GP) models to construct a global surrogate model of the uncertain SoS, that can be used to solve forward and backward UQ problems. The key idea of the proposed approach is to determine a local GP model for each solver constituting the SoS. These local GP models are built adaptively to satisfy criteria based on the global output error estimation. The error estimate can be decomposed into contributions from the individual GP models, enabling one to select the GP models to refine to efficiently reduce the global error. The framework is first tested on several analytical problems and subsequently applied to space object reentry simulations.

Concerning optimization under uncertainties, we have worked on the formulation of novel framework to perform multi-objective optimization [24], when considering an error on the objective functions. In many engineering optimization problems, the objective functions are affected by an error arising from the model employed for the computation of the functions. For example, in the case of uncertainty-based optimization the objective functions are statistics of a performance of interest which is uncertain due to the variability of the system input variables. These estimated objectives are affected by an error, which can be modeled with a confidence interval. The framework proposed here is general and aims at dealing with any error affecting a given objective function. The strategy is based on the extension of the Bounding-Box concept to the Pareto optima, where the error can be regarded with the abstraction of an interval (in one-dimensional problems) or a Bounding-Box (in multi-dimensional problems) around the estimated value. This allows the computation of an approximated Pareto front, whose accuracy is strongly dependent on the acceptable computational cost. This approach is then supplemented by the construction of an evolutive surrogate model on the objective functions, iteratively refined during the optimization process. This allows ultimately to further reduce the computation cost of the Pareto front with approximations of the objective functions at a negligible cost. Regarding optimization, we have also worked on the formulation of a novel optimization under uncertainty framework for the definition of optimal shapes for morphing airfoils, applied here to advancing/retreating 2D airfoils. In particular, the morphing strategy is conceived with the intent of changing the shape at a given frequency to enhance aerodynamic performance. The optimization of morphing airfoils pre- sented here only takes into account the aerodynamic performance. The paper [5] is then focused on an aerodynamic optimization to set the optimal shape with respect to performance, where technological aspects are inserted through geometrical constraints.

7.4. Modelling of free surface flows

- Participants: Luca Arpaia, Mathieu Colin, Andrea Filippini, Maria Kazolea, Luc Mieussens, and Mario Ricchiuto
- Corresponding member: Mario Ricchiuto

This year we continue our work on fully non-linear weakly dispersive wave models in two dimensional horizontal coordinates. The proposed framework in [92], to approximate the so-called 2D Green-Naghdi equations has been presented in ISOPE (citation) conference an new paper is under preparation.

We also continue our study on wave breaking techniques on BT models [48]. We studied weakly and fully nonlinear models representative of classical and well known models/codes such as BOUSS-2D [85], [86], Funwave [144], [129], Coulwave [109], [131], BOSZ [128], MIKE21 [83], TUCWave [100], [101], and others. We have in particular focused on the enhanced equations of Nwogu [117], and on a frequency enhanced version of the Green-Naghdi system in the form proposed in [69], [93]. We have compared the now popular hybrid closure initially proposed in [136], with an eddy viscosity closure based on an adaptation of the turbulent kinetic energy closure model of [118], modified to be consistent with the detection mechanisms proposed of [101], [93], and also used here. The study performed has involved: a systematic analysis of the behaviour of the two closures for different mesh sizes; the use of dissipation monitors, consistent with the available theory of entropy dissipation for conservation laws [132], [133], to study the dynamics of breaking for several cases; thorough evidence of the equivalent capabilities of the two approaches to provide satisfactory results. Our results indicate that indeed, at least with the (rather standard) implementation proposed here, both closure approaches allow to describe correctly wave transformation and breaking at large scales. We have shown that when using the TKE eddy viscosity closure the numerical dissipation plays a negligible role, which motivates

to look for non-dissipative/energy conserving numerical methods in the future. Also, the results clearly show the reduced sensitivity to the mesh of this approach compared to the hybrid one. The analysis of the wave breaking of solitary waves on a slope also has allowed to quantitatively study the interplay of the dissipation introduced by friction, eddy viscosity, and numerical dissipation. A research paper is under review.

Further more we continue our work for weakly non linear weakly dispersive models, on the transformation breaking and run-up of irregular waves. Its is the first time that an unstructured high-resolution FV numerical solver for the 2D extended BT equations of Nwogu is tested on the generation and propagation of irregular waves. A research paper is under review.

The tools developed have been also used intensively in funded research programs. Within the TANDEM project, several benchmarks relevant to tsunami modelling have been performed and several common publications with the project partners are submitted and/or in preparation [6], [140]. We also our code SLOWS, to study the conditions for tidal bore formation in convergent alluvial estuaries [70]. A new set of dimensionless parameters has been introduced to describe the problem, and the code SLOWS has been used to explore the space of these parameters allowing to determine a critical curve allowing to characterize an estuary as "bore forming" or not. Surprising physical behaviours, in terms of dissipation and nonlinearity of the tides, have been highlighted.

7.5. Wave energy conversion hydrodynamics

- Participants: Umberto Bosi, Mario Ricchiuto
- Corresponding member: Mario Ricchiuto

We have proposed an efficient nonlinear modelling tool for the analysis of wave body interaction based on Boussinesq-type equations. The approach develop here is based on a PDE formulation which model the flow under the body with a depth averaged system featuring an unknown pressure for which a Poisson type problem must solved by appropriately embedding the constraint on the position of the body. The PDE system is discretised by means of a high-order continuous spectral/hp element method in which the coupling between the free surface and floater domains is handled by means of numerical fluxes inspired by techniques used in the discontinuous Galerkin approach. The model is now fully validated both in the hydrostatic case and in the non-hydrostatic one [36], [37]. Several extensions are under way within the MIDWEST project.

7.6. Kinetic modelling of rarefied gases and space reentry

- Participants: Giorgio Martalò, Luc Mieussens, Julien Mathiaud
- Corresponding member: Luc Mieussens

After the end of the post-doc of Giorgio Martalò, a paper has been published [2], in which as presented the derivation of modified boundary conditions for the compressible Navier-Stokes equations to take into account rarefied flow effects. Another paper, related to some numerical aspects, should be submitted soon. Moreover, a paper written by J. Mathiaud and L. Mieussens [7]. This is an extension of their previous work on the modelling of collisions in gases by Fokker-Planck model to polyatomic gases.

Finally, Baranger et. al [53] have presented a way to obtain correct numerical boundary conditions for the approximation of the Boltzmann equation to take into account collisions of gas molecules with solid boundaries. Standard second order finite volume schemes degenerate to first order close to solid walls, but it has been shown in [53] that a suitable use of extrapolation and slope limiters can give second order accuracy. This greatly improves the computation of the heat flux on solid boundaries for atmospheric re-entry flow simulation, for instance.

Finally, the project of numerical and physical modelling of the liquid ablation (for atmospheric re-entry flows) has been concluded by the defense of Simon Peluchon in November 2017. So far, one paper has been published on this subject [8], but at least another one should be submitted soon. Note that Simon Peluchon will be hired in the CEA as a researcher-engineer in January 2018. This subject might induce new collaborations between the CEA and Cardamom, in particular for the use of unstructured grids.

We have developed some activities concerning the application of UQ analysis to aerospace problems. First, we have illustrated in [44] how to perform a Bayesian calibration of the free stream parameters of a hypersonic high-enthalpy flow around a cylinder, exploiting active subspaces for the reduction of the dimensionality of the input space. The configuration taken into account was the HEG I configuration, known in literature as a validation test-case for hypersonic CFD. The goal of the Bayesian inversion was to show the feasibility in using measurements of pressure and heat flux at the stagnation point for rebuilding freestream velocity and density.

Then, we have realized several studies concerning ablation and the characterization of ablative materials. In particular, in [16] and [15], we have illustrated a proof-of-concept of the coupling between a thermochemical ablation model and modern uncertainty quantification techniques with the aim of rebuilding the ablative material tests performed in the inductively coupled Plasmatron facility at the von Karman Institute. Finally, in [14], we have shown how an approach that uses uncertainty quantification methodology can be used in order to rigorously compute error bars on numerically rebuilt values of enthalpy and catalycity from the Plasmatron facility.

7.7. Modelling of icing/de-icing

- Participants: Héloise Beaugendre, Léo Nouveau, Cécile Dobrzynski and Mario Ricchiuto
- Corresponding member: Héloise Beaugendre

The final public workshop of the European STORM project took place at the end of March 2017. The novel high-fidelity approach, based on penalization, proposed by Cardamom to model ice block trajectories has been compared to a low-fidelity approach from Airbus, to two chimera grids approaches from DLR and ONERA and to an experimental database elaborated during the project. The preliminary results are encouraging and commit to further developments of the method.

CQFD Project-Team

7. New Results

7.1. Asymptotically optimal open-loop load balancing

In many distributed computing systems, stochastically arriving jobs need to be assigned to servers with the objective of minimizing waiting times. Many existing dispatching algorithms are basically included in the SQ(d) framework: Upon arrival of a job, $d \ge 2$ servers are contacted uniformly at random to retrieve their state and then the job is routed to a server in the best observed state. One practical issue in this type of algorithm is that server states may not be observable, depending on the underlying architecture. In [3], we investigate the assignment problem in the open-loop setting where no feedback information can flow dynamically from the queues back to the controller, i.e., the queues are unobservable. This is an intractable problem, and unless particular cases are considered, the structure of an optimal policy is not known. Under mild assumptions and in a heavy-traffic many-server limiting regime, our main result proves the optimality of a subset of deterministic and periodic policies within a wide set of (open-loop) policies that can be randomized or deterministic and can be dependent on the arrival process at the controller. The limiting value of the scaled stationary mean waiting time achieved by any policy in our subset provides a simple approximation for the optimal system performance.

Author: J. Anselmi (Inria CQFD).

7.2. The economics of the cloud: price competition and congestion

The work developed in [4] proposes a model to study the interaction of price competition and congestion in the cloud computing marketplace. Specifically, we propose a three-tier market model that captures a marketplace with users purchasing services from Software-as-Service (SaaS) providers, which in turn purchase computing resources from either Provider-as-a-Service (PaaS) providers or Infrastructure-as-a-Service (IaaS) providers. Within each level, we define and characterize competitive equilibria. Further, we use these characterizations to understand the relative profitability of SaaSs and PaaSs/IaaSs, and to understand the impact of price competition on the user experienced performance, i.e., the 'price of anarchy' of the cloud marketplace. Our results highlight that both of these depend fundamentally on the degree to which congestion results from shared or dedicated resources in the cloud.

Authors: J. Anselmi (Inria CQFD), D. Ardagna, J.C.S. Lui, A. Wierman, Y. Xu and Z. Yang.

7.3. A new characterization of the jump rate for piecewise-deterministic Markov processes with discrete transitions

Piecewise-deterministic Markov processes form a general class of non-diffusion stochastic models that involve both deterministic trajectories and random jumps at random times. In [5], we state a new characterization of the jump rate of such a process with discrete transitions. We deduce from this result a nonparametric technique for estimating this feature of interest. We state the uniform convergence in probability of the estimator. The methodology is illustrated on a numerical example.

Authors: A. Genadot (Inria CQFD) and R. Azais.

7.4. Linear minimum mean square filters for Markov jump linear systems

In [9], new linear minimum mean square estimators are introduced by considering a cluster information structure in the filter design. The set of filters constructed in this way can be ordered in a lattice according to the refines of clusters of the Markov chain, including the linear Markovian estimator at one end (with only one cluster) and the Kalman filter at the other hand (with as many clusters as Markov states). The higher is the number of clusters, the heavier are pre-computations and smaller is the estimation error, so that the cluster cardinality allows for a trade-off between performance and computational burden. In this paper we propose the estimator, give the formulas for pre-computation of gains, present some properties, and give an illustrative numerical example.

Authors: E. Costa and B. De Saporta (Inria CQFD).

7.5. Zero-sum discounted reward criterion games for piecewise deterministic Markov processes

In [10], we deal with zero-sum games with a discounted reward criterion for piecewise deterministic Markov process (PDMPs) in general Borel spaces. The two players can act on the jump rate and transition measure of the process, with the decisions being taken just after a jump of the process. The goal of this paper is to derive conditions for the existence of minâmax strategies for the infinite horizon total expected discounted reward function, which is composed of running and boundary parts. The basic idea is, by using the special features of the PDMPs, to re-write the problem via an embedded discrete-time Markov chain associated to the PDMP and re-formulate the problem as a discrete-stage zero sum game problem.

Authors: O. Costa and F. Dufour (Inria CQFD).

7.6. Optimal strategies for impulse control of piecewise deterministic Markov processes

In [11], we deal with the general discounted impulse control problem of a piecewise deterministic Markov process. We investigate a new family of optimal strategies. The construction of such strategies is explicit and only necessitates the previous knowledge of the cost of the no-impulse strategy. In particular, it does not require the resolution of auxiliary optimal stopping problem or the computation of the value function at each point of the state space. This approach is based on the iteration of a single-jump-orintervention operator associated to the piecewise deterministic Markov process.

Authors: B. De Saporta, F. Dufour and A. Geeraert. All authors are members of CQFD at Inria.

7.7. Partially observed optimal stopping problem for discrete-time Markov processes

In [12], we have investigated of a new numerical method to approximate the optimal stopping problem for a discrete-time continuous state space Markov chain under partial observations. It is based on a two-step discretization procedure based on optimal quantization. First, we discretize the state space of the unobserved variable by quantizing an underlying reference measure. Then we jointly discretize the resulting approximate filter and the observation process. We obtain a fully computable approximation of the value function with explicit error bounds for its convergence towards the true value fonction.

Authors: B. De Saporta, F. Dufour and C. Nivot. All authors are members of CQFD at Inria.

7.8. On the stability and the uniform propagation of chaos of a class of extended ensemble Kalman–Bucy filters

The result published in [15] deals with the exponential stability and the uniform propagation of chaos properties of a class of Extended Ensemble Kalman-Bucy filters with respect to the time horizon. This class of nonlinear filters can be interpreted as the conditional expectations of nonlinear McKean Vlasov type diffusions with respect to the observation process. In contrast with more conventional Langevin nonlinear drift type processes, the mean field interaction is encapsulated in the covariance matrix of the diffusion. The main results discussed in the article are quantitative estimates of the exponential stability properties of these nonlinear diffusions. These stability properties are used to derive uniform and non asymptotic estimates of the propagation of chaos properties of Extended Ensemble Kalman filters, including exponential concentration inequalities. To our knowledge these results seem to be the first results of this type for this class of nonlinear ensemble type Kalman-Bucy filters.

Authors: P. Del Moral (Inria CQFD), A. Kurtzmann and J. Tugaut.

7.9. Exponential mixing properties for time inhomogeneous diffusion processes with killing

In [16], we consider an elliptic and time-inhomogeneous diffusion process with time-periodic coefficients evolving in a bounded domain of R d with a smooth boundary. The process is killed when it hits the boundary of the domain (hard killing) or after an exponential time (soft killing) associated with some bounded rate function. The branching particle interpretation of the non absorbed diffusion again behaves as a set of interacting particles evolving in an absorbing medium. Between absorption times, the particles evolve independently one from each other according to the diffusion evolution operator; when a particle is absorbed, another selected particle splits into two offsprings. This article is concerned with the stability properties of these non absorbed processes. Under some classical ellipticity properties on the diffusion process and some mild regularity properties of the hard obstacle boundaries, we prove an uniform exponential strong mixing property of the process conditioned to not be killed. We also provide uniform estimates w.r.t. the time horizon for the interacting particle interpretation of these non-absorbed processes, yielding what seems to be the first result of this type for this class of diffusion processes evolving in soft and hard obstacles, both in homogeneous and non-homogeneous time settings.

Authors: P. Del Moral (Inria CQFD) and D. Villemonais.

7.10. Averaging for some simple constrained Markov processes

In [17], we study a class of piecewise deterministic Markov processes with underlying fast dynamic. Using a " penalty method ", an averaging result is obtained when the underlying dynamic is infinitely accelerated. The features of the averaged process, which is still a piecewise deterministic Markov process, are fully described.

Authors: A. Genadot (Inria CQFD).

7.11. Nonasymptotic analysis of adaptive and annealed Feynman–Kac particle models

Sequential and quantum Monte Carlo methods, as well as genetic type search algorithms can be interpreted as a mean field and interacting particle approximations of Feynman-Kac models in distribution spaces. The performance of these population Monte Carlo algorithms is strongly related to the stability properties of nonlinear Feynman–Kac semigroups. In [18], we analyze these models in terms of Dobrushin ergodic coefficients of the reference Markov transitions and the oscillations of the potential functions. Sufficient conditions for uniform concentration inequalities w.r.t. time are expressed explicitly in terms of these two quantities. We provide an original perturbation analysis that applies to annealed and adaptive Feynman-Kac models, yielding what seems to be the first results of this kind for these types of models. Special attention is

devoted to the particular case of Boltzmann-Gibbs measures' sampling. In this context, we design an explicit way of tuning the number of Markov chain Monte Carlo iterations with temperature schedule. We also design an alternative interacting particle method based on an adaptive strategy to define the temperature increments. The theoretical analysis of the performance of this adaptive model is much more involved as both the potential functions and the reference Markov transitions now depend on the random evolution on the particle model. The nonasymptotic analysis of these complex adaptive models is an open research problem. We initiate this study with the concentration analysis of a simplified adaptive models based on reference Markov transitions that coincide with the limiting quantities, as the number of particles tends to infinity.

Authors: F. Giraud and P. Del Moral (Inria CQFD).

7.12. A comparison of fitness-case sampling methods for genetic programming

Genetic programming (GP) is an evolutionary computation paradigm for automatic program induction. GP has produced impressive results but it still needs to overcome some practical limitations, particularly its high computational cost, overfitting and excessive code growth. Recently, many researchers have proposed fitness-case sampling methods to overcome some of these problems, with mixed results in several limited tests. In [20], we present an extensive comparative study of four fitness-case sampling methods, namely: Interleaved Sampling, Random Interleaved Sampling, Lexicase Selection and Keep-Worst Interleaved Sampling. The algorithms are compared on 11 symbolic regression problems and 11 supervised classification problems, using 10 synthetic benchmarks and 12 real-world data-sets. They are evaluated based on test performance, overfitting and average program size, comparing them with a standard GP search. Comparisons are carried out using non-parametric multigroup tests and post hoc pairwise statistical tests. The experimental results suggest that fitness-case sampling methods are particularly useful for difficult real-world symbolic regression problems, improving performance, reducing overfitting and limiting code growth. On the other hand, it seems that fitness-case sampling cannot improve upon GP performance when considering supervised binary classification.

Authors: Y. Martinez, E. Naredo, L. Trujillo, P. Legrand (Inria CQFD) and U. Lopez.

7.13. Stochastic control of observer trajectories in passive tracking with acoustic signal propagation optimization

In [23], we present a numerical method which computes the optimal trajectory of a underwater vehicle subject to some mission objectives. The method is applied to a submarine whose goal is to best detect one or several targets, or/and to minimize its own detection range perceived by the other targets. The signal considered is acoustic propagation attenuation. Our approach is based on dynamic programming of a finite horizon Markov decision process. A quantization method is applied to fully discretize the problem and allows a numerically tractable solution. Different scenarios are considered. We suppose at first that the position and the velocity of the targets are known and in the second we suppose that they are unknown and estimated by a Kalman type filter in a context of bearings-only tracking.

Authors: H. Zhang (Inria CQFD), B. De Saporta (Inria CQFD), F. Dufour (Inria CQFD), D. Laneuville and A. Nègre.

7.14. Use of local Search in Genetic Programming

There are two important limitations of standard tree-based genetic programming (GP). First, GP tends to evolve unnecessarily large programs, what is referred to as bloat. Second, GP uses inefficient search operators that focus on modifying program syntax. The first problem has been studied in many works, with many bloat control proposals. Regarding the second problem, one approach is to use alternative search operators, for instance geometric semantic operators, to improve convergence. In [36], our goal is to experimentally show that both problems can be effectively addressed by incorporating a local search optimizer as an additional search operator. Using real-world problems, we show that this rather simple strategy can improve the convergence and performance of tree-based GP, while reducing program size. Given these results, a question arises: why are local search strategies so uncommon in GP? A small survey of popular GP libraries suggests to us that local search is underused in GP systems.

Authors: Leonardo Trujillo, Emigdio Z-Flores, Perla S. Juarez Smith, Pierrick Legrand (Inria CQFD), Sara Silva, Mauro Castelli, Leonardo Vanneschi, Oliver Schutze and Luis Munoz.

7.15. Hierarchical clustering with spatial constraints

In [8], we propose a Ward-like hierarchical clustering algorithm including spatial/geographical constraints. Two dissimilarity matrices D_0 and D_1 are inputted, along with a mixing parameter $\alpha \in [0, 1]$. The dissimilarities can be non-Euclidean and the weights of the observations can be non-uniform. The first matrix gives the dissimilarities in the "feature space" and the second matrix gives the dissimilarities in the "constraint space". The criterion minimized at each stage is a convex combination of the homogeneity criterion calculated with D_0 and the homogeneity criterion calculated with D_1 . The idea is then to determine a value of α which increases the spatial contiguity without deteriorating too much the quality of the solution based on the variables of interest i.e. those of the feature space. This procedure is illustrated on a real dataset using the R package ClustGeo.

Authors: Marie Chavent (Inria CQFD), Vanessa Kuentz, Amaury Labenne, Jérôme Saracco (Inria CQFD).

7.16. Variable importance assessment in sliced inverse regression for variable selection

In [19], we are interested in treating the relationship between a dependent variable y and a multivariate covariate x in a semiparametric regression model. Since the purpose of most social, biological, or environmental science research is the explanation, the determination of the importance of the variables is a major concern. It is a way to determine which variables are the most important when predicting y. Sliced inverse regression methods allows to reduce the space of the covariate x by estimating the directions β that form an effective dimension reduction (EDR) space. The aim of this article is to propose a computational method based on importance variable measure (only relying on the EDR space) in order to select the most useful variables. The numerical behavior of this new method, implemented in R, is studied on a simulation study. An illustration on a real data is also provided.

Authors: Ines Jlassi, Jérôme Saracco (Inria CQFD).

7.17. Group-sparse block PCA and explained variance

In [46], we address the simultaneous determination of group-sparse loadings by block optimization, and the correlated problem of defining explained variance for a set of non orthogonal components. We give in both cases a comprehensive mathematical presentation of the problem, which leads to propose i) a new formulation/algorithm for group-sparse block PCA and ii) a framework for the definition of explained variance with the analysis of five definitions. The numerical results i) confirm the superiority of block optimization over deflation for the determination of group-sparse loadings, and the importance of group information when available, and ii) show that ranking of algorithms according to explained variance is essentially independent of the definition of explained variance. These results lead to propose a new optimal variance as the definition of choice for explained variance.

Authors: Guy Chavent, Marie Chavent (Inria CQFD).

7.18. Multivariate Analysis of Mixed Data

In [47], we focus on mixed data that arise when observations are described by a mixture of numerical and categorical variables. The R package PCAmixdata extends standard multivariate analysis methods to incorporate this type of data. The key techniques/methods included in the package are principal component analysis for mixed data (PCAmix), varimax-like orthogonal rotation for PCAmix, and multiple factor analysis for mixed multi-table data. This paper gives a synthetic presentation of the three algorithms with details to help the user understand graphical and numerical outputs of the corresponding R functions. The three main methods are illustrated on a real dataset composed of four data tables characterizing living conditions in different municipalities in the Gironde region of southwest France.

Authors: Marie Chavent (Inria CQFD), Vanessa Kuentz, Amaury Labenne, Jérôme Saracco (Inria CQFD).

7.19. A Smooth Nonparametric Estimator of a Conditional Quantile

In [50], we propose a new smooth nonparametric estimator of conditional quantile of Y for a given value of X using a kernel type of estimators. A numerical study to examine the performance of our estimator as well as a theoretical asymptotic study have been conducted.

Authors: Ines Jlassi, Jérôme Saracco (Inria CQFD).

7.20. Perturbations and projections of Kalman-Bucy semigroups

The purpose of the work published in [40] is to analyse the effect of various perturbations and projections of Kalman-Bucy semigroups and Riccati equations. The original motivation was to understand the behaviour of various regulation methods used in ensemble Kalman filtering (EnKF). For example, covariance inflation-type methods (perturbations) and covariance localisation methods (projections) are commonly used in the EnKF literature to ensure well-posedness of the sample covariance (e.g. sufficient rank) and to 'move' the sample covariance closer (in some sense) to the Riccati flow of the true Kalman filter. In the limit, as the number of samples tends to infinity, these methods drive the sample covariance toward a solution of a perturbed, or projected, version of the standard (Kalman-Bucy) differential Riccati equation. The behaviour of this modified Riccati equation is investigated here. Results concerning continuity (in terms of the perturbations), boundedness, and convergence of the Riccati flow to a limit are given. In terms of the limiting filters, results characterising the error between the perturbed/projected and nominal conditional distributions are given. New projection-type models and ideas are also discussed within the EnKF framework; e.g. projections onto socalled Bose-Mesner algebras. This work is generally important in understanding the limiting bias in both the EnKF empirical mean and covariance when applying regularisation. Finally, we note the perturbation and projection models considered herein are also of interest on their own, and in other applications such as differential games, control of stochastic and jump processes, and robust control theory, etc.

Authors: Pierre Del Moral (Inria CQFD), Adrian Bishop and Sahani Pathiraja.

7.21. Probabilistic Safety Analysis of the Collision Between a Space Debris and a Satellite with an Island Particle Algorithm

Collision between satellites and space debris seldom happens, but the loss of a satellite by collision may have catastrophic consequences both for the satellite mission and for the space environment. To support the decision to trigger o a collision avoidance manoeuver, an adapted tool is the determination of the collision probability between debris and satellite. This probability estimation can be performed with rare event simulation techniques when Monte Carlo techniques are not enough accurate. In this chapter, we focus on analyzing the inuence of dierent simulation parameters (such as the drag coecient) that are set for to simplify the simulation, on the collision probability estimation. A bad estimation of these simulation parameters can strongly modify rare event probability estimations. We design here a new island particle Markov chain Monte Carlo algorithm to determine the parameters that, in case of bad estimation, tend to increase the collision probability value. This algorithm also gives an estimate of the collision probability maximum taking into account the likelihood of the parameters. The principles of this statistical technique are described throughout this chapter.

Authors: Pierre Del Moral (Inria CQFD), Christelle Vergé, Jérôme Morio and Juan Carlos Dolado Pérez.

7.22. Biased online parameter inference for state-space models

We consider Bayesian online static parameter estimation for state-space models. This is a very important problem, but is very computationally challenging as the state-of-the art methods that are exact, often have a computational cost that grows with the time parameter; perhaps the most successful algorithm is that of SM C2 (Chopin et al., J R Stat Soc B 75: 397–426 2013). We present a version of the SM C2 algorithm

which has computational cost that does not grow with the time parameter. In addition, under assumptions, the algorithm is shown to provide consistent estimates of expectations w.r.t. the posterior. However, the cost to achieve this consistency can be exponential in the dimension of the parameter space; if this exponential cost is avoided, typically the algorithm is biased. The bias is investigated from a theoretical perspective and, under assumptions, we find that the bias does not accumulate as the time parameter grows. The algorithm is implemented on several Bayesian statistical models.

Authors: Pierre Del Moral (Inria CQFD), Ajay Jasra and Yan Zhou.

7.23. Multilevel Sequential Monte Carlo Samplers for Normalizing Constants

This work considers the sequential Monte Carlo (SMC) approximation of ratios of normalizing constants associated to posterior distributions which in principle rely on continuum models. Therefore, the Monte Carlo estimation error and the discrete approximation error must be balanced. A multilevel strategy is utilized to substantially reduce the cost to obtain a given error level in the approximation as compared to standard estimators. Two estimators are considered and relative variance bounds are given. The theoretical results are numerically illustrated for the example of identifying a parametrized permeability in an elliptic equation given point-wise observations of the pressure.

Authors: Pierre Del Moral (Inria CQFD), Ajay Jasra, Kody Law and Yan Zhou.

7.24. Multilevel sequential Monte Carlo: Mean square error bounds under verifiable conditions

In this article, we consider the multilevel sequential Monte Carlo (MLSMC) method of Beskos et al. (Stoch. Proc. Appl. [to appear]). This is a technique designed to approximate expectations w.r.t. probability laws associated to a discretization. For instance, in the context of inverse problems, where one discretizes the solution of a partial differential equation. The MLSMC approach is especially useful when independent, coupled sampling is not possible. Beskos et al. show that for MLSMC the computational effort to achieve a given error, can be less than independent sampling. In this article we significantly weaken the assumptions of Beskos et al., extending the proofs to non-compact state-spaces. The assumptions are based upon multiplicative drift conditions as in Kontoyiannis and Meyn (Electron. J. Probab. 10 [2005]: 61–123). The assumptions are verified for an example.

Authors: Pierre Del Moral (Inria CQFD), Ajay Jasra and Kody Law.

7.25. Biased Online Parameter Inference for State-Space Models

We consider Bayesian online static parameter estimation for state-space models. This is a very important problem, but is very computationally challenging as the state-of-the art methods that are exact, often have a computational cost that grows with the time parameter; perhaps the most successful algorithm is that of SM C2 (Chopin et al., J R Stat Soc B 75: 397–426 2013). We present a version of the SM C2 algorithm which has computational cost that does not grow with the time parameter. In addition, under assumptions, the algorithm is shown to provide consistent estimates of expectations w.r.t. the posterior. However, the cost to achieve this consistency can be exponential in the dimension of the parameter space; if this exponential cost is avoided, typically the algorithm is biased. The bias is investigated from a theoretical perspective and, under assumptions, we find that the bias does not accumulate as the time parameter grows. The algorithm is implemented on several Bayesian statistical models.

Authors: Pierre Del Moral (Inria CQFD), Ajay Jasra and Yan Zhou.

7.26. Valuation of Barrier Options using Sequential Monte Carlo

Sequential Monte Carlo (SMC) methods have successfully been used in many applications in engineering, statistics and physics. However, these are seldom used in financial option pricing literature and practice. This paper presents SMC method for pricing barrier options with continuous and discrete monitoring of the barrier condition. Under the SMC method, simulated asset values rejected due to barrier condition are re-sampled from asset samples that do not breach the barrier condition improving the efficiency of the option price estimator; while under the standard Monte Carlo many simulated asset paths can be rejected by the barrier condition making it harder to estimate option price accurately. We compare SMC with the standard Monte Carlo method and demonstrate that the extra effort to implement SMC when compared with the standard Monte Carlo is very little while improvement in price estimate can be significant. Both methods result in unbiased estimators for the price converging to the true value as $1/\sqrt{M}$, where M is the number of simulations (asset paths). However, the variance of SMC estimator is smaller and does not grow with the number of time steps when compared to the standard Monte Carlo. In this paper we demonstrate that SMC can successfully be used for pricing barrier options. SMC can also be used for pricing other exotic options and also for cases with many underlying assets and additional stochastic factors such as stochastic volatility; we provide general formulas and references.

Authors: Pierre Del Moral (Inria CQFD) and Pavel V. Shevchenko.

GEOSTAT Project-Team

7. New Results

7.1. Multifractal desynchronization of the cardiac excitable cell network during atrial fibrillation

Participants: G. Attuel, H. Yahia.

We compute the so-called multifractal spectra using two variants of the wavelet transform modulus maxima method, the moment (partition function) method and the magnitude cumulant method. Application of these methods to long time series recorded in a patient with chronic AF provides quantitative evidence of the multifractal intermittent nature of the electric energy of passing cardiac impulses at low frequencies, i.e. for times (> 0.5s) longer than the mean interbeat. We also report the results of a two-point magnitude correlation analysis which infers the absence of a multiplicative time-scale structure underlying multifractal scaling. The electric energy dynamics looks like a "multifractal white noise" with quadratic (log-normal) multifractal spectra. These observations challenge concepts of functional reentrant circuits in mechanistic theories of AF, still leaving open the role of the autonomic nervous system (ANS). A transition is indeed observed in the computed multifractal spectra which group according to two distinct areas, consistently with the anatomical substrate binding to the CS, namely the left atrial posterior wall, and the ligament of Marshall which is innervated by the ANS. In a companion paper (II. Modeling), we propose a mathematical model of a denervated heart where the kinetics of gap junction conductance alone induces a desynchronization of the myocardial excitable cells, accounting for the multifractal spectra found experimentally in the left atrial posterior wall area.

Publication: G. Attuel *et al.*, Multifractal desynchronization of the cardiac excitable cell network during atrial fibrillation. I. Multifractal analysis of clinical data. Accepted in **Frontiers in Physiology**, publication beginning 2018.

7.2. Super-resolution

Participant: N. Brodu.

N. Brodu, "Super-Resolving Multiresolution Images With Band-Independent Geometry of Multispectral Pixels". IEEE Transactions on Geoscience and Remote Sensing 55(8), p4610-4617, 2017. In this work we introduce a new method to enhance Sentinel-2 images. Sentinel is a leading ESA program in satellite-based environmental monitoring. N. Brodu has included a super-resolution algorithm in the form of a plug-in in the ESA SNAP software. SNAP is a major reference in satellite image processing. Free software (Cecill-C and GPL), APP deposit under way, https://bil.inria.fr/fr/software/view/2984/tab. Publication: [15].

7.3. Surface mixing and biological activity in the Northwest African upwelling

Participants: A. El Aouni, K. Daoudi, H. Yahia.

The aim of this work is to study the horizontal stirring and mixing in different upwelling areas of the Northwest African margin using attracting/repelling Lagrangian coherent structures (LCS) obtained as subsets of hyperstreamline of the Cauchy-Green strain tensor, whose normal repulsion rate is larger than tangential stretch over backward/forward time interval, and their link to the chlorophyll fronts concentrations, based on 10 years satellite data. The temporal variability of surface stirring is compared to the fronts chlorophyll concentration. Two of the four studied areas exhibit negative correlation between mixing and the chlorophyll concentration. The other two regions show similar seasonal variations, nearly coincident maxima and minima, leading to a global positive correlation. These results are compared to other works that make use of Finite Size Lyapunov Exponent (FSLE) whose output is a plot of scalar distributions. Furthermore, we compare the chlorophyll concentrations with both compressing and stretching lines. Results show different regions with different properties. The surface mixing and chlorophyll concentrations is governed by stretching lines in two regions, by compressing lines in one region with, while no different is shown between attracting and repelling LCS in the northenrn region of the studied area.

7.4. Spatio-Temporal Dynamics of Floods

Participant: N. Brodu.

The floods are an annual phenomenon on the Pacific Coast of Ecuador and can become devastating during El Niño years, especially in the Guayas watershed (32,300 km²), the largest drainage basin of the South American western side of the Andes. As limited information on flood extent in this basin is available, this study presents a monitoring of the spatio-temporal dynamics of floods in the Guayas Basin, between 2005 and 2008, using a change detection method applied to ENVISAT ASAR Global Monitoring SAR images acquired at a spatial resolution of 1 km.

Publication: [16].

7.5. Effect of wind stress forcing on ocean dynamics at Air-Sea Interface

Participant: H. Yahia.

At first order, oceanic currents are generated by the balance of Coriolis and pressure gradient (geostrophic current) and the balance of Coriolis and the frictional force dominated by wind stress in the surface ocean (Ekman current). We aim at studying the difference in term of turbulent hydrodynamics carried by the wind forcing at the air-sea interface. We explore the statistical properties of singularity spectra computed from velocity norms and vorticity data, notably in relation with kurtosis information to underline differences in the turbulent regimes associated with both kinds of velocity fields. This study is conducted over 1 year of daily data and demonstrates the differences in terms of turbulent property of wind forcing. Publication: [24].

7.6. Ocean dynamics: frontal activity

Participant: H. Yahia.

A highresolution (1km spatial and daily temporal resolutions) dataset of 11 years (2003 to 2013) remotely sensed SST by MODIS sensor onboard Aqua and Terra satellites has been investigated and compared with coastal numerical model experiments. The detection and characterization fronts with fluctuating amplitudes is achieved through the Singularity Analysis (i.e. the process of calculating the degree of regularity or irregularity of a function at each point in a domain).

Publication: [18].

7.7. Pathological speech processing

Participants: K. Daoudi, G. Li, Q. Robin, F. G. Satsou.

- Small amount of training data in learning robust classifiers for differential diagnosis between progressive supranuclear palsy (PSP) and multiple system atrophy (MSA). We showed that factorial discriminant analysis and logistic regression can lead to such robust classifiers. Moreover, we showed that these models provide good insights on the multivariate variability and (un)correlation of acoustic features, which can facilitate clinical interpretation.
- We investigated the problem of extracting ground thruth of glottal closure instants (GCI) from electroglottographic (EGG) signals of healthy and pathological speakers. We carried out a large experimental study which showed that existing methods are not robust to recording settings and material. We then proposed a method to overcome this problem. On the other hand, this problem highlighted the non robustness of state of the art methods in automatic detection of GCI from speech.
- We made an experimental evaluation of state of the art methods in automatic extraction of the excitation source from voiced speech. To carry out this evaluation, we used a very recent source-filter model of sustained phonations. The results showed that these methods are reliable only in very particular cases and fail in most.
- Matching pursuit (MP), particularly using the Gammatones dictionary, has become a popular tool in sparse representations of speech/audio signals. The classical MP algorithm does not however take into account psychoacoustical aspects of the auditory system. Recently two algorithms, called PAMP and PMP have been introduced in order to select only perceptually relevant atoms during MP decomposition. We compared the performance these two algorithms on few speech sentences. The results showed that PMP, which also has the strong advantage of including an implicit stop criterion, always outperforms PAMP as well as classical MP. We then raised the question of whether the Gammatones dictionary is the best choice when using PMP. We thus compared it to the popular Gabor and damped-Sinusoids dictionaries. The results showed that Gammatones always outperform damped-Sinusoids, and that Gabor yield better reconstruction quality but with higher atoms rate.

Publications: [22], [23], [21], [19].

MEMPHIS Project-Team

7. New Results

7.1. Fluid-structure interaction and a monolithic scheme

Fluid-structure interaction (FSI) problems are still today difficult to solve on the numerical point of view. Memphis team works on the development of a new numerical method for the simulation of these phenomenas. This method relies on a FSI coupling scheme called "monolithic", in which an eulerian hyperelastic model (Mooney-Rivlin) predicts the behaviour of an elastic structure, all of this in the context of an implicit inclusion of the geometry. A 2D axi-symmetric incompressible Navier-Stokes model is used to follow the behaviour of a newtonian fluid, interacting with this elastic body.

With this coupling method, the solid and fluid problems are solve as a unique numerical solver. This approach has already been studied in the Memphis team for compressible fluids. This process seems to be interesting while it competes on the accuracy point of view with the partitioned approches, commonly used in the literature. More over, an eulerian formalism releases us from the constraints related to the tracking of the fluid-structure interface, which remains the key difficulty for lagrangian methods. This implicit consideration is therefore coherent from the perspective of including complexe geometries. In responding to difficulties related to the monolithic scheme, we employ a kind of meshing, particularly adapted to AMR (Adaptative Mesh Refinement). Developed by the OPTIMAD society, the library PABLO offers the ability to build conceptually simple meshes, natively parallel, and convenient to use. The hierarchical cartesian meshes are also particularly adapted to complex geometries.

The fluid-structure interface is followed via a level-set function. This one is transported in time with a 2nd order semi-lagrangian scheme which is volume conservative, and it is frequently reinitialize with a redistanciation algorithm. A linear extrapolation algorithm (Aslam) is besides added as a complement to the elastic model in order to limit the "non physical" effects introduced by the monolithic coupling scheme. Finally, a contact model is employed to model the collision between an elastic solid and a rigid solid which can occur in particular in a cardiac pump based on oscillating membranes.



Figure 8. Two FSI problems. On the left: elastic cylinder colliding a rigid plate; on the right: hyperelastic membrane immersed in a pump geometry, moving thanks to a mechanical oscillating actuator.

7.2. A Local Lubrication Model for Spherical Particles within an Incompressible Navier-Stokes Flow

The lubrication effects are short-range hydrodynamic interactions essential to the suspension of the particles, and are usually underestimated by direct numerical simulations of particle laden flows.

A local lubrication correction model for particle laden flow of spherical solid particles has been presented and validated. Interactions between a particle and an obstacle (another particle or a wall) can be decomposed into three types: long range hydrodynamics, short range hydrodynamics also called lubrication effects, and mechanical solid-solid contacts.

Long range hydrodynamic interaction are fully resolved by the Volume Penalization method (VP). The incompressible Navier-Stokes equations have been discretized in time using a scalar projection method and in space with a fully second order penalty method.

Due to unresolved scales associated with the grid, short range hydrodynamic interactions are only partially captured by the numerical approach. We thus introduce a local lubrication model. This correction is based on asymptotic expansions of analytical solutions of particle-particle or particle-wall interactions, assuming that the flow within the gap between the particle and the obstacle is in the Stokes regime. Lubrication forces and torques are corrected in a neighborhood of the contact point of two interacting particles where lubrication is poorly captured, as long as the normalized gap width ϵ is smaller than a critical length ϵ_{lub} (a model parameter).

Finally, a linear soft-sphere collision model is used for solid-solid contacts. This model, widely used in the literature [Costa15,Izard14], represents mechanical contacts as two spring-dashpot systems connected at the contact point. The model allows stretching the collision time, to avoid computational overhead in the calculation of the collision force, making the method particularly efficient.

Our local lubrication correction model have been validated on several benchmarks. First, we considered a single particle falling onto a wall at various approach velocities. The comparison with experimental results [Harada01,Joseph01] enables us to validate the dominant lubrication component resulting from the squeezing of the fluid in the gap. The lubrication force and the torque created by the shearing of the fluid in the gap have been validated on oblique particle-wall collisions in dry and wet systems proposed by Joseph and Hunt [Joseph04]. Since lubrication corrections are made locally, our lubrication model does not required tabulation and is compatible to non-spherical particles. The model will be tested for polydisperse flow of ellipsoidal particles in future works.

7.3. Incompressible flow schemes on octrees.

The incompressible Navier-Stokes solver was validated in 2D last years thanks to the computation of the order of convergence. This year, a comparison has been done with data from literature. A first test-case was the flow around a 2-D cylinder. On the figure 9 can be seen a comparison between results from the developed solver and data from literature [Ploumhans (2000)].

A second test case was the flow around a Naca0012 airfoil. The figure 10 shows the X-Velocity around the airfoil at Re = 1000 with an angle of attack of 15°. A QuadTree grid has been used as can be seen in figure 11 . The aerodynamic coefficients have been computed for this test-case and have been compared with data from literature [D. F. Kurtulus (2015)]. With $C_D mean = 0.3$ and $C_L mean = 0.6$ the results are in good agreement with $C_D mean = 0.3$ and $C_L mean = 0.7$ from literature data gathered in [D. F. Kurtulus (2015)].

Then, a grid adaptation process has been developed. It allows for exemple to deal with moving bodies and to focus on interesting areas in the computational domain. With user defined criteria, the grid is indeed automatically refined or coarsened. So, this code allows a fast meshing of the computational domain thanks to the penalization approach. An interesting compromise between computational time and accuracy is also reached thanks to the adaptive mesh refinement process. A validation of the adaptive mesh refinement process has been done with a comparison between 2 cases : the case of the flow around a fixed body with an inflow of $1m.s^{-1}$ and the case of a moving body with a velocity of $1m.s^{-1}$ in a fixed flow. It can be seen on figure 12.

An extension of the code to 3-D has been developed and validated. Again, 2 different test-cases has been chosen for the validation. First, the flow around a sphere has been computed at different Reynolds number and a comparison has been done with several data from literature as shown in table 1. The figure 13 shows the X-Velocity of the flow around a sphere at Re = 500 with Octree grid. A LES turbulence model has been implemented with a Vreman subgrid model. So, the second test-case is the flow around a cylinder at



Figure 9. Comparison of drag coefficients between the code developped and data from literature for the flow around a cylinder at Re = 550



Figure 10. X-Velocity around a Naca0012 airfoil at Re = 1000 with an angle of attack of 15°



Figure 11. QuadTree grid around a Naca0012 airfoil



Figure 12. Comparison with drag coeficient obtained with a fixed body in a flow and with a moving body with a velocity of $1m.s^{-1}$ in a fixed flow

Re = 3900 with LES. The wake profile at different positions has been compared with experimental data as can be seen in figure 14.

Re	Present work	[Campregher	[Fornberg	[Fadlun et al.	[Kim et al.	[Subramanian
		(2009)]	(1988)]	(2000)]	(2001)]	(2003)]
300	0.6268	0.675	-	-	0.657	0.653
500	0.5488	0.52	0.4818	0.476	-	0.555

Table 1. Comparison of drag coefficients with data from literature at different Reynolds Number



Figure 13. X-Velocity around a sphere at Re = 500

As the overall aim is to simulate the aeroelastic effects on a wind turbine subjected to gusts, a dynamic beam model with axial, torsional and flexural deformations have been implemented and coupled with the Octree Navier-Stokes solver.

7.4. Validation of NaSCar at higher Reynolds numbers and Aeroelastic coupling

A beam finite element model has been implemented in order to study the dynamic behavior of the wind turbine blade. The structural model is linear and can describe bending, torsion and axial deformation. There is the possibility to take into account some coupling effects between bending-torsion and torsion-axial deformation. The implementation of the structural model has been validated by means of different static and dynamic tests. In **15** the Fast Fourier Transform of the tip deflection history is reported: the frequency of the predicted peaks is in good agreement with the theoretical values.

The structural model has been coupled with two different computational fluid dynamics codes: a cartesian code (NASCAR3d) and an octree code (developed by Claire Taymans during her PhD). The coupling requires to compute the loads for the structural model by performing an integral of the fluid forces on a surface mesh. The surface mesh is updated at each time step according to the displacement of the structure and this allows to update the level set which is used to impose the effects of the body on the fluid.

28



Figure 14. Wake profile at different positions behind a cylinder at Re = 3900 obtained by averaging Velocity after a preliminary simulation



Figure 15. spectrum of the tip deflection

In order to focus the attention on a single blade of the rotor, the inertial terms (centrifugal and Coriolis forces) have been added in both the fluid solver and in the structural model. This makes it possible to perform a preliminary study of the behavior of a single elastic blade by neglecting the interactions between the different blades and the wind turbine's tower (see 16).



Figure 16. velocity vector field and q-criterion vorticity contour around the turbine blade

The turbulent flow around the blade is studied by means of the Vreman Large Eddy Simulation model 0 which has been tested on the flow around a cylinder at Re=3900 and Re=140000. The validation of the model for high reynolds flows required the use of a very fine mesh in order to appropriately simulate turbulent disspation and accurately predict the mean flow field, the results obtained are in good agreement with the experimental data of Cantwell et al 0 , as reported in 17.



Figure 17. Cylinder mean wake velocity profiles, Re=140000

⁰A.W. Vreman, "'An eddy-viscosity subgrid-scale model for turbulent shear flow: Algebraic theory and applications,"' Phys. Fluids 16,10 (2004)

⁰B. Cantwell, D. Coles, "'An experimental study of entrainment and transport in the turbulent near wake of a circular cylinder,"' J. Fluid Mech. (1983), vol. 136, pp. 321-374

In order to extend the capability of the code to high Reynolds number a wall function approach has been implemented following the guidelines of De Tullio⁰. The main idea of this approach is to impose the value of the velocity in the first fluid cells close to the wall by performing a non-linear interpolation based on wall function which represents the velocity distribution in the turbulent boundary layer (see 18).



Figure 18. wall correction

7.5. Thoracic implant

We are interested in the simulation of elastic tissus deformation in order to simulate the skin deformation due to the pose of a thoracic implant. These implants are used to fill the sternum cavity of patients affected by Pectus Excavatum syndrom. As a first step, we simulated the skin deformation with a single layer elastic model from the real bones, skin and implant geometries imported from STL files. The implant geometry has been designed on-demand by Anatomik Modeling. The single layer elastic model representing an underskin implant, has been implemented on an octree grid to easily and automatically refine around the different geometries and keep accuracy.

The results obtained were qualitatively validated by Anatomik Modeling. The implant actually lays on the rib cage, under the muscles. The next step will be then to include a multi-layer elasticity model to take into account the muscles and other biological soft tissus.

Another problem linked to custom made thoracic implants is the extraction of the so-called surgical plan. It is a critical step necessary to design the implant. This plan corresponds to the surface of the rib cage. To extract it, a mass-spring model has been developed and integrated in a software prototype with a graphical interface. The resulting prototype can be used easily from on any rib cage described by a STL file.

⁰De Tullio, M.D. (2006) *Development of an Immersed Boundary method for the solution of the preconditioned Navier-Stokes equations* (PhD thesis)



Figure 19. Left: Automatic refinement around a part of the rib cage. Right: Slice of the signed distance fonction from rib cage and skin with automatic refinement





Figure 20. Left: Skin without implant. Right: Skin simulation with implant under skin

REALOPT Project-Team

7. New Results

7.1. Improving Branch-and-Price Methods

We have made progress on stabilization techniques and math-heuristics that are essential components for generic Branch-and-Price methods.

The convergence of a column generation algorithm can be improved in practice by using stabilization techniques. Smoothing and proximal methods based on penalizing the deviation from the incumbent dual solution have become standards of the domain. Interpreting column generation as cutting plane strategies in the dual problem, we have analyzed [15] the mechanisms on which stabilization relies. In particular, the link is established between smoothing and in-out separation strategies to derive generic convergence properties. For penalty function methods as well as for smoothing, we describe proposals for parameter self-adjusting schemes. Such schemes make initial parameter tuning less of an issue as corrections are made dynamically. Such adjustments also allow to adapt the parameters to the phase of the algorithm. Extensive test reports validate our self-adjusting parameter scheme and highlight their performances. Our results also show that using smoothing in combination with penalty function yields a cumulative effect on convergence speed-ups.

Math heuristics have become an essential component in mixed integer programming (MIP) solvers. Extending MIP based heuristics, we have studied [17] generic procedures to build primal solutions in the context of a branch-and-price approach. As the Dantzig-Wolfe reformulation of a problem is typically tighter than that of the original compact formulation, heuristics based on rounding its linear programing (LP) solution can be more competitive. We focus on the so-called diving methods that used re-optimization after each LP rounding. We explore combination with diversification- intensification paradigms such as Limited Discrepancy Search, sub-MIPing, relaxation induced neighbourhood search, local branching, and strong branching. The dynamic generation of variables inherent to a column generation approach requires specific adaptation of heuristic paradigms. We manage to use simple strategies to get around these technical issues. Our numerical results on generalized assignment, cutting stock, and vertex coloring problems sets new benchmarks, highlighting the performance of diving heuristics as generic procedures in a column generation context and producing better solutions than state-of-the-art specialized heuristics in some cases.

7.2. Aggregation Techniques

We have developed [13] a general solution framework based on aggregation techniques to solve NP-Hard problems that can be formulated as a circulation model with specific side constraints. The size of the extended Mixed Integer Linear Programming formulation is generally pseudo-polynomial. To efficiently solve exactly these large scale models, we propose a new iterative aggregation and disaggregation algorithm. At each iteration, it projects the original model onto an aggregated one, producing an approximate model. The process iterates to refine the current aggregated model until the optimality is proved.

The computational experiments on two hard optimization problems (a variant of the vehicle routing problem and the cutting-stock problem) show that a generic implementation of the proposed framework allows us to outperform previous known methods.

We have applied this aggregation method to reduce the size of column generation (CG) models for covering problems in which the feasible subsets depend on a resource constraint [16]. The aggregation relies on a correlation between the resource consumption of the elements and the corresponding optimal dual values. The resulting aggregated dual model is a restriction of the original one, and it can be rapidly optimized to obtain a feasible dual solution. A primal bound can also be obtained by restricting the set of columns to those saturated by the dual feasible solution obtained by aggregation. The convergence is realized by iterative disaggregation until the gap is closed by the bounds. Computational results show the usefulness of our method for different cutting-stock problems. An important advantage is the fact that it can produce high-quality dual bounds much faster than the traditional Lagrangian bound used in stabilized column generation.

We have developed an algorithm for the exact solution of the Temporal Knapsack Problem [29], [24]. We proposed a dynamic programming formulation for the problem, whose size is exponential in the size of the input data. To cope with the curse of dimensionality, we based our algorithm on the Successive Sublimation Dynamic Programming method. We generalized it to allow more precise aggregation of the state space of the dynamic program. Several application-specific feasibility tests and dominance relations, based on aggregated information, are used to derive an efficient implementation of the method. The algorithms compares favorably with the literature, solving several open instances.

7.3. Revisiting Benders Decomposition & Enhancing the Algorithm

In Benders decomposition approach to mixed integer programs, the optimization is carried in two stages: key first-stage decision variables are optimized using a polyhedral approximation of the full-blown problem projection, then a separation problem expressed in the second-stage variables is solved to check if the current first-stage solution is truly feasible, and otherwise, it produces a violated inequality. Such cutting-plane algorithms suffer from several drawbacks and may have very bad convergence rates. We have reviewed [98] the battery of approaches that have been proposed in the literature to address these drawbacks and to speed-up the algorithm. Our contribution consists in explaining these techniques in simple terms and unified notations, showing that in several cases, different proposals of the literature boil down to the same key ideas. We classify methods into specific initialization mode, stabilization techniques, strategies to select the separation point, and cut generation strategies. We have contributed to enhance convergence of Benders cutting plane algorithm by a mixture of smoothing techniques and proximal approaches. Our numerical benchmarking is still on going [18].

7.4. Routing Problems

Given a directed graph G = (V, A), a cost function c associated with the arcs of A, and a set of precedence constraints $B \subset V \times V$, the Precedence Constrained Asymmetric Traveling Salesman Problem (PCATSP) seeks for a minimum cost Hamiltonian circuit, starting at node 1, and such that for each $(i, j) \in B$, the node i is visited before node j. There are many ways of modelling the ATSP and several for the PCATSP. In [14], [25] we present new formulations for the two problems that can be viewed as resulting from combining precedence variable based formulations with network flow based formulations. Indeed, the former class of formulations permits to integrate linear ordering constraints. The motivating formulation for this work is a complicated and "ugly" formulation that results from the separation of generalized subtour elimination constraints presented. This so called "ugly" formulation exhibits, however, one interesting feature, namely the "disjoint subpaths" property that is further explored to create more complicated formulations that combine two (or three) "disjoint path" network flow based formulations and have a stronger linear programming bound. Some of these stronger formulations are related to the ones presented for the PCATSP and can be viewed as generalizations in the space of the precedence based variables. Several sets of projected inequalities in the space of the arc and precedence variables are obtained by projection from these network flow based formulations. Computational results for the ATSP and PCATSP evaluate the quality of the new models and inequalities.

In [36] we deal with the Minimum Latency Problem (MLP), another variant of the well-known Traveling Salesman Problem in which the objective is to minimize the sum of waiting times of customers. This problem arises in many applications where customer satisfaction is more important than the total time spent by the server. This paper presents a novel branch-and-price algorithm for MLP that strongly relies on new features for the ng-path relaxation, namely: (1) a new labeling algorithm with an enhanced dominance rule named multiple partial label dominance; (2) a generalized definition of ng-sets in terms of arcs, instead of nodes; and (3) a strategy for decreasing ng-set sizes when those sets are being dynamically chosen. Also, other elements of efficient exact algorithms for vehicle routing problems are incorporated into our method, such as reduced cost fixing, dual stabilization, route enumeration and strong branching. Computational experiments over TSPLIB instances are reported, showing that several instances not solved by the current state-of-the-art method can now be solved.

In [37], [31] we consider the Resource Constrained Shortest Path problem arising as a subproblem in stateof-the-art Branch-Cut-and-Price algorithms for vehicle routing problems. We propose a variant of the bidirectional label correcting algorithm in which the labels are stored and extended according to so-called bucket graph. Such organization of labels helps to decrease significantly the number of dominance checks and the running time of the algorithm. We also show how the forward/backward route symmetry can be exploited and how to filter the bucket graph using reduced costs. The proposed algorithm can be especially beneficial for vehicle routing instances with large vehicle capacity and/or with time constraints. Computational experiments were performed on instances from the distance constrained vehicle routing problem, including multi-depot and site-depended variants, on the vehicle routing problem with time windows, and on the "nightmare" instances of the heterogeneous fleet vehicle routing problem. Very significant improvements over the best algorithms in the literature were achieved and many instances could be solved for the first time.

We also considered a family of Vehicle Routing Problem (VRP) variants that generalize the classical Capacitated VRP by taking into account the possibility that vehicles differ by capacity, costs, depot allocation, or even by the subset of customers that they can visit. In [33], [30], [23] we propose a branch-cut-and-price algorithm that adapts advanced features found in the best performing exact algorithms for homogeneous fleet VRPs. The original contributions include: (i) the use of Extended Capacity Cuts, defined over a pseudo-polynomially large extended formulation, together with Rank-1 Cuts, defined over the Set Partitioning Formulation; (ii) the concept of vehicle-type dependent memory for Rank-1 Cuts; and (iii) a new family of lifted Extended Capacity Cuts that takes advantage of the vehicle-type dependent route enumeration. The algorithm was extensively tested in instances of the literature and was shown to be significantly better than previous exact algorithms, finding optimal solutions for many instances with up to 200 customers and also for some larger instances. Several new best solutions were found too.

7.5. Machine Scheduling Problems

In [21] we consider the unrelated parallel machine scheduling problem with setup times to minimize a general objective function. In this work we present a novel exact algorithm that is capable of solving this problem $R|r_j, s_{ij}^k| \sum f_j(C_j)$ and the large class of problems that can be derived as particular cases from it. The proposed algorithm consists of a branch-cut-and-price approach that combines several features such as non-robust cuts, strong branching, reduced cost fixing and dual stabilization. To our knowledge, this is the first exact algorithm for unrelated machines with earliness and/or tardiness criteria that can solve consistently instances with more than 20 jobs. We report improved bounds for instances of problems $R|r_j, s_{ij}^k| \sum w'_j E_j + w_j T_j$ and $R||\sum w'_j E_j + w_j T_j$ with up to 80 and 120 jobs, respectively.

7.6. Scheduling Strategies for Runtime Systems

We consider the design of low cost but guaranteed approximation algorithms in the context of the runtime StarPU in [20]. In High Performance Computing, heterogeneity is now the norm with specialized accelerators like GPUs providing efficient computational power. The added complexity has led to the development of task-based runtime systems, which allow complex computations to be expressed as task graphs, and rely on scheduling algorithms to perform load balancing between all resources of the platforms. Developing good scheduling algorithms , even on a single node, and analyzing them can thus have a very high impact on the performance of current HPC systems. The special case of two types of resources (namely CPUs and GPUs) is of practical interest. HeteroPrio is such an algorithm which has been proposed in the context of fast multipole computations, and then extended to general task graphs with very interesting results. In this paper, we provide a theoretical insight on the performance of HeteroPrio, by proving approximation bounds compared to the optimal schedule in the case where all tasks are independent and for different platform sizes. Interestingly, this shows that spoliation allows to prove approximation ratios for a list scheduling algorithm on two unrelated resources, which is not possible otherwise. We also establish that almost all our bounds are tight. Additionally, we provide an experimental evaluation of HeteroPrio on real task graphs from dense linear algebra computation, which highlights the reasons explaining its good practical performance.

7.7. Matrix Partitioning for Parallel Computing on Heterogeneous Platforms

We consider the combinatorial optimization problem that arises in the context of matrix multiplication in [40]. The problem of partitioning a matrix into a set of sub-matrices has received increased attention recently and is crucial when considering dense linear algebra and kernels with similar communication patterns on heterogeneous platforms. The problem of load balancing and minimizing communication is traditionally reducible to an optimization problem that involves partitioning a square into rectangles. This problem has been proven to be NP-Complete for an arbitrary number of partitions. In this paper, we present recent approaches that relax the restriction that all partitions be rectangles. The first approach uses an original mathematical technique to find the exact optimal partitioning. Due to the complexity of the technique, it has been developed for a small number of partitions only. However, even at a small scale, the optimal partitions found by this approach are often non-rectangular and sometimes non-intuitive. The second approach is the study of approximate partitioning methods by recursive partitioning algorithms. In particular we use the work on optimal partitioning to improve pre-existing algorithms. In this paper we discuss the different perspectives it opens and present two algorithms, SNRPP which is a sqrt(3/2) approximation, and NRPP which is a 2/sqrt(3) approximation. While sub-optimal, this approach works for an arbitrary number of partitions. We use the first exact approach to analyse how close to the known optimal solutions the NRRP algorithm is for small numbers of partitions. In order to validate above approach, we consider in [41] how to allocate data when performing matrix multiplication on a heterogeneous node, with multicores and GPUs. Classical (cyclic) allocations designed for homogeneous settings are not appropriate, but the advent of task-based runtime systems makes it possible to use more general allocations. Previous theoretical work has proposed square and cube partitioning algorithms aimed at minimizing data movement for matrix multiplication. We propose techniques to adapt these continuous square partitionings to allocating discrete tiles of a matrix, and strategies to adapt the static allocation at run-time. We use these techniques in an implementation of Matrix Multiplication based on the StarPU runtime system, and we show through extensive experiments that this implementation allows to consistently obtain a lower communication volume while improving slightly the execution time, compared to standard state-of-the-art dynamic strategies.

7.8. Convergence between HPC and Data Science

We consider the use of replication when scheduling independent identical tasks in [34]. MapReduce is a well-know framework for distributing data-processing computations onto parallel clusters. In MapReduce, a large computation is broken into small tasks that run in parallel on multiple machines, and scales easily to very large clusters of inexpensive commodity computers. Before the Map phase, the original dataset is split into data chunks that are replicated (a constant number of times, usually 3) and distributed randomly onto computing nodes. During the Map phase, local tasks (i.e., tasks whose data chunks are stored locally) are assigned in priority when processors request tasks. In this paper, we provide the first complete theoretical analysis of data locality in the Map phase of MapReduce, and more generally, for bag-of-tasks applications that behave like MapReduce. We prove that if tasks are homogeneous (in terms of processing time), as soon as the replication factor is larger than 2, FindAssignment, a matching based algorithm, achieves a quasi-perfect makespan (i.e., optimal up to an additive constant of one step) using a sophisticated matching algorithm. Above result is proved with high probability when the number of tasks becomes arbitrarily large, and we therefore complement theoretical results with simulations that corroborate them even for small number of tasks. We also show that the matching-based approach leads to an improvement of data locality during the Map phase and therefore decreases the amount of communications needed to achieve perfect makespan, compared to the classical MapReduce greedy approach. In the context of the convergence between HPC and Data Science, we investigate the use of Burst Buffers for HPC applications in [38]. Burst-Buffers are high throughput, small size intermediate storage systems typically based on SSDs or NVRAM that are designed to be used as a potential buffer between the computing nodes of a supercomputer and its main storage system consisting of hard drives. Their purpose is to absorb the bursts of I/O that many HPC applications experience (for example for saving checkpoints or data from intermediate results). In this paper, we propose a probabilistic model for evaluating the performance of Burst-Buffers. From a model of application and a data management strategy, we build a Markov chain based model of the system, that allows to quickly answer issues about dimensioning of the
system: for a given set of applications, and for a given Burst-Buffer size and bandwidth, how often does the buffer overflow? We also provide extensive simulation results to validate our modeling approach.

7.9. Network Design Problems

The delivery of freight from manufacturing platforms to demand zones is often managed through one or more intermediate locations where storing, merging, transshipment and consolidation activities are performed. In [22], we design a Two-Echelon Distribution Network that helps synchronise different flows of product. Under demand uncertainty, our model integrates decisions on the locations and the size of second echelon facilities an decisions on the flows assignment between the echelons, and on delivery routes to serve the demand zones.

7.10. Two-dimensional Guillotine Cutting Problems

The two-dimensional knapsack problem consists in packing a set of small rectangular items into a given large rectangle while maximizing the total reward associated with selected items. In [13], we restrict our attention to packings that emanate from a k-stage guillotine-cut process. We introduce a generic model where a knapsack solution is represented by a flow in a directed acyclic hypergraph. This hypergraph model derives from a forward labeling dynamic programming recursion that enumerates all non-dominated feasible cutting patterns. To reduce the hypergraph size, we make use of further dominance rules and a filtering procedure based on Lagrangian reduced costs fixing of hyperarcs. Our hypergraph model is (incrementally) extended to account for explicit bounds on the number of copies of each item. Our exact forward labeling algorithm is numerically compared to solving the max-cost flow model in the base hyper-graph with side constraints to model production bounds. Benchmarks are reported on instances from the literature and on datasets derived from a real-world application.

Also we consider a variant of two-dimensional guillotine cutting-stock problem that arises when different bills of order (or batches) are considered consecutively. The raw material leftover of the last cutting pattern is not counted as waste as it can be reused for cutting the next batch. The objective is thus to maximize the length of the leftover. In [35], [32] we propose a diving heuristic based on a Dantzig-Wolfe reformulation solved by column generation in which the pricing problem is solved using dynamic programming (DP). This DP generates so-called non-proper columns, *i.e.* cutting patterns that cannot participate in a feasible integer solution of the problem. We show how to adapt the standard diving heuristic to this "non-proper" case while keeping its effectiveness. We also introduce the partial enumeration technique, which is designed to reduce the number of non-proper patterns in the solution space of the dynamic program. This technique helps to strengthen the lower bounds obtained by column generation and improve the quality of solutions found by the diving heuristic. Computational results are reported and compared on classical benchmarks from the literature as well as on new instances inspired from industrial data. According to these results, proposed diving algorithms outperform constructive and evolutionary heuristics.

7.11. On sets avoiding distance 1

In a joint work with C. Bachoc, T. Bellitto and P. Moustrou [39], we consider the maximum density of sets avoiding distance 1 in \mathbb{R}^n . Let ||.|| be a norm of R^n and $G_{||.||}$ be the so-called unit distance graph with the points of R^n as vertex set and for edge set, the set of pairs $\{x, y\}$ such that ||x - y|| = 1. An independent set of $G_{||.||}$ is said to avoid distance 1.

Let $||.||_E$ denote the Euclidean norm. For n = 2, the chromatic number of $G_{||.||_E}$ is still wide open: it is only known that $4 \le \chi \left(G_{||.||_E} \right) \le 7$ (Nelson, Isbell 1950). The *measurable* chromatic number χ_m of the graph $G_{||.||}$ is the minimal number of *measurable* stable sets of $G_{||.||}$ needed to cover all its vertices. Obviously, we have $\chi \left(G_{||.||_E} \right) \le \chi_m \left(G_{||.||_E} \right)$. For $n = 2, 5 \le \chi_m \left(G_{||.||_E} \right)$ (Falconer 1981).

Let $m_1(G_{||.||})$ denote the maximum density of a measurable set avoiding distance 1. We have $\frac{1}{m_1(G_{||.||})} \leq \chi_m(G_{||.||})$. We study the maximum density m_1 for norms defined by polytopes: if P is a centrally symmetric polytope and x is a point of \mathbb{R}^n , $||x||_P$ is the smallest positive real t such that $x \in tP$. Polytope norms include some usual norms such as the L^1 and L^∞ norms.

If P tiles the space by translation, then it is easy to see that $m_1(G_{||.||_P}) \ge \frac{1}{2^n}$. C. Bachoc and S. Robins conjectured that equality always holds. We show that this conjecture is true for n = 2 and for some polytopes in higher dimensions.

7.12. Separating Codes and Traffic Monitoring

The paper [12] studies the problem of traffic monitoring which consists of differentiating a set of walks on a directed graph by placing sensors on as few arcs as possible. The problem of characterising a set of individuals by testing as few attributes as possible is already well-known, but traffic monitoring presents new challenges that the previous models of separation fall short from modelling such as taking into account the multiplicity and order of the arcs in a walk. We introduce a new and stronger model of separation based on languages that generalises the traffic monitoring problem. We study three subproblems with practical applications and develop methods to solve them by combining integer linear programming, separating codes and language theory.

CARMEN Project-Team

7. New Results

7.1. A parameter optimization method to solve the ECG inverse problem

Existing electrocardiographic inverse models express their results either in terms of potentials on the heart surface or in terms of activation times in the heart. G. Ravon developed a new method which gives a potentially more useful answer in terms of three parameters of the underlying action potentials in the heart [27]. Since there are more parameters, care had to be taken to avoid overfitting. Tests on in-silico and ex-vivo data showed good results: the method gave better activation maps than the method of fundamental solutions to which it was compared, and fitted the repolarization phase of the ECG accurately. Figure shows an example of an inversely estimated repolarization map.



Figure 3. Reference repolarization map and inversely estimated maps, using the newly developed method (middle) and the method of fundamental solutions to which it was compared (right). The new method results in more realistic patterns.

7.2. Optimal control to Bidomain-Bath model

This project is concerned with the study of the convergence analysis for an optimal control of a bidomain-bath model. The bidomain-bath model equations describe the cardiac bioelectric activity at the tissue bath volumes where the control acts at the boundary of the tissue domain. In recent work [13] [44], we established the well-posedeness of the direct bidomain-bath model by a discrete Galerkin approach. The convergence proof is based on deriving a series of a priori estimates and using a general L^2 -compactness criterion. Moreover, the well-posedeness of the adjoint problem and the first order necessary optimality conditions are shown. Comparing to the direct problem, the convergence proof of the adjoint problem is based on using a general L^1 -compactness criterion. The numerical tests are demonstrated which achieve the successful cardiac defibrillation by utilizing less total current. Finally, the robustness of the Newton optimization algorithm is presented for different finer mesh geometries.

7.3. Bidomain Calcium Dynamics in Cardiac Cell

In our project [35], we are interested in modeling the interaction of Calcium dynamics in a bidomain medium including Sarcolemma and Sarcoplasmic reticulum. The governing equations consist of a nonlinear reaction-diffusion system representing the various calcium fuxes and theirs buffers in the two medias. A priori stability

bounds and the solvability of the system is analyzed using a fixed-point approach. We introduce a finite element method to numerically solve our model equations. Moreover, we establish existence of discrete solutions and show convergence to a weak solution of the original problem. Finally, we report several 2D and 3D numerical experiments illustrating the behavior of the proposed scheme.

7.4. Cardiac electromechanics with physiological ionic model

This project [37] is concerned with the mathematical analysis of a coupled elliptic-parabolic system modeling the interaction between the propagation of electric potential coupled with general physiological ionic models and subsequent deformation of the cardiac tissue. A prototype system belonging to this class is provided by the electromechanical bidomain model, which is frequently used to study and simulate electrophysiological waves in cardiac tissue. The coupling between muscle contraction, biochemical reactions and electric activity is introduced with a so-called active strain decomposition framework, where the material gradient of deformation is split into an active (electrophysiology-dependent) part and an elastic (passive) one. We prove existence of weak solutions to the underlying coupled electromechanical bidomain model under the assumption of linearized elastic behavior of the updated nonlinear diffusivities. The proof of the existence result is proved by means of a non-degenerate approximation system, the Faedo-Galerkin method, and the compactness method.

7.5. Electrocardiographic lead fields

Currently a monodomain reaction-diffusion model is a well-established method to simulate the electrical activity of the heart [58], [59], even more so because it can be adapted to approximate a bidomain model very closely [46], [49]. Computing the electrocardiogram (ECG) from the results of such models is harder because it requires large linear systems to be solved, and does not scale well to large numbers of processors. A possible solution is to use so-called lead fields, the electrocardiographic term for a linear combination of Green's functions that express the ECG potential as an integral over a field of electric current dipoles. M. Potse has implemented and tested methods to compute and use lead fields for ECG simulation with the Propag code. It turned out that this classical method is practical and sufficiently accurate, and gives a huge scaling advantage on modern highly parallel computers. This result is of practical importance for our applied work, and a journal manuscript on this topic will be submitted in January 2018.

7.6. Rapid localization of arrhythmia

Our pilot studies using model data [25] have shown promising results for a proposed simple and rapid localization method to be used in the catheterization laboratory. We have found that even with only a few ECG electrodes accuracies in the order of millimeters can be achieved for the position of an arrhythmia origin with respect to a catheter position. A journal manuscript on this topic is expected to be complete in February 2018.

7.7. Bilayer model

The rigorous proof that mathematically found the bilayer model from S. Labarthe PhD thesis [56] was actually published in SIAM Journal of Applied Math [15]. Based on sophisticated energy estimates, it proves that the bilayer model is the rigorous limit of the underlying three-dimensional model.

7.8. multi-electrode array measurement

In the context of the CardioXComp project, in collaboration with the REO team and the company Notocord, we proposed a strategy to analyze the signals acquired by multi-electrode array (MEA) on cultures of hiPSC-CMs cells with drug compounds, and to automatically deduce the channels affected by the drug. First, in [10], we study how MEA meausrement can be modeled in such a way that the produced in-silico signals are comparable to real ones. The main problems concern the heterogenities of cell cultures. Then, in [20] a method based on parameter identification, by comparing the in-silico and real signals, was used on signals acquired with commercial systems on cell culture with various drugs. The IC-50 and dose-response of several drugs could be assessed. This kind of techniques could contribute to promote the technology based on MEA and hiPSC-CMs.

7.9. High-order integration methods for ion channel models

On November 15, 2017, C. Douanla Lontsi defended his PhD thesis [9] on the numerical analysis of timestepping methods for the cardiac monodomain equations. A huge amount of work was carried out in the thesis. The thesis builds on the seminal Rush-Larsen technique [62], [57], and the recent novel computational interest in exponential integrator methods. Two new exponential methods of arbitrarily high order are proposed (EABk and RLk). Most notably, Rush-Larsen techniques of order k = 2, 3, 4 were entirely explicited. The theory was adapted to analyse these methods and convergence proofs were derived. The complete Dahlquist stability region of these methods was documented. Finally, the methods were integrated into an IMEX strategy to solve the monodomain equation in 1D to 3D problems, with two ionic models (BR and TNNP). The results essentially show that order at least 3 is required to leas to reasonably accurate simulations. Three journal papers were submitted in 2017, [39], [41], [40].

7.10. High-order finite-volume discretizations

Y. Coudière and R. Turpault proposed new simple and efficient high-order finite volume discretizations to be applied to the monodomain equation. They showed how the method can be easily implemented up to the order 6, with very good results, also for simulation of complex propagation patterns. The results were published in [16].

7.11. Identification of multiple space dependent ionic parameters in cardiac electrophysiology modelling

In this paper, we consider the inverse problem of space dependent multiple ionic parameters identification in cardiac electrophysiology modelling from a set of observations. We use the monodomain system known as a state-of-the-art model in cardiac electrophysiology and we consider a general Hodgkin-Huxley formalism to describe the ionic exchanges at the microscopic level. This formalism covers many physiological transmembrane potential models including those in cardiac electrophysiology. Our main result is the proof of the uniqueness and a Lipschitz stability estimate of ion channels conductance parameters based on some observations on an arbitrary subdomain.

MAGIQUE-3D Project-Team

6. New Results

6.1. Seismic Imaging and Inverse Problems

6.1.1. Mathematical determination of the Fréchet derivative with respect to the domain for a fluid-structure scattering problem. Case of polygonal-shaped domains.

Participants: Hélène Barucq, Elodie Estecahandy.

The characterization of the Fréchet derivative of the elasto-acoustic scattered field with respect to Lipschitz continuous polygonal domains is established. The considered class of do- mains is of practical interest since two-dimensional scatterers are always transformed into polygonal- shaped domains when employing finite element methods for solving direct and inverse scattering problems. The obtained result indicates that the Fr´echet derivative with respect to the scatterer of the scattered field is the solution of the same elasto-acoustic scattering problem but with addi- tional right-hand side terms in the transmission conditions across the fluid-structure interface. This characterization has the potential to advance the state-of-the-art of the solution of inverse obstacle problems.

This work has been done in collaboration with Prof. Rabia Djellouli (California State University at Northridge) and has been accepted for publication in Siam Journal of Applied Mathematics [16].

6.1.2. Shape-reconstruction and parameter identification of an elastic object immersed in a fluid

Participants: Izar Azpiroz Iragorri, Hélène Barucq, Julien Diaz, Kevin Lagnoux.

We have developed a procedure to reconstruct the shape and material parameters of an elastic obstacle immersed in a fluid medium from some external measurements given by the so called far-field pattern. It is a nonlinear and ill-posed problem which is solved by applying a Newton-like iterative method involving the Fréchet derivatives of the scattered field. These derivatives express the sensitivity of the scattered field with respect to the parameters of interest. They are defined as the solution of boundary value problems which differ from the direct one only at the right-hand sides level. We have been able to establish the well-posedness of each problem in the case of a regular obstacle and it would be interesting in the near future to extend those results to the case of scatterers with polygonal boundaries. It requires to work with less regular Sobolev spaces for which the definition of traces is not obvious. We have also provided an analytical representation of the Fréchet derivatives in the case of a circle. This provide a way of validating the numerical experiments and it would be interesting to obtain their expression in the case of elliptical scatterers or spherical ones. It is worth mentioning that this work has been done only in the case of isotropic media and it would be interesting to extend field in anisotropic media as well. It requires to establish analytic representations of the scattered field in anisotropic media which is more difficult because it involves more parameters.

We have studied the response of the data to the different parameters. It turns out that the sensitivity of the far field pattern is very different regarding the shape or the material parameters. We have delivered a sensitivity analysis which has been essential for understanding that the reconstruction of the material parameters is conditioned by the recovering of the shape parameters. This makes the full reconstruction very difficult and sometimes unstable. In particular, in the case of a disk-shaped obstacle, when addressing the role of the frequency in the reconstruction, we have been faced to the issue of the existence of Jones modes which had been already observed by Elodie Estecahandy in her PhD thesis. Next, we have introduced a series of numerical experiments that have been performed by applying two algorithms which propose two strategies of full reconstruction regarding the material parameters are retrieved simultaneously with the shape or not. It turns out that both work similarly delivering the same level of accuracy but the simultaneously. Since realistic configurations include noisy data, we have performed some simulations for the reconstruction of the shape along with the Lamé coefficients for different noise levels. Other interesting experiments have been carried out using a multistage procedure where the parameters of interest are the density of the solid interior, the shape of the obstacle and its position. We have considered the case of Limited Aperture Data in back-scattering configurations, using multiple incident plane waves, mimicing a physical disposal of non-destructive testing. This is an encouraging ongoing work which deserves to be completed by considering a wide range of examples including more general geometries of the scatterer. It should also be extended by dealing with limited aperture data using only one incident wave (which will probably require multiple frequency data).

These results have been obtained in collaboration with Rabia Djellouli (California State University at Northridge, USA) and were presented to the Waves 2017 conference.

6.1.3. Shape-reconstruction and parameter identification of an anisotropic elastic object immersed in a fluid

Participants: Izar Azpiroz Iragorri, Hélène Barucq, Julien Diaz.

We extended the solution methodology for reconstructing the shape and material parameters of an elastic obstacle (see 6.1.2) to the case of anisotropic media. This is a very challenging case which still deserves further works. We have obtained some results but since the impact of some of the anisotropic parameters on the FFP is even weaker than the Lamé coefficients, the reconstruction of these parameters together with the shape parameters requires several frequencies and carefully adapted regularization parameters. It is in particular difficult to retrieve the Thomsen parameters ϵ and δ because their reconstruction requires to have an accurate adjustment on the rest of material and shape parameters. The recovery process is thus computationally intensive and some efforts should be done in the near future to decrease the computational costs. We were able to recover all the anisotropic parameters when the shape were assumed to be known. However, when trying to recover both shape and material parameters, we could only recover the shape and some of the physical parameters (namely the three most important ones : the density and the two velocities V_p and V_s). We should now find a way to determine all the Thomsen parameters together with the shape. Then, we will have to deal with more complex media such as TTI media (this will add the angle of anisotropy as additional parameter). The last step will be to consider general anisotropy, which could be done by recovering each element of the elastic stiffness tensor. This is simple to implement, since the derivative of the stiffness tensor with respect to one of its component is easily computable (it is a tensor composed of zeroes and ones). However, the stability of the reconstruction is not guaranteed, since we will strongly increase the number of components to be retrieved.

These results have been obtained in collaboration with Rabia Djellouli (California State University at Northridge, USA).

6.1.4. Mathematical analysis and solution methodology for a class of inverse spectral problems arising in the design of optical waveguides

Participant: Hélène Barucq.

We analyze mathematically the problem of determining refractive index profiles from some desired/measured guided waves propagating in optical fibers. We establish the uniqueness of the solution of this inverse spectral problem assuming that only one guided mode is known. Then, we propose an iterative computational procedure for solving numerically the considered inverse spectral problem. Numerical results are presented to illustrate the potential of the proposed regularized Newton algorithm to efficiently and accurately retrieve the refractive index profiles even when the guided mode measurements are highly noisy.

This work has been submitted for publication in a peer-reviewed journal. It has been done in collaboration with Rabia Djellouli (California State University at Northridge, USA) and Chokri Bekkey (University of Monastir, Tunisia)

6.1.5. Time-harmonic seismic inverse problem with Cauchy data

Participant: Florian Faucher.

This work is a collaboration with Giovanni Alessandrini (Università di Trieste), Maarten V. de Hoop (Rice University), Romina Gaburro (University of Limerick) and Eva Sincich (Università di Trieste).

We study the performance of Full Waveform Inversion (FWI) from time-harmonic Cauchy data via conditional well-posedness driven iterative regularization. The Cauchy data can be obtained with dual sensors measuring the pressure and the normal velocity. We define a novel misfit functional which, adapted to the Cauchy data, allows the independent location of experimental and computational sources. The conditional well-posedness is obtained for a hierarchy of subspaces in which the inverse problem with partial data is Lipschitz stable. Here, these subspaces yield piecewise linear representations of the wave speed on given domain partitions. Domain partitions can be adaptively obtained through segmentation of the gradient. The domain partitions can be taken as a coarsening of an unstructured tetrahedral mesh associated with a finite element discretization of the Helmholtz equation. We illustrate the effectiveness of the iterative regularization through computational experiments with data in dimension three. In comparison with earlier work, the Cauchy data do not suffer from eigenfrequencies in the configurations.

The resulting paper is [47] and is also connected to the following conference presentations, [36], [27].

6.1.6. Quantitative Convergence of Full Waveform Inversion in the Frequency Domain

Participants: Hélène Barucq, Florian Faucher.

This work is a collaboration with Guy Chavent (Inria Rocquencourt).

We study the convergence of the inverse problem associated with the frequency domain wave equations for the recovery of subsurface parameters. The numerical method selected for the resolution is the Full Waveform Inversion (FWI), which designs an iterative minimization algorithm. We study the convergence of the scheme in the context of least squares minimization. We establish numerical estimates based on the Fréchet derivatives for the radius of curvature and the deflection. We quantify the (complex) frequency progression to select to foster the convergence, and illustrate the effect of the subsurface geometry. From the curvature estimates, we also provide an insight of the robustness with noise depending on the situation. We supplement the numerical analysis with numerical experiments to demonstrate the results.

The results have been presented in the following conference, [36], [27], [26], [25].

6.1.7. Contributions to seismic full waveform inversion for time harmonic wave equations: stability estimates, convergence analysis, numerical experiments involving large scale optimization algorithms

Participants: Hélène Barucq, Florian Faucher.

In this project, we investigate the recovery of subsurface Earth parameters. We consider the seismic imaging as a large scale iterative minimization problem, and deploy the Full Waveform Inversion (FWI) method. The reconstruction is based on the wave equations because the characteristics of the measurements indicate the nature of the medium in which the waves propagate. First, the natural heterogeneity and anisotropy of the Earth require numerical methods that are adapted and efficient to solve the wave propagation problem. In this study, we have decided to work with the harmonic formulation, i.e., in the frequency domain.

The inverse problem is then established in order to frame the seismic imaging. It is a nonlinear and illposed inverse problem by nature, due to the limited available data, and the complexity of the subsurface characterization. However, we obtain a conditional Lipschitz-type stability in the case of piecewise constant model representation. We derive the lower and upper bound for the underlying stability constant, which allows us to quantify the stability with frequency and scale. It is of great use for the underlying optimization algorithm involved to solve the seismic problem. We review the foundations of iterative optimization techniques and provide the different methods that we have used in this project. The Newton method, due to the numerical cost of inverting the Hessian, may not always be accessible. We propose some comparisons to identify the benefits of using the Hessian, in order to study what would be an appropriate procedure regarding the accuracy and time. We study the convergence of the iterative minimization method, depending on different aspects such as the geometry of the subsurface, the frequency, and the parametrization. In particular, we quantify the frequency progression, from the point of view of optimization, by showing how the size of the basin of attraction evolves with frequency. Following the convergence and stability analysis of the problem, the iterative minimization algorithm is conducted via a multi-level scheme where frequency and scale progress simultaneously. We perform a collection of experiments, including acoustic and elastic media, in two and three dimensions. The perspectives of attenuation and anisotropic reconstructions are also introduced.

6.1.8. Quantitative localization of small obstacles with single-layer potential fast solvers

Participants: Hélène Barucq, Florian Faucher, Ha Pham.

In this work, we numerically study the inverse problem of locating small circular obstacles in a homogeneous medium using noisy backscattered data collected at several frequencies. The main novelty of our work is the implementation of a single-layer potential based fast solver (called FSSL) in a Full-Waveform inversion procedure, to give high quality reconstruction with low-time cost. The efficiency of FSSL was studied in our previous works. We show reconstruction results with up to 12 obstacles in structured or random configurations with several initial guesses, all allowed to be far and different in nature from the target. This last assumption is not expected in results using nonlinear optimization schemes in general. For results with 6 obstacles, we also investigate several optimization methods, comparing between nonlinear gradient descent and quasi-Newton, as well as their convergence with different line search algorithms.

The resulting research report is [45].

6.2. Mathematical modeling of multi-physics involving wave equations

6.2.1. Atmospheric Radiation Boundary Conditions for the Helmholtz Equation

Participants: Hélène Barucq, Juliette Chabassier, Marc Duruflé.

An article is to be published in M2AN, see [14]. This work offers some contributions to the numerical study of acoustic waves propagating in the Sun and its atmosphere. The main goal is to provide boundary conditions for outgoing waves in the solar atmosphere where it is assumed that the sound speed is constant and the density decays exponentially with radius. Outgoing waves are governed by a Dirichlet-to-Neumann map which is obtained from the factorization of the Helmholtz equation expressed in spherical coordinates. For the purpose of extending the outgoing wave equation to axisymmetric or 3D cases, different approximations are implemented by using the frequency and/or the angle of incidence as parameters of interest. This results in boundary conditions called Atmospheric Radiation Boundary Conditions (ARBC) which are tested in ideal and realistic configurations. These ARBCs deliver accurate results and reduce the computational burden by a factor of two in helioseismology applications. This work has been done in collaboration with Laurent Gizon and Michael Leguèbe (Max-Planck-Institut für Sonnensystemforschung, Gottingen, Germany).

6.2.2. Atmospheric radiation boundary conditions for high frequency waves in time-distance helioseismology

Participants: Hélène Barucq, Juliette Chabassier, Marc Duruflé.

An article has been published in Astronomy and Astrophysics [22]. The temporal covariance between seismic waves measured at two locations on the solar surface is the fundamental observable in time-distance helioseismology. Above the acoustic cutoff frequency (5.3 mHz), waves are not trapped in the solar interior and the covariance function can be used to probe the upper atmosphere. We wish to implement appropriate radiative boundary conditions for computing the propagation of high-frequency waves in the solar atmosphere. We consider the radiative boundary conditions recently developed by Barucq et al. (2017) for atmospheres in which sound-speed is constant and density decreases exponentially with radius. We compute the cross-covariance function using a finite element method in spherical geometry and in the frequency domain. The ratio between first-and second-skip amplitudes in the time-distance diagram is used as a diagnostic to compare boundary conditions and to compare with observations. We find that a boundary condition applied 500 km above the photosphere and derived under the approximation of small angles of incidence accurately reproduces the 'infinite atmosphere' solution for high-frequency waves. When the radiative boundary condition is applied 2 Mm above the photosphere, we find that the choice of atmospheric model affects the time-distance diagram.

In particular, the time-distance diagram exhibits double-ridge structure when using a VAL atmospheric model. This is a collaboration with Damien Fournier, Laurent Gizon, Chris Hanson and Michael Leguèbe (Max-Planck-Institut für Sonnensystemforschung, Gottingen, Germany).

6.2.3. Computational helioseismology in the frequency domain: acoustic waves in axisymmetric solar models with flows

Participants: Hélène Barucq, Juliette Chabassier, Marc Duruflé.

An article has been published in Astronomy and Astrophysics [23]. Context. Local helioseismology has so far relied on semi-analytical methods to compute the spatial sensitivity of wave travel times to perturbations in the solar interior. These methods are cumbersome and lack flexibility. Aims. Here we propose a convenient framework for numerically solving the forward problem of time-distance helioseismology in the frequency domain. The fundamental quantity to be computed is the cross-covariance of the seismic wavefield. Methods. We choose sources of wave excitation that enable us to relate the cross-covariance of the oscillations to the Green's function in a straightforward manner. We illustrate the method by considering the 3D acoustic wave equation in an axisymmetric reference solar model, ignoring the effects of gravity on the waves. The symmetry of the background model around the rotation axis implies that the Green's function can be written as a sum of longitudinal Fourier modes, leading to a set of independent 2D problems. We use a high-order finite-element method to solve the 2D wave equation in frequency space. The computation is embarrassingly parallel, with each frequency and each azimuthal order solved independently on a computer cluster. Results. We compute travel-time sensitivity kernels in spherical geometry for flows, sound speed, and density perturbations under the first Born approximation. Convergence tests show that travel times can be computed with a numerical precision better than one millisecond, as required by the most precise travel-time measurements. Conclusions. The method presented here is computationally efficient and will be used to interpret travel-time measurements in order to infer, e.g., the large-scale meridional flow in the solar convection zone. It allows the implementation of (full-waveform) iterative inversions, whereby the axisymmetric background model is updated at each iteration. This work is a collaboration with Aaron Birch, Damien Fournier, Laurent Gizon, Chris Hanson, Michael Leguèbe and Emanuele Papini(Max-Planck-Institut für Sonnensystemforschung, Gottingen, Germany) and with Thorsten Hohage (Gôttingen University, Germany).

6.2.4. The virtual workshop : towards versatile optimal design of musical wind instruments for the makers

Participants: Juliette Chabassier, Robin Tournemenne.

Our project aims at proposing optimization solutions for wind instrument making. Our approach is based on a strong interaction with makers and players, aiming at defining interesting criteria to optimize from their point of view. After having quantified those criteria under the form of a cost function and a design parameters space, we wish to implement state-of-the-art numerical methods (finite elements, full waveform inversion, neuronal networks, diverse optimization techniques...) that are versatile (in terms of models, formulations, couplings...) in order to solve the optimization problem. More precisely, we wish to take advantage of the fact that sound waves in musical instruments satisfy the laws of acoustics in pipes (PDE), which gives us access to the full waveform inversion technique, usable in harmonic or temporal regime. The methods that we want to use are attractive because the weekly depend on the chosen criterion, and they are easily adaptable to various physical situations (multimodal decomposition in the pipe, coupling with the embouchure, ...), which can therefore be modified a posteriori. The goal is to proceed iteratively between instrument making and optimal design (the virtual workshop) in order to get close to tone quality related and playability criteria.

6.2.5. Energy based model and simulation in the time domain of linear acoustic waves in a radiating pipe

Participants: Juliette Chabassier, Robin Tournemenne.

We model in the time domain linear acoustic waves in a radiating pipe without daming. The acoustic equations system in formulated in flow and pressure, which leads to a first order space time equations system. The radiation condition is also written as a first order in time equation, and is parametrized by two real coefficients. Moreover, an auxiliary variable is introduced at the radiating boundary. The choice of this variable is adapted to the considered source type in order to ensure the model stability by energy techniques, under some conditions on the radiating condition. We then propose a stable space time explicit discretization, which ensures the dissipation of a discrete energy. The novelty of the discretization lies, on the one hand, in the variational nature of the space approximation (which leads to artibrary order finite elements with no required matrix inversion), and on the other hand, on the definition of the auxiliary variable for any acoustic source type (which leads to the decay of a well defined energy). Finaly, we quantify the frequential domain of validity of the used radiation condition by comparison with theoretical and experimental models of the litterature. This is a collaboration with Morgane Bergot (Université Claude Bernard, Lyon 1).

6.2.6. Computation of the entry impedance of a dissipative radiating pipe

Participants: Juliette Chabassier, Robin Tournemenne.

Modeling the entry impedance of wind instruments pipes is essential for sound synthesis or instrument qualification. We study this modeling with the finite elements method in one dimension (FEM1D) and with the more classically used transfer matrix method (TMM). The TMM gives an analytical formula of the entry impedance depending on the bore (intern geometry of the instrument) defined as a concatenation of simple elements (cylinders, cones, etc). The FEM1D gives the entry impedance for any instrument geometry. The main goals of this work are to assess the viability of the FEM1D and to study the approximations necessary for the TMM in dissipative pipes. First, lossless Weber's equation in one dimension is studied with arbitrary radiation conditions. In this context and for cylinders or cones, the TMM is exact. We verify that the error made with FEM1D for fine enough elements is as small as desired. When we consider viscothermal losses, the TMM does not solve the classical Kirchhoff model because two terms are supposed constant. In order to overcome this model approximation, simple elements, on which are based the TMM, are decomposed into much smaller elements. The FEM1D does not necessitate any model approximation, and it is possible to show that it solves the dissipative equation with any arbitrarily small error. With this in hand, we can quantify the TMM model approximation error.

6.2.7. Hybrid discontinuous finite element approximation for the elasto-acoustics.

Participants: Hélène Barucq, Julien Diaz, Elvira Shishenina.

Discontinuous Finite Element Methods (DG FEM) have proven their numerical accuracy and flexibility. However, numerically speaking, the high number of degrees of freedom required for computation makes them more expensive, compared to the standard techniques with continuous approximation. Among the different variational approaches to solve boundary value problems there exists a distinct family of methods, based on the use of trial functions in the form of exact solutions of the governing equations. The idea was first proposed by Trefftz in 1926, and since then it has been largely developed and generalized. By its definition, Trefftz-DG methods reduce numerical cost, since the vari- ational formulation contains the surface integrals only. Thus, it makes possible exploration of the meshes with different geometry, in order to create more realistic application. Trefftz-type approaches have been widely used for time-harmonic problems, while their implementation is still limited in time domain. The particularity of Trefftz-DG methods applied to the time-dependent formulations consists in the use of space-time meshes. Even though it creates another computational difficulty, due to a dense form of the matrix, which represents the global linear system, the inversion of the full "space-time" matrix can be reduced to the inversion of one block-diagonal matrix, which corresponds to the interactions in time. In the present work, we develop a theory for solving the coupled elasto-acoustic wave propagation system. We study well-posedness of the problem, based on the error estimates in mesh-dependent norms. We consider a space-time polynomial basis for numerical discretization. The obtained numerical results are validated with analytical solutions. Regarding the advantages of the method, following properties have been proven by the numerical tests: high flexibility in the choice of basis functions, better order of convergence, low dispersion. These results have been obtained in collaboration with Henri Calandra (TOTAL) and have been published in a research report [43]. A paper has been submitted and a second one is being prepared.

6.2.8. Construction of stabilized high-order hybrid Galerkin schemes.

Participants: Hélène Barucq, Aurélien Citrain, Julien Diaz.

We have compared the perfomance of Discontinuous Galerkin Methods and Spectral Element Methods on academic benchmark and on realistic geophysical model in two dimensions. We have shown that, for a given accuracy, SEm on quadrilateral meshes could be 10 times faster than DGm, which justifies our strategy to consider SEm wherever it is possible to use quadrilateral/hexahedral cells. These first results have been presented in Matthias conference. Then, we have considered the SEM/DG coupling proposed for electromagnetics in [78] and we have implemented it in our acoustics code. We are now analyzing the performance of this strategy and we are extending it to deal with elastodynamic and elasto-acoustic coupling. The following steps will be the extension of the analysis to 3D dimensional problems and the application to realistic test case. The main bottleneck is obviously to the definition of an efficient strategy to couple tetrahedra and hexahedra. Indeed, if in the 2D case, the edges of both triangles and quadrilaterals are all segments the faces of tetrahedra are triangle while the faces of hexahedra are quadrilaterals. Hence, in 2D it sufficed to define integration on segment, while in 3D it will be necessary to consider integration of various polygon resulting of the intersection of triangle and quadrilaterals. Once this strategy is defined and implemented, we expect to be able to reduce the computational of the plateform that we develop jointly with Total by a factor between 5 and 10. These results have been obtained in collaboration with Henri Calandra (TOTAL) and Christian Gout (INSA Rouen).

6.2.9. Modeling of dissipative porous media.

Participants: Juliette Chabassier, Julien Diaz, Fatima Jabiri.

In this work we have considered the modeling of 1D acoustic wave propagation coupled with visco-thermical losses that occur in porous media. We have proposed a family of dissipative models from which we have been able to obtain a quasi-constant quality factor (which is an indicator of the dissipation as a function of the frequency). We have derived stability conditions on the parameters of the model thanks to an energy analysis and we have rewritten the problem of designing a quasi-constant quality factor as a contrained least-square optimization problem. The parameters to optimize are the parameters of the family of dissipative models and the constraints are the stability of the final model. We are now considering the extension of the family to more general formulations and to heterogeneous media, before tackling multidimensional problems. These results have been obtained during the Master internship of Fatima Jabiri, in collaboration with Sébastien Imperiale (Inria Project-Team M3DISIM)

6.2.10. Asymptotic models for the electric potential across a highly conductive casing. Application to the field of resitivity measurements.

Participants: Hélène Barucq, Aralar Erdozain, Victor Péron.

A configuration that involves a steel-cased borehole is analyzed, where the casing that covers the borehole is considered as a highly conductive thin layer. Asymptotic techniques are presented as the suitable tool for deriving reduced problems capable of dealing with the numerical issues caused by the casing when applying the traditional numerical methods. The derivation of several reduced models is detailed by employing two different approaches, each of them leading to different classes of models. The stability and convergence of these models is studied and uniform estimates are proved. The theoretical orders of convergence are supported by numerical results obtained with the finite element method. We develop an application to the field of resistivity measurements. The second derivative of the potential which solves a reduced model has been employed to recover the resistivity of rock formations. These results are in accordance with an experiment of Kaufmann for the reference solution and have been obtained in collaboration with David Pardo (UPV/EHU).

6.2.11. Semi-Analytical Solutions for the Electric Potential across a Highly Conductive Casing. Participants: Hélène Barucq, Aralar Erdozain, Victor Péron.

A transmission problem for the electric potential is considered, where one part of the domain is a highconductive casing. Semi-analytical solutions are derived for several asymptotic models. These asymptotic models are designed to replace the casing by appropriate impedance conditions in order to avoid numerical instabilities. A decomposition in Fourier series of the solution to these asymptotic models is characterized. As an application we reproduced successfully the experiment of Kaufmann, using his same parameters, but computing with a fourth order asymptotic model. This experiment allows to recover the resitivity of rock formations employing a second derivative of the potential along the vertical direction. These results have been obtained in collaboration with Ignacio Muga (Pontificia Universidad Catolica of Valparaiso).

6.2.12. Asymptotic modeling of the electromagnetic wave scattering problem by a small sphere perfectly conducting.

Participants: Justine Labat, Victor Péron, Sébastien Tordeux.

In the context of non-destructive testing in medical imaging or civil engineering, the detection of small heterogeneities can be a difficult task in three dimensional domains. The complexity for solving numerically the direct problem both in terms of computation time and memory cost is due to the small size of obstacles in comparison with the incident wavelength and the large size of the domain of interest. Then the fine mesh size makes unsuitable or too expensive the use of classical numerical methods type continuous and discontinuous finite element methods or boundary element methods. The use of reduced models allows to get an approximation of the exact solution at a certain accuracy with a lower cost. We develop a Matched Asymptotic Expansions method to solve a time-harmonic electromagnetic scattering problem by a small sphere. This method allows to replace the scatterer by an equivalent asymptotic point source. In practice, it consists in defining an approximate solution using multi-scale expansions over far and near fields, related in a matching area. When the scatterer is a sphere, we make explicit the asymptotic expansions until the second order of approximation, relatively to the sphere radius. Numerical results make evident the convergence rate with respect to the sphere radius. Reference solutions are analytical solutions computed thanks to Montjoie Code. This work has been presented in the *Caleta Numerica* seminar, Pontificia Universidad Catolica of Valparaiso, Chili [48].

6.2.13. Asymptotic Models and Impedance Conditions for Highly Conductive Sheets in the Time-Harmonic Eddy Current Model.

Participant: Victor Péron.

This work is concerned with the time-harmonic eddy current problem for a medium with a highly conductive thin sheet. We present asymptotic models and impedance conditions up to the second order of approximation for the electromagnetic field. The conditions are derived asymptotically for vanishing sheet thickness ϵ where the skin depth is scaled like ϵ . The first order condition is the perfect electric conductor boundary condition. The second order condition turns out to be a Poincaré-Steklov map between tangential components of the magnetic field and the electric field [49]. Numerical experiments have been performed to assess the accuracy of the second order model. Complementary simulations will be conducted to study the robustness with respect to the sheet conductivity and the convergence of the modelling error. These results have been obtained in collaboration with Mohammad Issa and Ronan Perrussel (LAPLACE, CNRS/IMPT/UPS, Univ. de Toulouse) and this work has been presented in the international conference ACOMEN 2017 [33].

6.2.14. Numerical robustness of single-layer method with Fourier basis for multiple obstacle acoustic scattering in homogeneous media

Participants: Hélène Barucq, Juliette Chabassier, Ha Pham, Sébastien Tordeux.

We investigate efficient methods to simulate the multiple scattering of obstacles in homogeneous media. With a large number of small obstacles on a large domain, optimized pieces of software based on spatial discretization such as Finite Element Method (FEM) or Finite Difference lose their robustness. As an alternative, we work with an integral equation method, which uses single-layer potentials and truncation of Fourier series to describe the approximate scattered field. In the theoretical part of the paper, we describe in detail the linear

systems generated by the method for impenetrable obstacles, accompanied by a well-posedness study. For the numerical performance study, we limit ourselves to the case of circular obstacles. We first compare and validate our codes with the highly optimized FEM-based software Montjoie. Secondly, we investigate the efficiency of different solver types (direct and iterative of type GMRES) in solving the dense linear system generated by the method. We observe the robustness of direct solvers over iterative ones for closely-spaced obstacles, and that of GMRES with Lower–Upper Symmetric Gauss–Seidel and Symmetric Gauss–Seidel preconditioners for far-apart obstacles.

This work has been published in the journal Wave Motion, [15] and is also connected to the following conference presentations, [31], [41].

6.2.15. A study of the Numerical Dispersion for the Continuous Galerkin discretization of the one-dimensional Helmholtz equation

Participants: Hélène Barucq, Ha Pham, Sébastien Tordeux.

This work is a collaboration with Henri Calandra (TOTAL).

Although true solutions of Helmholtz equation are non-dispersive, their discretizations suffer from a phenomenon called numerical dispersion. While the true phase velocity is constant, the numerical one changes with the discretization scheme, order and mesh size. In our work, we study the dispersion associated with classical finite element. For arbitrary order of discretization, without using an Ansatz, we construct the numerical solution on the whole \mathbb{R} , and obtain an asymptotic expansion for the phase difference between the exact wavenumber and the numerical one. We follow an approach analogous to that employed in the construction of true solutions at positive wavenumbers, which involves Z-transform, contour deformation and limiting absorption principle. This perspective allows us to identify the numerical wavenumber with the angle of analytic poles. Such an identification is useful since the latter (analytic poles) can be numerically evaluated by an algorithm, which then yields the value of numerical wavenumber.

This work is detailed in the research report [44].

6.3. Supercomputing for Helmholtz problems

6.3.1. Numerical libraries for hybrid meshes in a discontinuous Galerkin context

Participants: Hélène Barucq, Lionel Boillot, Aurélien Citrain, Julien Diaz.

Elasticus team code has been designed for triangles and tetrahedra mesh cell types. The first part of this work was dedicated to add quadrangle libraries and then to extend them to hybrid triangles-quadrangles (so in 2D). This implied to work on polynomials to form functions basis for the (discontinuous) finite element method, to finally be able to construct reference matrices (mass, stiffness, ...).

A complementary work has been done on mesh generation. The goal was to encircle an unstructured triangle mesh, obtained by third-party softwares, with a quadrangle mesh layer. At first, we built scripts to generate structured triangle meshes, quadrangle meshes and hybrid meshes (triangles surrounded by quadrangles). We are now able to couple unstructured triangle mesh with structured quadrangle mesh, and we are now working on the implementation of the coupling between Discontinuous Galerkin methods (for the triangles) and Spectral Element methods (for the quadrangles).

6.3.2. Hybridizable Discontinuous Galerkin methods for solving the elastic Helmholtz equations

Participants: Marie Bonnasse-Gahot, Julien Diaz.

The advantage of performing seismic imaging in frequency domain is that it is not necessary to store the solution at each time step of the forward simulation. Unfortunately, the drawback of the Helmholtz equations, when considering 3D realistic elastic cases, lies in solving large linear systems. This represents today a challenging task even with the use of High Performance Computing (HPC). To reduce the size of the global linear system, we developed a Hybridizable Discontinuous Galerkin method (HDGm). It consists in expressing the unknowns of the initial problem in function of the trace of the numerical solution on each face of the mesh cells. In this way the size of the matrix to be inverted only depends on the number of degrees of freedom on each face and on the number of the faces of the mesh, instead of the number of degrees of freedom on each cell and on the number of the cells of the mesh as we have for the classical Discontinuous Galerkin methods (DGm). The solution to the initial problem is then recovered thanks to independent elementwise calculation. These results have been published in [18]. This is a collaboration with Henri Calandra (Total) and Stéphane Lanteri (Inria Project Team Nachos)

6.3.3. Scalability of linear solvers for Hybridizable Discontinuous Galerkin methods

Participants: Marie Bonnasse-Gahot, Julien Diaz.

We coupled our HDG code with tested two linear solvers: a parallel sparse direct solver MUMPS (MUltifrontal Massively Parallel sparse direct Solver) and a hybrid solver MaPHyS (Massively Parallel Hybrid Solver) which combines direct and iterative methods. In the framework of the european project HPC4E, we analyzed the scalability of the two solvers on the plateform Plafrim We compared the performances of the two solvers when solving 3D elastic waves propagation over HDGm. These comparisons were presented at the 2017 EAGE Workshop on High Performance Computing for Upstream and at MATHIAS 2017 conferences. This is a collaboration with Henri Calandra (Total), Luc Giraud, Mathieu Kuhn (Inria Project-Team Hiepacs) and Stéphane Lanteri (Inria Project Team Nachos).

6.4. Hybrid time discretizations of high-order

6.4.1. Construction and analysis of a fourth order, energy preserving, explicit time discretization for dissipative linear wave equations.

Participants: Juliette Chabassier, Julien Diaz, Anh-Tuan Ha.

We submitted a paper to M2AN. This paper deals with the construction of a fourth order, energy preserving, explicit time discretization for dissipative linear wave equations. This scheme is obtained by replacing the inversion of a matrix, that comes naturally after using the technique of the Modified Equation on the second order Leap Frog scheme applied to dissipative linear wave equations, by an explicit approximation of its inverse. The stability of the scheme is studied first using an energy analysis, then an eigenvalue analysis. Numerical results in 1D illustrate the good behavior regarding space/time convergence and the efficiency of the newly derived scheme compared to more classical time discretizations. A loss of accuracy is observed for non smooth profiles of dissipation, and we propose an extension of the method that fixes this issue. Finally, we assess the good performance of the scheme for a realistic dissipation phenomenon in Lorentz's materials. This work has been done in collaboration with Sébastien Imperiale (Inria Project-Team M3DISIM) and Alain Anh-Tuan Ha (Internship at Magique 3D in 2016).

6.4.2. Higher-order optimized explicit Runge-Kutta schemes for linear ODEs

Participants: Hélène Barucq, Marc Duruflé, Mamadou N'Diaye.

In this work, we have constructed optimized explicit Runge-Kutta schemes for linear ODEs that we called Linear-ERK. Theses schemes can be applied to the following ODE

$$M_h \frac{dU}{dt} = K_h U + F(t)$$

where M_h is the mass matrix, K_h the stiffness matrix and F(t) a source term. Linear-ERK schemes are constructed using polynomial stability functions which are obtained by maximizing the CFL number. We have considered a polynomial stability function based on the Taylor series expansion of an exponential function. Then, we have added extra terms beyond the terms of the Taylor expansion without changing the order of accuracy. The coefficients of those extra terms have been computed by optimizing the CFL number such that the stability region of the developed scheme includes a typical spectrum. This spectrum has been obtained by computing eigenvalues of the matrix $M_h^{-1}K_h$ for the wave equation solved on a square with Hybrid Discontinuous Formulation (HDG). The optimization is performed by using the algorithm developed by D. Ketcheson and coworkers. By proceeding this way, we have obtained optimized explicit schemes up to order 8. We have also determined the CFL number and the efficiency on the typical spectrum for each explicit scheme. We have provided algorithms to implement these schemes and numerical results to compare them.

This work is a chapter of the thesis defended by Mamadou N'diaye on December 8, 2017, under the joint supervision of Hélène Barucq and Marc Duruflé.

6.4.3. High-order locally implicit time schemes for linear ODEs

Participants: Hélène Barucq, Marc Duruflé, Mamadou N'Diaye.

In this work we have proposed a method that combines optimized explicit schemes and implicit schemes to form locally implicit schemes for linear ODEs, including in particular ODEs coming from the space discretization of wave propagation phenomena. This method can be applied to the following ODE

$$M_h \frac{dU}{dt} = K_h U + F(t)$$

Like in the local time-stepping developed by Grote and co-workers, the computational domain is split into a fine region and a coarse region. The matrix A_h is given as

$$A_h = M_h^{-1} K_h = A_h P + A_h (I - P)$$

where P is the projector on the fine region of the computational domain. Then the proposed locally implicit method is obtained from the combination of the A-stable implicit schemes we have developed in 2016 (Padé schemes or Linear-SDIRK schemes detailed in [17]) on the fine region and explicit schemes with optimal CFL number in the coarse region. The developed method has been used to solve the acoustic wave equation and we have checked the convergence in time of these schemes for order 4, 6 and 8.

This work has been presented at the Mathias annual Total seminar and is a chapter of the thesis defended by Mamadou N'diaye on December 8, 2017, under the joint supervision of Hélène Barucq and Marc Duruflé.

MNEMOSYNE Project-Team

7. New Results

7.1. Overview

This year we have explored two main cortico-basal loops of cerebral architecture, the limbic and motor loops, and their associated memory mechanisms. The limbic loop (cf. § 7.2) concerns the taking into account of the emotional and motivational aspects by the respondant and operant conditioning and their relations with the semantic and episodic memories. The motor loop (cf. § 7.3) considers the evolution of sensorimotor learning, from goal-directed behaviors to habitual behaviors.

We have also worked on the systemic integration of our models (*cf.* § 7.4), raising the question of the conditions of autonomous learning and certain global characteristics such as neuromodulation.

Finally, we study the links between our bio-inspired modeling work and Machine Learning (*cf.* § 7.5), revisiting this latter domain in the light of the principles highlighted by our models.

7.2. The limbic loop

We explore the limbic loop by studying a series of neural mechanisms that propose how respondant conditioning results from interactions between the amygdala, the nucleus accumbens and the limbic pole of the frontal cortex. In our models, this learning is also fed by exchanges with the hippocampus (episodic memory) [6] and the sensory cortex (semantic memory). We have also addressed the difficult question of the articulation between the respondant and operant conditioning in particular in the nucleus accumbens.

Also in connection with this loop, we studied the dynamics of dopamine release in the midbrain, considered to play an essential role in the coding of the prediction error. This model [12] developed in the framework of our collaboration with India (*cf.* § 9.3) proposes to introduce into the classical circuit, new actors (such as the pedunculopontine tegmental nucleus in the brainstem) and new functions (dissociation of amplitude and timing of the reward), that we will seek to corroborate in the future.

7.3. The motor loop

The nervous system structures involved in decision making constitute a circuit formed by the basal ganglia, the cortex, the thalamus and their numerous interconnections. This circuit can be described as a set of loops operating in parallel and interacting at different points. The decisions and therefore the actions of an individual emerge from the interactions between these loops and the plasticity of their connections. These emerging behaviors and arising learning processes are addressed through a closed-loop approach in which the theoretical model is in constant interaction with the environment of the task. To this end, neural modeling and dedicated analysis software tools were developed in the laboratory, at the level of the neuronal circuit.

7.4. Systemic integration

Systemic integration promotes the idea of developing large models that associate several cortico-basal loops and even other cerebral structures and more generally takes into account the influence of the body on this network [19]. This requires to propose a global picture for the organization and functional association between all these elements [18] and to analyze its consequences from a representational point of view [1] and also concerning autonomous learning [7].

It also requires to evaluate the properties of such systems from their interactions with the body and the environment, as we have done this year using the VirtualEnaction platform.

7.5. Machine Learning

In this section, we report on some neuronal adaptive mechanisms, that we develop at the frontier between Machine Learning and Computational Neuroscience. Our goal is to consider and adapt models in Machine Learning for their integration in a bio-inspired framework. We were interested this year in three paradigms of computation.

The first paradigm concerns the manipulation of temporal sequences. In a perspective of better understanding how the brain learns structured sequences we work on a model on syntax acquisition and Human-Robot Interaction using the Reservoir Computing framework (using random recurrent networks) [24], [15], [17] with our collaborators at the University of Hamburg (*cf.* § 9.3). A syntactic re-analysis system [15], which corrects syntax errors in speech recognition hypotheses, was built in order to enhance vocal Human-Robot Interaction and to enhance the previously developped model [40]. Additionally, the ability to deal with several languages (from different language families) of this later model of sentence parsing [40] was evaluated. We showed that it can successfully learn to parse sentences related to home scenarios in fifteen languages originating from Europe and Asia [24]. In a different perspective, in order to try to overcome word misrecognition at a more basic level, we tested whether the same architecture was able to process directly phonemes instead of grammatical constructions [17]. Applied on a small corpus, we see that the model has similar performance.

In an industrial application for the representation of electrical diagrams (*cf.* § 8.1), we also study how recurrent layered models can be trained to run through these schemes for prediction and sequence representation tasks [10].

The second paradigm concerns the extraction of characteristics and the use of hierarchical networks, as in the case of deep networks. An industrial application (*cf.* § 9.2) allows us to revisit these models to make them more easily usable in constrained frameworks, for example with limited size corpuses, and more interpretable introducing a new notion of prototypes and exploring the capability to learn the network architecture itself (using shortcuts) [11]. In order to push the sate of the art, the next step is going to consider not only feed-forward but also recurrent architecture, and to this end neural network recurrent weight estimation through backward tuning has been revisited [21].

The third paradigm is about spatial computation. We have designed a graphical method originating from the computer graphics domain that is used for the arbitrary and intuitive placement of cells over a two-dimensional manifold. Using a bitmap image as input, where the color indicates the identity of the different structures and the alpha channel indicates the local cell density, this method guarantees a discrete distribution of cell position respecting the local density function. This method scales to any number of cells, allows to specify several different structures at once with arbitrary shapes and provides a scalable and versatile alternative to the more classical assumption of a uniform non-spatial distribution. This preliminary work will be used in the design of a new class of model where explicit topography allows to connect structure according to known pathways.

MONC Project-Team

6. New Results

6.1. Mathematical modeling of tumor-tumor distant interactions supports a systemic control of tumor growth

Authors: Sébastien Benzekry, Clare Lamont, Dominique Barbolosi, Lynn Hlatky, Philip Hahnfeldt. Paper published in Cancer Research, https://hal.inria.fr/hal-01566947.

Interactions between different tumors within the same organism have major clinical implications, especially in the context of surgery and metastatic disease. Three main explanatory theories (competition, angiogenesis inhibition and proliferation inhibition) have been proposed but precise determinants of the phenomenon remain poorly understood. Here we formalized these theories into mathematical models and performed biological experiments to test them with empirical data. In syngeneic mice bearing two simultaneously implanted tumors, growth of only one of the tumors was significantly suppressed (61% size reduction at day 15, p<0.05). The competition model had to be rejected while the angiogenesis inhibition and proliferation inhibition models were able to describe the data. Additional models including a theory based on distant cytotoxic log-kill effects were unable to fit the data. The proliferation inhibition model was identifiable and minimal (4 parameters), and its descriptive power was validated against the data, including consistency in predictions of single tumor growth when no secondary tumor was present. This theory may also shed new light on single cancer growth insofar as it offers a biologically translatable picture of how local and global action may combine to control local tumor growth, and in particular, the role of tumor-tumor inhibition. This model offers a depiction of concomitant resistance that provides an improved theoretical basis for tumor growth control and may also find utility in therapeutic planning to avoid post-surgery metastatic acceleration. Major findings In mice bearing two tumors implanted simultaneously, tumor growth was suppressed in one of the two tumors. Three theories of this phenomenon were advanced and assessed against the data. As formalized, a model of competition for nutrients was not able to explain the growth behavior as well as indirect, angiogenesis-regulated inhibition or a third model based on direct systemic inhibition. This last model offers a depiction of concomitant resistance that provides an improved theoretical basis for tumor growth control and may also find utility in therapeutic planning to avoid post-surgery metastatic acceleration.

6.2. Precision of manual two-dimensional segmentations of lung and liver metastases and its impact on tumour response assessment using RECIST 1.1

Authors: *François Cornelis, Marie Martin, Olivier Saut*, Xavier Buy, *Michèle Kind*, Jean Palussiere, *Thierry Colin*. Paper published in European Radiology Experimental, https://hal.inria.fr/hal-01634849.

Response evaluation criteria in solid tumours (RECIST) has significant limitations in terms of variability and reproducibility, which may not be independent. The aim of the study was to evaluate the precision of manual bi-dimensional segmentation of lung, liver metastases, and to quantify the uncertainty in tumour response assessment. Methods: A total of 520 segmentations of metastases from six livers and seven lungs were independently performed by ten physicians and ten scientists on CT images, reflecting the variability encountered in clinical practice. Operators manually contoured the tumours, firstly independently according to the RECIST and secondly on a preselected slice. Diameters and areas were extracted from the segmentations. Mean standard deviations were used to build regression models and 95% confidence intervals (95% CI) were calculated for each tumor size and for limits of progressive disease (PD) and partial response (PR) derived from RECIST 1.1. Results: Thirteen aberrant segmentations (2.5%) were observed without significant differences between the physicians and scientists; only the mean area of liver tumours (p = 0.034) and mean diameter of lung tumours (p = 0.021) differed significantly. No difference was observed between the methods. Interobserver agreement was excellent (intra-class correlation >0.90) for all variables. In liver, overlaps of the 95% CI with the 95% CI of limits of PD or PR were observed for diameters above 22.7 and 37.9 mm, respectively. An overlap of 95% CIs was systematically observed for area. No overlaps were observed in lung. Conclusions: Although the experience of readers might not affect the precision of segmentation in lung and liver, the results of manual segmentation performed for tumor response assessment remain uncertain for large liver metastases.

6.3. Dose- and time-dependence of the host-mediated response to paclitaxel therapy: a mathematical modeling approach

Authors: Madeleine Benguigui, Dror Alishekevitz, Michael Timaner, Dvir Shechter, Ziv Raviv, *Sebastien Benzekry*, Yuval Shaked. Paper published in OncoTarget, https://hal.inria.fr/hal-01672568.

It has recently been suggested that pro-tumorigenic host-mediated processes induced in response to chemotherapy counteract the anti-tumor activity of therapy, and thereby decrease net therapeutic outcome. Here we use experimental data to formulate a mathematical model describing the host response to different doses of paclitaxel (PTX) chemotherapy as well as the duration of the response. Three previously described host-mediated effects are used as readouts for the host response to therapy. These include the levels of circulating endothelial progenitor cells in peripheral blood and the effect of plasma derived from PTX-treated mice on migratory and invasive properties of tumor cells in vitro. A first set of mathematical models, based on basic principles of pharmacokinetics/pharmacodynamics, did not appropriately describe the dose-dependence and duration of the host response regarding the effects on invasion. We therefore provide an alternative mathematical model with a dose-dependent threshold, instead of a concentration-dependent one, that describes better the data. This model is integrated into a global model defining all three host-mediated effects. It not only precisely describes the data, but also correctly predicts host-mediated effects at different doses as well as the duration of the host response. This mathematical model may serve as a tool to predict the host response to chemotherapy in cancer patients, and therefore may be used to design chemotherapy regimens with improved therapeutic outcome by minimizing host mediated effects.

6.4. Tumor growth model of ductal carcinoma: from in situ phase to stroma invasion

Authors: *Olivier Gallinato, Thierry Colin, Olivier Saut, Clair Poignard* Paper published in Journal of Theoretical Biology, https://hal.inria.fr/hal-01598837.

This paper aims at modeling breast cancer transition from the in situ stage –when the tumor is confined to the duct- to the invasive phase. Such a transition occurs thanks to the degradation of the duct membrane under the action of specific enzymes so-called matrix metallo-proteinases (MMPs). The model consists of advection-reaction equations that hold in the duct and in the surrounding tissue, in order to describe the proliferation and the necrosis of the cancer cells in each subdomain. The divergence of the velocity is given by the increase of the cell densities. Darcy law is imposed in order to close the system. The key-point of the modeling lies in the description of the transmission conditions across the duct. Nonlinear Kedem-Katchalsky transmission conditions across the membrane describe the discontinuity of the pressure as a linear function of the flux. These transmission conditions make it possible to describe the transition from the in situ stage to the invasive phase at the macroscopic level. More precisely, the membrane permeability increases with respect to the local concentration of MMPs. The cancer cells are no more confined to the duct and the tumor invades the surrounding tissue. The model is enriched by the description of nutrients concentration, tumor necrosis factors, and MMPs production. The mathematical model is implemented in a 3D C ++-code, which is based on well-adapted finite difference schemes on Cartesian grid. The membrane interface is described by a levelset, and the transmission conditions are precisely approached at the second order thanks to well-suited sharp stencils. Our continuous approach provides new significant insights in the macroscopic modeling of the breast cancer phase transition, due to the membrane degradation by MMP enzymes.

6.5. Superconvergent second order Cartesian method for solving free boundary problem for invadopodia formation

Authors: Olivier Gallinato, Clair Poignard. Paper published in Journal of Computational Physics, https://hal. inria.fr/hal-01483484.

In this paper, we present a superconvergent second order Cartesian method to solve a free boundary problem with two harmonic phases coupled through the moving interface. The model recently proposed by the authors and colleagues describes the formation of cell protrusions. The moving interface is described by a level set function and is advected at the velocity given by the gradient of the inner phase. The finite differences method proposed in this paper consists of a new stabilized ghost fluid method and second order discretizations for the Laplace operator with the boundary conditions (Dirichlet, Neumann or Robin conditions). Interestingly, the method to solve the harmonic subproblems is superconvergent on two levels, in the sense that the first and second order derivatives of the numerical solutions are obtained with the second order of accuracy, similarly to the solution itself. We exhibit numerical criteria on the data accuracy to get such properties and numerical simulations corroborate these criteria. In addition to these properties, we propose an appropriate extension of the velocity of the level-set to avoid any loss of consistency, and to obtain the second order of accuracy of the complete free boundary problem. Interestingly, we highlight the transmission of the superconvergent properties for the static subproblems and their preservation by the dynamical scheme. Our method is also well suited for quasistatic Hele-Shaw-like or Muskat-like problems.

6.6. A Voronoi Interface approach to cell aggregate electropermeabilization

Authors: Arthur Guittet, Clair Poignard, Frederic Gibou.

In this work, a Voronoi Interface approach to the study of cell electropermeabilization is presented. We consider the nonlinear electropermeabilization model of Poignard et al., which takes into account the jump in the voltage potential across cells' membrane. The jump condition is imposed in a sharp manner, using the Voronoi Interface Method of Guittet et al., while adaptive Quad-/Oc-tree grids are employed to automatically refine near the cells boundary for increased accuracy. Numerical results are provided to illustrate the accuracy of the methods. We also carry out simulations in three spatial dimensions to investigate the influence of shadowing and of the cells shape on the degree of permeabilization.

6.7. Revisiting bevacizumab + cytotoxics scheduling using mathematical modeling: proof of concept study in experimental non-small cell lung carcinoma

Authors: D.C. Imbs, R. El Cheikh, A. Boyer, J. Ciccolini, C. Mascaux, B. Lacarelle, F. Barlesi, D. Barbolosi and *S. Benzekry*. Paper published in CPT Pharmacometrics Syst Pharmacol, https://hal.inria.fr/hal-01624423.

Concomitant administration of bevacizumab and pemetrexed-cisplatin is a common treatment for advanced nonsquamous non-small cell lung cancer (NSCLC). Vascular normalization following bevacizumab administration may transiently enhance drug delivery, suggesting improved efficacy with sequential administration. To investigate optimal scheduling, we conducted a study in NSCLC-bearing mice. First, experiments demonstrated improved efficacy when using sequential vs. concomitant scheduling of bevacizumab and chemotherapy. Combining this data with a mathematical model of tumor growth under therapy accounting for the normalization effect, we predicted an optimal delay of 2.8 days between bevacizumab and chemotherapy. This prediction was confirmed experimentally, with reduced tumor growth of 38% as compared to concomitant scheduling, and prolonged survival (74 vs. 70 days). Alternate sequencing of 8 days failed in achieving a similar increase in efficacy, thus emphasizing the utility of modeling support to identify optimal scheduling. The model could also be a useful tool in the clinic to personally tailor regimen sequences.

PLEIADE Team

6. New Results

6.1. Alcyone system for repeatable e-science

One of PLEIADE's goals is to assist scientific users in deploying analysis software in their desktop environments. Increasingly, this is not a question of installing software packages locally, but of building bespoke environments that comprise many cooperating software tools. A typical example is a local Galaxy instance, communicating with a project-specific database that is shared with visualization and analysis tools, and cooperating with an electronic notebook such as Jupyter. In order to foster repeatable science, the configuration of each such environment should be reliably recorded, in a way that allows it to be redeployed in the future or shared with a colleague.

PLEIADE's **Alcyone** system provides a mechanism for specifying and deploying software environments for scientific users in bioinformatics and biodiversity. Alcyone offers three facilities:

- 1. A *specification* using configuration-by-convention style, combining specification files in YAML format and raw data files.
- 2. A collection of Dockerized services that can be chosen in the specification.
- 3. A *deployment system* that compiles the specification into a master container image, which orchestrates the deployment and management of the service containers.

The user's environment is fully specified in files that can be archived and shared, allowing future reuse. The use of Docker containers guarantees that future deployments run exactly as before, since the precise versions of the service containers are recorded.

Furthermore, Alcyone specifications are files, that can be managed by the Git source code control system. Different versions of the environment, including different analysis pipelines and intermediate results, are stored in the Git history and any version can be resurrected and deployed. Git branches can also be used to share configurations between users in the same lab.

Alcyone is being tested internally by PLEIADE and is undergoing intense development. Existing service containers are PLEIADE's Magus knowledge base, Magecal gene prediction pipeline, and Mimoza metabolic network explorer; as well as third-party tools Galaxy, Gbrowse, and Jbrowse.

6.2. Clavispora lusitaniae

Clavispora lusitaniae, an environmental saprophytic yeast belonging to the CTG clade of *Candida* and a teleomorph of *Candida lusitaniae*, is an environmentally ubiquitous ascomycetous yeast with no known specific ecological niche. It can be isolated from different substrates, such as soils, waters, plants, and gastrointestinal tracts of many animals including birds, mammals, and humans. In immunocompromised hosts, *C. lusitaniae* can be pathogenic and is responsible for about 1% of invasive candidiasis, particularly in pediatric and hematology-oncology patients.

The Laboratoire de Microbiologie Fondamentale et Pathogénicité UMR-CNRS 5234 and PLEIADE sequenced and annotated the genome of *C. lusitaniae* type strain CBS 6936, and analyzed it in comparison with the strains ATCC 42720, isolated from the blood of a patient with myeloid leukemia, and MTCC 1001, a self-fertile strain isolated from citrus. In spite of a conserved genome structure, the genomes have undergone significant divergence. In particular the SNP density of 1 SNP per 90 bp is twice the level observed between strains SC5314 and WO-1 of *Candida albicans*, which are members of different subgroups within the species and qualified as having diverged relatively recently.

This work contributes to PLEIADE's long-term goal of developing understanding how diversity measured at the genome level can be made to correspond with observed functional diversity.

6.3. Introgressions as a source of diversity

Several prominent mechanisms of genomic evolution have been described for the yeasts, among them interspecific hybridization, reticulated evolution, aneuploidization, recent or ancient poly-ploidization events, large chromosomal duplication or more limited gene duplication, and horizontal transfer. These mechanisms are usually so closely intertwined that it is difficult to determine which ones are causes or consequences. Regardless of mechanisms the result has been a drastic reshaping of yeasts genome along evolution. Understanding these mechanisms is important, not only for strain construction in biotechnology, but also more fundamentally for insight into the causes and effects of genome reshaping on much shorter time scales.

Introgression, the transfer of large or more limited genetic information from one species to another, is an evolutionary mechanism of particular interest in industrial applications such as wine making where large vat cultures are used. Introgression results in mosaic genomes, and can be the result of interspecific hybridization fol- lowed by the extensive loss of one parental genome, either through repeated backcross with one parental species or through missegregation of the hybrid at meiosis.

In collaboration with the Institut des Science de la Vigne et du Vin and Bordeaux Sciences Agro, PLEIADE developed tools to rapidly assess the presence of introgressed regions in a large population of *Saccharomyces uvarum* isolates (104 strains), focusing on Holarctic isolates from natural, cider and wine environments since introgressed regions are absent in Southern hemisphere isolates. The overall number of introgressed regions is significantly higher in cider-associated strains compared to wild strains, and is higher in wine isolates. However, only a subset of the introgressed regions were found to be overrepresented in anthropic activities and their number and quality varied between cider- and wine-making processes.

Paradoxically, the low Holarctic genetic diversity observed in [1] contrasts with the relative high phenotypic diversity found for technological traits. This contradiction suggests that interspecific introgressions found among Holarctic *S. uvarum* strains could be the most important source of genetic, and by extention of phenotypic, diversity.

6.4. New results Biodiversity

The activity of PLEIADE in computational biodiversity has consisted mainly in reinforcing a cooperation with actors in High Performance Computing, namely Inria team Hiepacs, for method developments in metabarcoding. Metabarcoding is a supervised or unsupervised statistical learning method, to build taxonomic inventories from so called environmental samples, i.e. sets of short reads of a same marker for a whole community or guild. Most of tools used therefore still rely on some classical ones shaped in Multivariate Data Analysis. Those tools are indeed well known, but still are often behind the scene in current developments in Machine Learning (like kernel PCA, Support Vector Machines, etc. ...). Most of them, if not all, are based on Singular Value Decomposition of a matrix. If p features are observed on n items, the size of the matrix is $n \times p$. The complexity of such algorithms is in $O(p^3)$. The recent development of NGS data has had as a consequence to multiply by a factor $10^2/10^3$ the size of data sets. This leads to a factor $10^6/10^9$ of required computation time. Reaching such a goal is beyond resources currently offered by parallelization. Hence, a new approach has been selected, by using other methods. Indeed, it has been known for some years now that concentration of measure phenomena (a sort of extension of law of large numbers) leads to a blessing of dimensionality, i.e. some randomized methods are available as heuristics to make some matrix computations efficiently and accurately. This is the case for running SVD. Therefore, a cooperation has been set up between HiePacs and PLEIADE through Pierre Blanchard (a former Hiepacs PhD student who has held a post-doc position during 7 months in PLEIADE) to implement those methods in the framework of metabarcoding. Former work in PLEIADE had led (with a DARI project 2014-2016) to the production of many high-dimensional pairwise distance matrices of DNA environmental samples (amplicon based metabarcoding). Classical Multidimensional Scaling of some of those matrices has been programmed in C++, with dedicated libraries in domain of so called random projection, or column selection (fmr library). This has permitted to build a point cloud of an environmental sample of 1.2×10^5 reads, and see its "shape", with eyes, from projections on first axis, and build a low dimensional approximation of it. The outcome is twofolds: (i) build a point cloud attached to an environmental sample, for further ecological studies and (ii) delivery of a scientific library in High Performance Computing for randomized matrix computations. These research lines will be carried on in 2018, and the cooperation extended to mésocentre GRICAD in Grenoble for HPC and C++ code development.

PLEIADE has carried on statistical learning methods, both supervised and unsupervised in metabarcoding. A cooperation with IMBE at Marseille has permitted to associate MDS as developed above with graph based methods (building connected components of a graph built from pairwise distance matrices after thresholding), and test these methods for unsupervised statistical learning (OTU building) of data sets from an ongoing PhD in Marseille Bay. Cooperation with Institut Pasteur at Cayenne has lead to a joint publication [12] for a proof of concept of an inventory by metagenomics of viromes of bats in French Guiana, with two objectives: (i) detect as soon as possible some strains which could potentially be transmitted to man and (ii) develop a viral ecology by studying further how environmental factors and nature of the host drive the virome composition.

Meanwhile, PLEIADE has carried on cooperation with SLU University at Uppsala especially on metabarcoding of diatom communities in rivers and lakes in Sweden (co-direction of a PhD student located at Uppsala in SLU), and first steps in biogeography of diatoms in Fennoscandia (cooperation with a PostDoc in SLU).

SISTM Project-Team

7. New Results

7.1. Statistical and mechanistic modeling

- Prague M, Commenges D, Gran JM, Ledergerber B, Young J, Furrer H, Thiébaut R. Dynamic models for estimating the effect of HAART on CD4 in observational studies: Application to the Aquitaine Cohort and the Swiss HIV Cohort Study. Biometrics. 2017;73:294-304. [26]
 Comparison of descriptive models (Marginal structural models) and mechanistic models (Ordinary differential equations with mixed effect models on parameters) performances for estimating treatment effect from observational studies.
- Turner EL, Li F, Gallis JA, Prague M, Murray DM. Review of recent methodological developments in group-randomized trials: part 1 design. American Journal of Public Health 107(6), 907-915. [27]
- Mélanie Prague, Elisabeth Turner, Gallis John, Li Fan, Murray David. Review of recent Methodological Developments in group-randomized trials: Part 2 - Analysis. American Journal of Public Health, American Public Health Association, 2017. [28]

Review of the literature on how to analyse data from a cluster randomised trial

Review of literature regarding Design of cluster randomised trials.

- Jarne Ana, Daniel Commenges, Mélanie Prague, Yves Levy, Rodolphe Thiébaut and inspire 2/3 study group. Modelling CD4+ T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous interleukin 7. Annals of applied statistics 11(3) 1593-1616. [21]
 Modelling the disability of CD4 restauration by repeated cycles of Intereukine-7 injections using mechanistic models.
- Prague M., Wang R., Stephens A., Tchetgen Tchetgen E, DeGruttola V. Accounting for interference variables using semi-parametric augmentation for improving efficiency in clustered randomized trials with missing at random outcomes. Biometrics 72(4) 1066-1077. [29] Doubly robust approach to estimate the treatment effect in Cluster randomised trials.

7.2. Statistical learning methods for high-dimensional data

- Genuer R, Poggi J-M, Tuleau-Malot C, Villa-Vialaneix N. Random Forests for Big Data. Big Data Research, 9 (2017). [18]
 Addresses the analysis of Big Data with Random Forests, review of existing algorithms, simulation study and recommandations.
- Agniel D and Hejblum BP, Variance component score test for time-course gene set analysis of longitudinal RNA-seq data, Biostatistics, 18(4):589–604, 2017.[16]
 We propose tcgsaseq, a principled, model-free, and efficient method for detecting longitudinal changes in RNA-seq gene sets defined a priori. Applied to both simulated data and two real datasets, tcgsaseq is shown to exhibit very good statistical properties, with an increase in stability and power when compared to state-of-the-art methods
- Hejblum BP, Alkhassim C, Gottardo R, Caron F, Thiébaut R. Sequential Dirichlet Process Mixtures of Multivariate Skew t-distributions for Model-based Clustering of Flow Cytometry Data, preprint on ArXiv. [39]

We propose to use a Bayesian nonparametric approach with Dirichlet process mixture of multivariate skew t-distributions to perform model based clustering of flow-cytometry data, robustly estimating the number of cell populations from the data.

7.3. Software tools

Mougin F, Auber D, Bourqui R, Diallo G, Dutour I, Jouhet V, Thiessard F, Thiébaut R, Thébault P. Visualizing omics and clinical data: Which challenges for dealing with their variety? Methods. 2017.
 [23]

This is a review on the methods to visualize the big data in the context of clinical research.

• Prague M., Wang R. and de Grutolla V. CRTgeeDR: An R package for generalized estimating equations with missing data in cluster randomized trials. R journal (in press). [29] Diffusion of a package to estimate estimate the intervention effect of a prevention strategy agains epidemics in cluster randomised trials. Estimation is based on GEE

7.4. Analysis of results from Clinical trials and cohorts in HIV

- Bouteloup V, Sabin C, Mocroft A, Gras L, Pantazis N, Le Moing V, d'Arminio Monforte A, Mary-Krause M, Roca B, Miro JM, Battegay M, Brockmeyer N, Berenguer J, Morlat P, Obel N, De Wit S, Fätkenheuer G, Zangerle R, Ghosn J, Pérez-Hoyos S, Campbell M, Prins M, Chêne G, Meyer L, Dorrucci M, Torti C, Thiébaut R; Standard Reference Distribution of CD4 Response to HAART Project Team for the Collaboration of Observational HIV Epidemiological Research Europe (COHERE) in EuroCoord. Reference curves for CD4 T-cell count response to combination antiretroviral therapy in HIV-1-infected treatment-naïve patients. HIV Medicine. 2017;18:33-44. This is a tool that should help clinicians to evaluate the immunological response to antiretroviral therapy in HIV infected patients. Thanks to the analyse of one of the largest observational database in the world, we provide with an online tool references on the CD4 count during the first year of antiretroviral therapy.
- Picat MQ, Pellegrin I, Bitard J, Wittkop L, Proust-Lima C, Liquet B, Moreau JF, Bonnet F, Blanco P, Thiébaut R; ANRS CO3 Aquitaine Cohort. Integrative Analysis of Immunological Data to Explore Chronic Immune T-Cell Activation in Successfully Treated HIV Patients. PLoS One. 2017;12:e0169164. [25]

This is an analysis of data on gene expression and factors associated to the immune activation in HIV-infected patients. Using structural models, we disentangle the effect of factors such as CMV and the mediation through type 1 interferon pathway.

• Thiébaut R, Hue S, Le Marec F, Lelièvre JD, Dupon M, Foucat E, Lazaro E, Dabis F, Duffau P, Wittkop L, Surenaud M, Pellegrin I, Lacabaratz C, Bonnet F, Lévy Y; ANRS CO3 Aquitaine Cohort. Serum ST2 level is an independent predictor of all-cause mortality in HIV-infected patients. Aquitaine Cohort, France. AIDS. 2017 [32]

In this work, we have demonstrated the independent effect of the biomarker ST2 on the overall mortality in a large cohort of HIV infected patients.

• Vladimir Novitsky, Mélanie Prague, Sikhulile Moyo, Tendani Gaolathe, Mompati Mmalane, et al.. High HIV-1 RNA among Newly Diagnosed People in Botswana. AIDS Research and Human Retroviruses, Mary Ann Liebert, To appear.

7.5. Analysis of results from Clinical trials and cohorts in Ebola

Rechtien A, Richert L, Lorenzo H, Martrus G, Hejblum B, Dahlke C, Kasonta R, Zinser M, Stubbe H, Matschl U, Lohse A, Krähling V, Eickmann M, Becker S; VEBCON Consortium, Thiébaut R, Altfeld M, Addo MM. Systems Vaccinology Identifies an Early Innate Immune Signature as a Correlate of Antibody Responses to the Ebola Vaccine rVSV-ZEBOV. Cell Report. 2017;20:2251-2261. [30]

In this work, we have analyzed high-dimensional gene expression and cell characterization data. We showed the predictive capacity of the innate immune response to the Ebola vaccine to define the antibody response established beyond one month. This is a successful application of integrative analyses tools on high dimensional immunogenicity data from an Ebola vaccine trial with identification of early correlates of later antibody responses.

7.6. Analysis of results from clinical trials and cohorts in other fields (Epidemiology, Medical Sciences, Neuroimaging, Sport Sciences)

- Zago L, Hervé PY, Genuer R, Laurent A, Mazoyer B, Tzourio-Mazoyer N, Joliot M. Predicting hemispheric dominance for language production in healthy individuals using support vector machine. Hum Brain Mapp. 2017 Dec;38(12):5871-5889. doi: 10.1002/hbm.23770. [34] Joint work with the GIN-IMN team, application of a variable selection procedure based on SVM method to analyze functional MRI data.
- Tabue-Teguo M, Grasset L, Avila-Funes JA, Genuer R, Proust-Lima C, Péres K, Féart C, Amieva H, Harmand MG, Helmer C, Salles N, Rainfray M, Dartigues JF. Prevalence and Co-Occurrence of Geriatric Syndromes in People Aged 75 Years and Older in France: Results From the Bordeaux Three-city Study. J Gerontol A Biol Sci Med Sci. (2017) [31]
 Application of Multiple Correspondence Analysis which enlights frailty and dependent profile of people from the Three-city study.
- Née M, Avalos M, Luxcey A, Contrand B, Salmi LR, Fourrier-Réglat A, Gadegbeku B, Lagarde E, Orriols L. Prescription medicine use by pedestrians and the risk of injurious road traffic crashes: A case-crossover study. PLoS Medicine. Jul 18;14(7):e1002347 (2017) [24] Exploration of the association between the use of medicinal drugs and the risk of being involved in a road traffic crash as a pedestrian. We applied the Lasso methodology that we previously developed for the case-crossover design in a high-dimensional setting. This design controls for time-invariant factors by using each case as its own control. This study highlights the necessity of improving awareness of the effect of medicines on pedestrians.
- Hellard P, Scordia C, Avalos M, Mujika I, Pyne DB. Modelling of optimal training load patterns during the 11 weeks preceding major competition in elite swimmers. Applied Physiology, Nutrition, and Metabolism. Jun 26 (2017) [20]

Quantification of the relationships between the effects of periodization variables and competitive performance in elite swimmers using semiparametric mixed effects models. In the framework of the 2014-2016 R&D project "Quels schémas de périodisation pour la préparation des Jeux Olympiques à Rio ?" with the French Swimming Federation.

- Hejblum BP, Cui J, Lahey LJ, Cagan A, Sparks JA, Sokolove J, Cai T, Liao KP, Association between anti-citrullinated fibrinogen antibodies and coronary artery disease in rheumatoid arthritis, Arthritis Care & Research, in press, 2017. [19].
 We show that anti-cit-fibrinogen antibodies as a group were associated with CAD outcomes in our RA cohort, with the strongest signal for association arising from a subset of the autoantibodies.
- Liao KP, Sparks JA, Hejblum BP, Kuo IH, Cui J, Lahey LJ, Cagan A, Gainer VS, Liu W, Cai TT, Sokolove J, Cai T, Phenome-wide association study of autoantibodies to citrullinated and noncitrullinated epitopes in rheumatoid arthritis, Arthritis & Rheumatology, 69: 742–749, 2017. [22] We demonstrated application of a bioinformatics method, the PheWAS, to screen for the clinical significance of RA-related autoantibodies. Using the PheWAS approach, we identified potentially significant links between variations in the levels of autoantibodies and comorbidities of interest in RA.

7.7. Conferences

Members of the team were involved in 12 talks during conferences and colloquium.

Mélanie Prague has her work presented in 2017 in 2 peer-reviewed international conferences (Society of clinical trials Liverpool UK and Keystone symposium of mathematical modeling of virus infection, Este Park, May 2017).

Robin Genuer presented his work in the peer-reviewed International Conference of the European Research Consortium for Informatics and Mathematics Working Group (ERCIM WG) on Computational and Methodological Statistics, University of London, UK. Boris Hejblum presented his work in the peer-reviewed 38th Annual Conference of the International Society for Clinical Biostatistics.

Chloé Pasin presented her work in the peer-reviewed Systems Immunology & Vaccine Design symposium, Heidelberg, Germany and the French Applied and Industrial Mathematics Society (SMAI) conference (Ronce-les-Bains).

Members of the team participated in French conferences: GDR Stat santé Bordeaux, GDR mathematical modelling of life Lyon and Journées de la statistique Francaise, Avignon (Perrine Soret, Mélanie Prague, Boris Hejblum). Mélanie Prague and Boris Hejblum also presented 4 posters in workshops.

HIEPACS Project-Team

7. New Results

7.1. High-performance computing on next generation architectures

7.1.1. Bridging the gap between OpenMP and task-based runtime systems

With the advent of complex modern architectures, the low-level paradigms long considered sufficient to build High Performance Computing (HPC) numerical codes have met their limits. Achieving efficiency, ensuring portability, while preserving programming tractability on such hardware prompted the HPC community to design new, higher level paradigms while relying on runtime systems to maintain performance. However, the common weakness of these projects is to deeply tie applications to specific expert-only runtime system APIs. The OpenMP specification, which aims at providing common parallel programming means for shared-memory platforms, appears as a good candidate to address this issue thanks to the latest task-based constructs introduced in its revision 4.0. The goal of this paper is to assess the effectiveness and limits of this support for designing a high-performance numerical library, ScalFMM, implementing the fast multipole method (FMM) that we have deeply redesigned with respect to the most advanced features provided by OpenMP 4. We show that OpenMP 4 allows for significant performance improvements over previous OpenMP revisions on recent multicore processors and that extensions to the 4.0 standard allow for strongly improving the performance, bridging the gap with the very high performance that was so far reserved to expert-only runtime system APIs. More details on this work can be found in [17].

7.1.2. Modeling Irregular Kernels of Task-based codes: Illustration with the Fast Multipole Method

The significant increase of the hardware complexity that occurred in the last few years led the high performance community to design many scientific libraries according to a task-based parallelization. The modeling of the performance of the individual tasks (or kernels) they are composed of is crucial for facing multiple challenges as diverse as performing accurate performance predictions, designing robust scheduling algorithms, tuning the applications, etc. Fine-grain modeling such as emulation and cycle-accurate simulation may lead to very accurate results. However, not only their high cost may be prohibitive but they furthermore require a high fidelity modeling of the processor, which makes them hard to deploy in practice. In this paper, we propose an alternative coarse-grain, empirical methodology oblivious to both the target code and the hardware architecture, which leads to robust and accurate timing predictions. We illustrate our approach with a task-based Fast Multipole Method (FMM) algorithm, whose kernels are highly irregular, implemented in the **ScalFMM** library on top of the **StarPU** task-based runtime system and the simgrid simulator. More details on this work can be found in [41].

7.1.3. Task-based fast multipole method for clusters of multicore processors

Most high-performance, scientific libraries have adopted hybrid parallelization schemes - such as the popular MPI+OpenMP hybridization - to benefit from the capacities of modern distributed-memory machines. While these approaches have shown to achieve high performance, they require a lot of effort to design and maintain sophisticated synchronization/communication strategies. On the other hand, task-based programming paradigms aim at delegating this burden to a runtime system for maximizing productivity. In this article, we assess the potential of task-based fast multipole methods (FMM) on clusters of multicore processors. We propose both a hybrid MPI+task FMM parallelization and a pure task-based parallelization where the MPI communications are implicitly handled by the runtime system. The latter approach yields a very compact code following a sequential task-based programming model. We show that task-based approaches can compete with a hybrid MPI+OpenMP highly optimized code and that furthermore the compact task-based scheme fully matches the performance of the sophisticated, hybrid MPI+task version, ensuring performance while maximizing productivity. We illustrate our discussion with the ScalFMM FMM library and the StarPU runtime system. More details on this work can be found in [40].

7.1.4. Achieving high-performance with a sparse direct solver on Intel KNL

The need for energy-efficient high-end systems has led hardware vendors to design new types of chips for general purpose computing. However, designing or porting a code tailored for these new types of processing units is often considered as a major hurdle for their broad adoption. In this paper, we consider a modern Intel Xeon Phi processor, namely the Intel Knights Landing (KNL) and a numerical code initially designed for a classical multi-core system. More precisely, we consider the qr_mumps scientific library implementing a sparse direct method on top of the StarPU runtime system. We show that with a portable programming model (task-based programming), a good software support (a robust runtime system coupled with an efficient scheduler) and some well defined hardware and software settings, we are able to transparently run the exact same numerical code. This code not only achieves very high performance (up to 1 TFlop/s) on the KNL but also significantly outperforms a modern Intel Xeon multi-core processor both in terms of time to solution and energy efficiency up to a factor of 2.0. More details on this work can be found in [42].

7.2. High performance solvers for large linear algebra problems

7.2.1. Blocking strategy optimizations for sparse direct linear solver on heterogeneous architectures

The preprocessing steps of sparse direct solvers, ordering and block-symbolic factorization, are two major steps that lead to a reduced amount of computation and memory and to a better task granularity to reach a good level of performance when using BLAS kernels. With the advent of GPUs, the granularity of the block computation became more important than ever. In this paper, we present a reordering strategy that increases this block granularity. This strategy relies on the block-symbolic factorization to refine the ordering produced by tools such as METIS or Scotch, but it does not impact the number of operations required to solve the problem. We integrate this algorithm in the PaStiX solver and show an important reduction of the number of off-diagonal blocks on a large spectrum of matrices. This improvement leads to an increase in efficiency of up to 20% on GPUs.

These contributions have been published in SIAM Journal on Matrix Analysis and Applications [22].

7.2.2. Sparse supernodal solver using block low-rank compression

In the context of FASTLA associate team, during the last 4 years, we are collaborating with Eric Darve, professor in the Institute for Computational and Mathematical Engineering and the Mechanical Engineering Department at Stanford, on the design of a new efficient sparse direct solvers. We have been working on applying fast direct solvers for dense matrices to the solution of sparse direct systems. We observed that the extend-add operation (during the sparse factorization) is the most time-consuming step. We have therefore developed a series of algorithms to reduce this computational cost.

We presented two approaches using a Block Low-Rank (BLR) compression technique to reduce the memory footprint and/or the time-to-solution of the sparse supernodal solver PaStiX. This flat, non-hierarchical, compression method allows to take advantage of the low-rank property of the blocks appearing during the factorization of sparse linear systems, which come from the discretization of partial differential equations. The first approach, called *Minimal Memory*, illustrates the maximum memory gain that can be obtained with the BLR compression method, while the second approach, called *Just-In-Time*, mainly focuses on reducing the computational complexity and thus the time-to-solution. Singular Value Decomposition (SVD) and Rank-Revealing QR (RRQR), as compression kernels, are both compared in terms of factorization time, memory consumption, as well as numerical properties. Experiments on a single node with 24 threads and 128 GB of memory are performed to evaluate the potential of both strategies. On a set of matrices from real-life problems, we demonstrate a memory footprint reduction of up to 4 times using the *Minimal Memory* strategy and a computational time speedup of up to 3.5 times with the *Just-In-Time* strategy. Then, we study the impact of configuration parameters of the BLR solver that allowed us to solve a 3D laplacian of 36 million unknowns a single node, while the full-rank solver stopped at 8 million due to memory limitation.

66

These contributions have been presented at the PDSEC workshop of IPDPS'17 conference [30] and an extended version has been submitted in Journal of Computational Science [48].

7.2.3. Towards a hierachical symbol factorization for data sparse direct solvers

Hierarchical algorithms based on low-rank compression techniques have led to fully re-design the methods of solving dense linear systems at the dawn of the twenty-first century, significantly reducing the computational costs. However, their application to the treatment of sparse linear systems remains today a major challenge to which both the community of hierarchical matrices and that of the sparse matrices are tackling. For this purpose, a first class of approach has been developed by the community of hierarchical matrices to exploit the sparse matrix structure. If the strong point of these methods is that the resulting algorithm remains hierarchical, these do not manage exploit some zeros as naturally do sparse solvers. In contrast, the fact that a sparse factorization can be seen as a sequence of smaller, dense operations, the community of hollow treasure has explored this property to introduce hierarchical techniques within these elementary operations. However, the resulting algorithm loses the fundamental property of hierarchical algorithms, since the compression hierarchy is only local. As part of this doctorate, we introduce a new algorithm, performing a sparse hierarchical symbolic factorization that allows to exploit precisely the sparse structure the matrix and its factors while preserving a global hierarchical structure for to ensure effective compression. We have shown experimentally that this new approach allows us to obtain at the same time a reduced number of operations (because of its hierarchical character) and a number of non-zero elements as small as a hollow method (through the use of a symbolic factorization).

This work is developped in the A. Falco PhD thesis, it led to a publication in a national conference [31] and will give rise to a submission in an international journal in 2018

7.3. High performance fast multipole method for N-body problems

7.3.1. Modeling Irregular Kernels of Task-based codes

The significant increase of the hardware complexity that occurred in the last few years led the high performance community to design many scientific libraries according to a task-based parallelization. The modeling of the performance of the individual tasks (or kernels) they are composed of is crucial for facing multiple challenges as diverse as performing accurate performance predictions, designing robust scheduling algorithms, tuning the applications, etc. Fine-grain modeling such as emulation and cycle-accurate simulation may lead to very accurate results. However, not only their high cost may be prohibitive but they furthermore require a high fidelity modeling of the processor, which makes them hard to deploy in practice. In this paper, we propose an alternative coarse-grain, empirical methodology oblivious to both the target code and the hardware architecture, which leads to robust and accurate timing predictions. We illustrate our approach with a task-based Fast Multipole Method (FMM) algorithm, whose kernels are highly irregular, implemented in the **ScalFMM** library on top of the starpu task-based runtime system and the simgrid simulator. More details on this work can be found in [41].

7.3.2. Task-based fast multipole method for clusters of multicore processors

Most high-performance, scientific libraries have adopted hybrid parallelization schemes - such as the popular MPI+OpenMP hybridization - to benefit from the capacities of modern distributed-memory machines. While these approaches have shown to achieve high performance, they require a lot of effort to design and maintain sophisticated synchronization/communication strategies. On the other hand, task-based programming paradigms aim at delegating this burden to a runtime system for maximizing productivity. In this article, we assess the potential of task-based fast multipole methods (FMM) on clusters of multicore processors. We propose both a hybrid MPI+task FMM parallelization and a pure task-based parallelization where the MPI communications are implicitly handled by the runtime system. The latter approach yields a very compact code following a sequential task-based programming model. We show that task-based approaches can compete with a hybrid MPI+OpenMP highly optimized code and that furthermore the compact task-based scheme fully matches the performance of the sophisticated, hybrid MPI+task version, ensuring performance while

maximizing productivity. We illustrate our discussion with the ScalFMM FMM library and the StarPU runtime system. More details on this work can be found in [40].

7.4. Efficient algorithmic for load balancing and code coupling in complex simulations

7.4.1. Comparison of initial partitioning methods for multilevel direct k-way graph partitioning with fixed vertices

In scientific computing, load balancing is a crucial step conditioning the performance of large-scale applications. In this case, an efficient decomposition of the workload to a number of processors is highly necessary. A common approach to solve this problem is to use graph representation and perform a graph partitioning in k parts using the multilevel framework and the recursive bisection (RB) paradigm. However, in graph instances where fixed vertices are used to model additional constraints, RB often produces partitions of poor quality. In this paper, we investigate the difficulties of RB to handle fixed vertices and we compare its results with two different alternatives. The first one, called KGGGP is a direct k-way greedy graph growing partitioning that properly handles fixed vertices while the second one, introduced in kPaToH, uses RB and a post-processing technique to correct the obtained partition. Finally, experimental results on graphs that represent real-life numerical simulations show that both alternative methods provide improved partitions compared to RB. More details on this work can be found in [23].

7.5. Application Domains

7.5.1. Material physics

7.5.1.1. EigenSolver

The adaptive vibrational configuration interaction algorithm has been introduced as a new method to efficiently reduce the dimension of the set of basis functions used in a vibrational configuration interaction process. It is based on the construction of nested bases for the discretization of the Hamiltonian operator according to a theoretical criterion that ensures the convergence of the method. In the present work, the Hamiltonian is written as a sum of products of operators. The purpose of this paper is to study the properties and outline the performance details of the main steps of the algorithm. New parameters have been incorporated to increase flexibility, and their influence has been thoroughly investigated. The robustness and reliability of the method are demonstrated for the computation of the vibrational spectrum up to 3000 cm-1 of a widely studied 6-atom molecule (acetonitrile). Our results are compared to the most accurate up to date computation; we also give a new reference calculation for future work on this system. The algorithm has also been applied to a more challenging 7-atom molecule (ethylene oxide). The computed spectrum up to 3200 cm-1 is the most accurate computation that exists today on such systems. More details on this work can be found in [43], [21].

7.5.1.2. Dislocation

We have focused on the improvements of the parallel collision detection and of the accuracy in the force field computation in the **OPTIDIS** code.

- a new collision detection algorithm to reliably handle junction formation for Dislocation Dynamics using hybrid OpenMP + MPI parallelism has been developed. The enhanced precision and reliability of this new algorithm allows the use of larger time-steps for faster simulations. Hierarchical methods for collision detection, as well as hybrid parallelism are also used to improve performance;
- we observed that the force field computation depends on how the traversal of the segments list or boxes in the octree was done. New accurate formulas to remove this issue have been developed and we are implementing them in the code. They will be used in the Fast Multipole Method that we have developed previously.

Finally, a new distributed data structure has been developed to enhance the reliability and modularity of **OPTIDIS**. The new data structure provides an interface to modify safely and reliably the distributed dislocation mesh in order to enforce data consistency across all computation nodes. This interface also improves code modularity allowing the study of data layout performance without modifying the algorithms.

7.5.2. Co-design for scalable numerical algorithms in scientific applications

7.5.2.1. High performance simulation for ITER tokamak

Concerning the GYSELA global non-linear electrostatic code, the efforts during the period have concentrated on the design of a more efficient parallel gyro-average operator for the deployment of very large (future) GYSELA runs. The main unknown of the computation is a distribution function that represents either the density of the guiding centers, either the density of the particles in a tokamak. The switch between these two representations is done thanks to the gyro-average operator. In the previous version of GYSELA, the computation of this operator was achieved thanks to a Padé approximation. In order to improve the precision of the gyroaveraging, a new parallel version based on an Hermite interpolation has been done (in collaboration with the Inria TONUS project-team and IPP Garching). The integration of this new implementation of the gyroaverage operator has been done in GYSELA and the parallel benchmarks have been successful. This work is carried on in the framework of the PhD of Nicolas Bouzat (funded by IPL C2S@EXA) co-advised with Michel Mehrenberger from TONUS project-team and in collaboration with Guillaume Latu from CEA-IRFM. The scientific objectives of this work is first to consolidate the parallel version of the gyro-average operator, in particular by designing a scalable MPI+OpenMP parallel version and using a new communication scheme, and second to design new numerical methods for the gyro-average, source and collision operators to deal with new physics in GYSELA. The objective is to tackle kinetic electron configurations for more realistic complex large simulations.

In the context of the EoCoE project, we have collaborations with CEA-IRFM. First, with G. Latu, we have investigated the potential of using the last release of the PaStiX solver (version 6.0) on Intel KNL architecture, and more especially on the MARCONI machine (one of the PRACE supercomputers at Cineca, Italia). The results obtained on this architecture are really promising since we are able to reach more than 1 Tflops using a single node. Secondly, we also have a collaboration with P. Tamain and G. Giorgani on the TOKAM3X code to analyze the performance of using PaStiX as a preconditioner. Since a distributed memory is required during the simulation, the previous release of PaStiX is then used. Some difficulties regarding the Fortran wrapper and some memory issues should be fixed when we will have reimplemented the MPI interface in the current release.

7.5.2.2. High performance simulation for 3D frequency-domain Maxwell's equations

We also recently developed a collaboration with NACHOS on the HORSE (High Order solver for Radar cross Section Evaluation) simulation code. The aim was to integrate the PaStiX solver, with low-rank compression technique, in a domain decomposition framework to solve 3D frequency-domain Maxwell's equations. The results are promising since we were able to reduce by two the factorization and the solve time for each subdomain. And we were also able to reduce by two the memory requirements thanks to our compression techniques. This would allow us to consider larger subdomains with the same memory constraints that currently limit the simulations.

7.5.2.3. High performance simulation for atmospheric chemistry

We worked on the development and tests of the Adaptative Semi-Implicit Scheme (ASIS) solver for the simulation of atmospheric chemistry. To solve the Ordinary Differential Equation systems associated with the time evolution of the species concentrations, ASIS adopts a one step linearized implicit scheme with specific treatments of the Jacobian of the chemical fluxes. It conserves mass and has a time stepping module to control the accuracy of the numerical solution. In 5 idealized box model simulations ASIS gives results similar to the higher order implicit schemes derived from the Rosenbrock's and Gear's methods and requires less computation and run time at the moderate precision required for atmospheric applications. When implemented in the MOCAGE CTM and the LMD Mars GCM the ASIS solver performs well and reveals weaknesses and limitations of the original semi-implicit solvers used by these two models. ASIS can be easily adapted to

various chemical schemes and further developments are foreseen to increase its computational efficiency, and to include the computation of the 10 concentrations of the species in aqueous phase in addition to gas phase chemistry.

More details on this work can be found in [19].

PHOENIX Project-Team

7. New Results

7.1. Everyday Functioning Benefits from an Assisted Living Platform amongst Frail Older Adults and Their Caregivers

Ambient assisted living technologies (AAL) are regarded as a promising solution to support aging in place. Yet, their efficacy has to be demonstrated in terms of benefits for independent living and for work conditions of caregivers. Hence, the purpose of this study was to assess the benefits of a multi-task AAL platform for both Frail older Individuals (FIs) and professional caregivers with respect to everyday functioning and caregiver burden. In this context, a 6-month field study involved 32 FIs living at home (half of them were equipped by the platform and the remaining half were not, as a control condition) and their caregivers. Everyday functioning measures were reported by frail participants and caregiver's estimates of everyday functioning of equipped participants were unchanged across time, while they decreased for the control participants. Also, a reduction of self-reported objective burden was obtained after 6 months of AAL intervention for the equipped group, compared to the control group. Overall, these results highlighted the potential of AAL as a relevant environmental support for preventing both functional losses in FIs and objective burden professional caregiver.

7.2. Designing Parallel Data Processing for Enabling Large-Scale Sensor Applications

Masses of sensors are being deployed at the scale of cities to manage parking spaces, transportation infrastructures to monitor traffic, and campuses of buildings to reduce energy consumption. These large-scale infrastructures become a reality for citizens via applications that orchestrate sensors to deliver high-value, innovative services. These applications critically rely on the processing of large amounts of data to analyze situations, inform users, and control devices. This work proposes a design-driven approach to developing orchestrating applications for masses of sensors that integrates parallel processing of large amounts of data. Specifically, an application design exposes declarations that are used to generate a programming framework based on the MapReduce programming model. We have developed a prototype of our approach, using Apache Hadoop. We applied it to a case study and obtained significant speedups by parallelizing computations over twelve nodes. In doing so, we demonstrate that our design-driven approach allows to abstract over implementation details, while exposing architectural properties used to generate high-performance code for processing large datasets. Furthermore, we show that this high-performance support enables new, personalized services in a smart city. Finally, we discuss the expressiveness of our design language, identify some limitations, and present language extensions.

7.3. Internet of Things: From Small-to Large-Scale Orchestration

The domain of Internet of Things (IoT) is rapidly expanding beyond research, and becoming a major industrial market with such stakeholders as major manufacturers of chips and connected entities (i.e., things), and fast-growing operators of wide-area networks. Importantly, this emerging domain is driven by applications that leverage an IoT infrastructure to provide users with innovative, high-value services. IoT infrastructures range from small scale (e.g., homes and personal health) to large scale (e.g., cities and transportation systems). In this work, we argue that there is a continuum between orchestrating connected entities in the small and in the large. We propose a unified approach to application development, which covers this spectrum. To do so, we examine the requirements for orchestrating connected entities and address them with domain-specific design concepts. We then show how to map these design concepts into dedicated programming patterns and runtime mechanisms. Our work revolves around domain-specific concepts and notations, integrated into a tool-based design methodology and dedicated to develop IoT applications. We have applied our work across a spectrum of infrastructure sizes, ranging from an automated pilot in avionics, to an assisted living platform for the home of seniors, to a parking management system in a smart city.

71

7.4. Designing an Accessible and Engaging Email Application for Aging in Place

72

Supporting independent everyday functioning of older adults is a major challenge for aging in place. In particular, communication and social activities need support to prevent social isolation, cognitive and psychosocial well-being decline, and a risk of depression. This paper focuses on how technology can bring social support to isolated older-old adults (over 75 years old) and allow them to communicate with members of their social network. We present the design of an accessible and engaging email application dedicated to this population. We propose design principles based on the older adults' specificities and then use these principles to develop a tablet-based email application. We conducted a field study to evaluate our email application during 9 months. We equipped 13 community-dwelling old-older adults with a touchscreen tablet and our application at their home (compared to 13 control counterparts). This field study validates our design principles as shown by the effectiveness and efficiency gained by the participants in using our application. Moreover, we reveal the influence of health indicators in the usage behaviors and the long-term use of our application.

7.5. HomeAssist: An Assisted Living Platform for Aging in Place Based on an Interdisciplinary Approach

HomeAssist is an assisted living platform aims to support aging in place. This platform was designed using a human-centered approach. It offers assistive services, addressing the main aspects of daily life: activities of daily living, home and user safety, and social participation. HomeAssist introduces key novel features: (1) it covers multiple aspects of daily life, addressing a variety of needs of older adults; (2) it provides customization mechanisms, adapting assistance to the user's abilities while preventing autonomy losses; (3) it relies on context awareness, delivering timely assistance; and, (4) it revolves around a unified user interface to achieve usability. All these features play a key role towards achieving high acceptance of HomeAssist and supporting autonomy effectively, as shown by our field study.
STORM Project-Team

6. New Results

6.1. Distributed Sequential Task Flow with StarPU

The emergence of accelerators as standard computing resources on supercomputers and the subsequent architectural complexity increase revived the need for high-level parallel programming paradigms. Sequential task-based programming model has been shown to efficiently meet this challenge on a single multicore node possibly enhanced with accelerators, which motivated its support in the OpenMP 4.0 standard. We showed [5] that this paradigm can also be employed to achieve high performance on modern supercomputers composed of multiple such nodes, with extremely limited changes in the user code. To prove this claim, we have extended the StarPU runtime system with an advanced inter-node data management layer that supports this model by posting communications automatically. We illustrates our discussion with the task-based tile Cholesky algorithm that we implemented on top of this new runtime system layer. We showed that it enables very high productivity while achieving a performance competitive with both the pure Message Passing Interface (MPI)-based ScaLAPACK Cholesky reference implementation and the DPLASMA Cholesky code, which implements another (non-sequential) task-based programming paradigm.

6.2. Distributed StarPU on top of a High Performance Communication Library

A new implementation of the StarPU's distributed engine is being currently developed on top of the New-Madeleine library. The first version of this engine had been written directly on top of MPI. The performance were not as good as expected when dealing with applications exchanging huge number of messages, and we had to implement within StarPU mechanisms to control the memory subscription [14].

NewMadeleine is a high performance communication library for clusters developed in the Tadaam team. It applies optimization strategy on data flows through dynamic packet scheduling, and is usable on various high performance networks. The new implementation of the StarPU's distributed engine no longer has to deal with communication-related issues, and provides a better reactivity as the communications progress is dealt with by NewMadeleine itself. First experiments with the Chameleon solver show promising results.

6.3. Bridging the Gap between a Standard Parallel Language and a Task-based Runtime System

With the advent of complex modern architectures, the low-level paradigms long considered sufficient to build High Performance Computing (HPC) numerical codes have met their limits. Achieving efficiency, ensuring portability, while preserving programming tractability on such hardware prompted the HPC community to design new, higher level paradigms while relying on runtime systems to maintain performance. However, the common weakness of these projects is to deeply tie applications to specific expert-only runtime system APIs. The OpenMP specification, which aims at providing common parallel programming means for sharedmemory platforms, appears as a good candidate to address this issue thanks to the latest task-based constructs introduced in its revision 4.0. We assessed [4] the effectiveness and limits of this support for designing a high-performance numerical library, ScalFMM, implementing the fast multipole method (FMM) that we have deeply re-designed with respect to the most advanced features provided by OpenMP 4. We showed that OpenMP 4 allows for significant performance improvements over previous OpenMP revisions on recent multicore processors and that extensions to the 4.0 standard allow for strongly improving the performance. bridging the gap with the very high performance that was so far reserved to expert-only runtime system APIs. Our proposal for an OpenMP extension to let the programmer express the property of commutativity between multiple tasks has been presented by Inria and successfully voted-on and integrated as the notion of mutually exclusive input/output sets (mutexinoutset keyword) in OpenMP ARB's Technical Report 6: OpenMP Version 5.0 Preview 2, the last pre-version of the upcoming OpenMP 5.0 specification.

6.4. Combining a Component Model and a Task Parallelism Model

We demonstrated the feasibility of efficiently combining both a software component model and a task-based model [6]. Task based models are known to enable efficient executions on recent HPC computing nodes while component models ease the separation of concerns of application and thus improve their modularity and adaptability.

This paper describes a prototype version of the COMET programming model combining concepts of taskbased and component models, and a preliminary version of the COMET runtime built on top of StarPU and L2C. Evaluations of the approach have been conducted on a real-world use-case analysis of a sub-part of the production application GYSELA.

Results show that the approach is feasible and that it enables easy composition of independent software codes without introducing overheads. Performance results are equivalent to those obtained with a plain OpenMP based implementation.

6.5. Tackling the granularity problem

One of the main issues encountered when trying to exploit both CPUs and accelerators is that these devices have very different characteristics and requirements. Indeed, GPUs typically exhibit better performance when executing kernels applied to large data sets while regular CPU cores reach their peak performance with fine grain kernels working on a reduced memory footprint. To work around this granularity problem, task-based applications running on such heterogeneous platforms typically adopt a medium granularity, chosen as a trade-off between coarse-grain and fine-grain kernels. To tackle this granularity problem, we investigated different complementary technics. The first two technics are based on StarPU, performing both load-balancing and scheduling, the third one splits automatically kernels at compile-time and then performs load-balancing.

- The first technic is based on resource aggregation : we aggregate CPU cores to execute coarse grain tasks in a parallel manner. We have showed that this technic for a dense Cholesky factorization kernel outperforms state of the art implementations on a platform equipped with 24 CPU cores and 4 GPU devices (reaching a peak performance of 4.8 TFlop/s) and on the Intel KNL processor (reaching a peak performance 1.58 TFlop/s).
- The second technic splits dynamically coarse grain tasks when they are assigned to CPU cores. Tasks can be replaced by a subgraph of tasks of finer granularity, allowing for a finer handling of depencencies and a better pipelining of kernels. This mechanism allowing to deal with hierarchical task graphs has been designed within StarPU. Moreover, it allows to parallelize the task submission flow while preserving the simplicity of the sequential task flow submission paradigm. First experimental results for dense Cholesky factorization kernel show good performance improvements with respect to the native StarPU's implementation.
- The third technic extends our previous work that provides an automatic compiler and runtime technique to execute single OpenCL kernels on heterogeneous multi-device architectures. Our technique splits computation and data automatically across the multiple computing devices. OpenCL applications that consist in a chain of data-dependent kernels in an iterative computation are now considered.

The technique proposed is completely transparent to the user, and does not require off-line training or a performance model. It manages sometimes conflicting needs between load balancing each kernel in the chain and minimizing data transfer between consecutive kernels, taking data locality into account. Load-balancing issues, resulting from hardware heterogeneity, load imbalance within a kernel itself, and load variations between repeated executions are also managed.

Experiments on some benchmarks show the interest of our approach and we are currently implementing it in an OpenCL N-body computation with short-range interactions.

6.6. Interfacing MAQAO and BOAST Frameworks for Kernel Autotuning on ARM Platforms

In Project MontBlanc 2's deliverable D5.11 [13] we presented the integration of STORM's MAQAO software (a binary-level code analysis framework) with BOAST (an automatic performance tuning framework for metaprogramming and optimizing computing kernels) developed at LIG's NANOSIM. From source meta-kernels written in the RUBY language, BOAST generates multiple versions, in various target languages, optionally applying optimization transformations and strategies, and exploring the space of compiler flags combinations, to discover the most effective kernel tuning parameters. MAQAO offers a scriptable framework to disassemble kernel binaries, explore binary instruction flows, register-level data dependencies, program control structures, to patch, re-assemble and instrument kernel binaries for tracing data access patterns, and to process them from custom analyzers written in the LUA language. This integration work built on the complementarity of these two environments by enabling MAQAO to process binary kernels generated by BOAST, and lead developers in a guided tuning cycle.

6.7. Using heterogeneous memories

Heterogeneous memories, such as the MCDRAM in the Xeon Phi architecture, with different latency and bandwidth characteristics, complexify the way the users allocate and use memory. In 2017, we have designed, in collaboration with CEA, an automatic tool to characterize the bandwidth needs of an application, in particular finding the functions and the arrays in these functions that would benefit the most of a high bandwidth. This tool is a plugin of gcc, and has been applied successfully to large CORAL benchmarks (Lulesh, MiniFE, AMG2013, Mcb and Snap). This characterization is essential in the common case where all data cannot fit into the MCDRAM but a more selective use of the MCDRAM is needed. The transformation of the memory allocations is automatic, based on these metrics. The development of better metrics, allowing to choose the most appropriate array is on going work.

6.8. Rewriting System for Profile-Guided Data Layout Transformations on Binaries

Careful data layout design is crucial for achieving high performance. However exploring data layouts is time-consuming and error-prone, and assessing the impact of a layout transformation on performance is difficult without performing it. We proposed [7] a method and implemented a prototype to guide application programmers through data layout restructuring for improving kernel performance and SIMDizability, by providing a comprehensive multidimensional description of the initial layout, built from trace analysis, and then by giving a performance evaluation of the transformations tested and an expression of each transformed layout. The programmer can limit the exploration to layouts matching some patterns. We apply this method to two multithreaded applications. The performance prediction of multiple transformations matches within 5% the performance of hand-transformed layout code.

6.9. Correctness of HPC Applications

The current supercomputer hardware trends lead to more complex HPC applications (heterogeneity in hardware and combinations of parallel programming models) that pose programmability challenges. Furthermore, progress to exascale stresses the requirement for convenient and scalable debugging methods to help developers fully exploit the future machines. Despite advances in the domain, this still remains a manual complex task. We aim to develop tools and methods to aid developers with problems of correctness in HPC applications for exascale systems. There are several requirements for such tools: 1) precision - report and handle only real problems, areas of interest; 2) scalability in LoCs and execution time; 3) heterogeneity - ability to handle multiple languages, runtime and execution models; and 4) soundness - ability to prove code properties. In order to improve developer productivity, we aim to develop a combination of static and dynamic analyses. Static analysis techniques will enable soundness and scalability in execution time. Dynamic analysis techniques will enable precision, scalability in LoCs and heterogeneity for hybrid parallelism. The achieved results this year allow to perform an interprocedural static data- and control-flow analysis: its improves precision, by only detecting possible correctness issues related to MPI rank dependent variables. It improves scalability also by reducing the amount of dead-lock avoiding code added. This new method has been applied to CUDA, MPI, OpenMP and UPC parallel codes to detect collective deadlocks.

6.10. AMR-Based Dynamic Load Balancing for Molecular Dynamics Simulations

Modern parallel architectures require applications to generate enough parallelism to feed many cores, which require in turn regular data-parallel instructions to exploit large vector units. We revisit the extensively-studied Classical Molecular Dynamics N-body problem in the light of these hardware constraints. A new data layout is proposed with efficient force computation methods focusing on adaptive mesh refinement techniques, multi-threading, vectorization-friendly, using low memory footprint. Our design is guided by the need for load balancing and adaptivity raised by highly dynamic particle sets, as typically observed in simulations of strong shocks resulting in material micro-jetting. We analyze performance results on several simulation scenarios, over clusters equipped with Intel Xeon Phi knl processors. Performance obtained with our implementation using OpenMP is close to state-of-the-art implementations (LAMMPS) using MPI on steady particles simulations, and outperform them by 1.2 on micro-jetting simulations on Intel Xeon Phi (KNL).

6.11. Resource-Management Study in HPC Runtime-Stacking Context

With the advent of multicore and manycore processors as building blocks of HPC supercomputers, many applications shift from relying solely on a distributed programming model (e.g., MPI) to mixing distributed and shared-memory models (e.g., MPI+OpenMP), to better exploit shared-memory communications and reduce the overall memory footprint. One side effect of this programming approach is runtime stacking: mixing multiple models involve various runtime libraries to be alive at the same time and to share the underlying computing resources. This paper explores different configurations where this stacking may appear and introduces algorithms to detect the misuse of compute resources when running a hybrid parallel application. We have implemented our algorithms inside a dynamic tool that monitors applications and outputs resource usage to the user. We validated this tool on applications from CORAL benchmarks. This leads to relevant information which can be used to improve runtime placement, and to an average overhead lower than 1% of total execution time.

TADaaM Project-Team

7. New Results

7.1. Network Modeling

NETLOC (see Section 6.3) is a tool in HWLOC to discover the network topology. The information gathered and analysed are now saved in XML format. It brings more flexibility, readability and compatibility. Henceforth, in the display tool, we compute the positions of the nodes rather than use physics algorithm provided by vis.js library for node placement. Thus, it makes the visualization faster and we can display a fat-tree with around 41k nodes in less than 1 second.

Moreover, we can deal with other kinds of topologies. We handle topologies in a generic way and can have nested topologies. For the mapping, we build a deco graph in SCOTCH. Consequently, the mapping will be possible for any architecture. [17]

We have also optimized the mapping by giving a preconditioned matrix to SCOTCH, and by computing some metrics in order to evaluate mappings and keep the best one.

The part about discovering network have been improved and we support now, in addition to Infiniband, Omnipath fat-trees, Cray Torus.

7.2. Locality Aware Roofline Model

The trend of increasing the number of cores on-chip is enlarging the gap between compute power and memory performance. This issue leads to design systems with heterogeneous memories, creating new challenges for data locality. Before the release of those memory architectures, the Cache-Aware Roofline Model [47] (CARM) offered an insightful model and methodology to improve application performance with knowledge of the cache memory subsystem.

With the help of hwloc library, we are able to leverage the machine topology to extend the CARM for modeling NUMA and heterogeneous memory systems, by evaluating the memory bandwidths between all combinations of cores and NUMA nodes. The new Locality Aware Roofline Model [19] (LARM) scopes most contemporary types of large compute nodes and characterizes three bottlenecks typical of those systems, namely contention, congestion and remote access.

This work has been achieved in collaboration with the authors of the CARM and the source code of the associated tool is publicly available at https://github.com/NicolasDenoyelle/Locality-Aware-Roofline-Model.

In the future we plan to design and embed in the model an hybrid memory bandwidth model to provide an automatic roof matching feature.

7.3. Scalable Management of Platform Topologies

HWLOC (see Section 6.2) is used for gathering the topology of computing nodes. Those nodes are now growing to hundreds of cores, making the overall amount of topology information non-negligible. We studied the overhead of topology discovery on the overall execution time and showed that the Linux kernel is bottleneck on large nodes. It raised the need to use exported and/or abstracted topologies to factorize this overhead [22].

The memory footprint of locality information is also becoming an issue on large many-core. We designed a way to share this information between processes inside nodes so as to factorize this memory consumption [45].

7.4. New algorithm for I/O scheduling

We started working on I/O scheduling for HPC applications. HPC applications can be caracterized by I/O patterns that are repeated periodically. We showed in a simple context how this information can be taken into account to outperform state of the art I/O schedulers [15].

These preliminary results led to the obtention of the ANR DASH (see Section 9.1.2).

After which, we have performed a theoretical analysis to show how one should size the burst-buffers and the bandwidth to those buffers on a HPC system depending on the applications running. In our study we focused on one role of the buffers (namely the role of buffer to the PFS) [42]. This study is particularly important since those buffers are limited and can be used for many usage. Over or under booking the buffers for a specific use leads to an increase of congestion.

7.5. Topology-Aware Data Aggregation on Large-Scale Supercomputers

We have continue our work on on two-phase i/O and data aggregation. This strategy consists of selecting a subset of processes to aggregate contiguous pieces of data before performing reads/writes. In collaboration with Argonne National Lab, we have worked on TAPIOCA, an MPI-based library implementing an efficient topology-aware two-phase I/O algorithm. TAPIOCA can take advantage of double-buffering and one-sided communication to reduce as much as possible the idle time during data aggregation. We also introduce our cost model leading to a topology-aware aggregator placement optimizing the movements of data. We validate our approach at large scale on two leadership-class supercomputers: Mira (IBM BG/Q) and Theta (Cray XC40). On BG/Q+GPFS, for instance, our algorithm leads to a performance improvement by a factor of twelve while on the Cray XC40 system associated with a Lustre filesystem, we achieve an improvement of four [27]

7.6. Empirical Study of the Impact on Performance of Process Affinity and Metrics

Process placement, also called topology mapping, is a well-known strategy to improve parallel program execution by reducing the communication cost between processes. It requires two inputs: the topology of the target machine and a measure of the affinity between processes. In the literature, the dominant affinity measure is the communication matrix that describes the amount of communication between processes. We have studied the accuracy of the communication matrix as a measure of affinity. We have done an extensive set of tests with two fat-tree machines and a 3d-torus machine to evaluate several hypotheses that are often made in the literature and to discuss their validity. First, we check the correlation between algorithmic metrics and the performance of the application. Then, we check whether a good generic process placement algorithm never degrades performance. And finally, we see whether the structure of the communication matrix can be used to predict gain.

7.7. Automatic, Abstracted and Portable Topology-Aware Thread Placement

Efficiently programming shared-memory machines is a difficult challenge because mapping application threads onto the memory hierarchy has a strong impact on the performance. However, optimizing such thread placement is difficult: architectures become increasingly complex and application behavior changes with implementations and input parameters, e.g problem size and number of threads.We have worked on a fully automatic, abstracted and portable affinity module. It produces and implements an optimized affinity strategy that combines knowledge about application characteristics and the platform topology. Implemented in the back-end of our runtime system (ORWL), our approach was used to enhance the performance and the scalability of several unmodified ORWL-coded applications [23]

7.8. Process Placement with TreeMatch

We released TREEMATCH version 1.0 in June. The new feature are: a stabilize API, optional integration of SCOTCH, extensive testing of all the features.

7.9. Managing StarPU Communications with NewMadeleine

We have worked on the scalability with the number of communication requests in the NewMadeleine 6.4 communication library, so as to be able to manage communication patterns from the StarPU runtime. We have ported [44] StarPU on top of NewMadeleine so as to take benefit from NewMadeleine scalability in StarPU. Preliminary results are encouraging.

7.10. New abstraction to manage hardware topologies in MPI applications

Since the end of year 2016, we have been working on new abstractions and mechanisms that can allow the programmer to take advantage of the underlying hardware topology in their parallel applications developed in MPI. For instance, taking into account the intricate network/memory hierarchy can lead to substantial improvements in communication performance and reduce altogether the overall execution time of the application. However, it is important to find the relevant level of abstraction, as too much details are not usable practically because the programmer is not a a hardware specialist most of the time. Also, MPI being hardware-agnostic, it is important to find means to use the hardware specifics without being tied to a particular architecture or hardware design.

With these goals in mind, we proposed the HSPLIT (see Section 6.1) libray that implements a solution based on a well-known MPI concept, the *communicators* (that can be seen as groups of communicating processes). With HSPLIT, each level in the hardware hierarchy is accessible through a dedicated communicator. In this way, the programmer can leverage the underlying hierarchy in their application quite simply. The current implementation of HSPLIT is based on both HWLOC and NETLOC.

This work led to the creation of a new active working group within the MPI Forum, coordinated and lead by Inria.

7.11. Empirical Study of the Impact on Performance of Process Affinity and Metrics

Process placement, also called topology mapping, is a well-known strategy to improve parallel program execution by reducing the communication cost between processes. It requires two inputs: the topology of the target machine and a measure of the affinity between processes. In the literature, the dominant affinity measure is the communication matrix that describes the amount of communication between processes. We have studied the accuracy of the communication matrix as a measure of affinity. We have done an extensive set of tests with two fat-tree machines and a 3d-torus machine to evaluate several hypotheses that are often made in the literature and to discuss their validity. First, we check the correlation between algorithmic metrics and the performance of the application. Then, we check whether a good generic process placement algorithm never degrades performance. And finally, we see whether the structure of the communication matrix can be used to predict gain [35].

7.12. Gradient reconstruction in a legacy CFD application using task-based programming models

We investigated different runtime systems, namely StarPU and PaRSEC and their use in a legacy CFD code from EDF R&D. We assessed both runtimes in terms of performance, ease of implementation and various others criterion such as maintainability, documentation and team activity. By experimenting these solutions out of classical linear algebra problems, we push them out of their comfort zone into the common issues seen in Computational Fluid Dynamics codes with unstructured meshes [30].

7.13. Efficient multi-constraint graph partitioning algorithms

Although several tools provide multi-constraint graph partitioning features, this problem had not been thoroughly investigated. In the context of the PhD of Rémi Barat, several significant results were achieved regarding the multi-constraint graph partitioning problem.

Firstly, a theoretical analysis of the solution space of the mono-criterion, balanced graph bipartitioning problem showed that this space is strongly connected. Hence, local optimization algorithms may indeed succeed in finding paths to better solutions, from some existing solution. A conjecture on the multi-criteria case has been derived. These findings reversed our view on partitioning: while most tools try to find a possibly unbalanced partition of small cut, and then try to rebalance it, it is in fact possible to compute a balanced partition of arbitrary cut, and then to improve the cut.

Secondly, a thorough investigation of the multilevel framework, and of its implementations in several existing tools, allowed us to define the characteristics of an effective coarsening method, both in the mono-criterion and multi-criteria case. Also, new multi-criteria graph algorithms were designed for the initial partitioning and local optimization phases of the multilevel framework [43]. A new data structure has been devised, which speeds-up the computation of balanced partitions in the multi-criteria case.

Thirdly, all of the aforementioned algorithms were implemented in a prototype version of SCOTCH.

7.14. Progress threads placement for MPI Non-Blocking Collectives

MPI Non-Blocking Collectives (NBC) allow for communication overlap with computation. A good overlapping ratio is obtained when computation and communication are running in parallel. To achieve this, each MPI task generates a progress thread to manage communication tasks. The progression of these communications requires regular access to the processors. These threads compete with each other and with MPI tasks. In order to run threads with minimal disruption, we bound the progress threads on free cores when it is possible. Then, we showed that folding all progress threads on very few cores does not work for tree algorithms. The number of communication generated are too important. The solution that we propose is to perform a number of levels (S) of the dependency tree on MPI tasks. We get a reasonable execution time (less than compute time + communication time) while reserving fewer cores for progress threads. All these methods have been implemented in the MPC framework, which contributes to its development.

7.15. Use of PaMPA on large-scale simulations

Many improvements have been brought to PaMPA this year, to improve its robustness and scalability, and to extend its features. In the context of a joint work with CERFACS, PaMPA was subsequently used to remesh the mesh of a helicopter turbine combustion chamber, up to 1 billion elements. This allowed to run a Large-Eddy Simulation (LES) simulation that was out of reach of previous state-of-the-art remeshing software [38].

7.16. Co-scheduling applications on cache-partitioned systems

Cache-partitioned architectures allow subsections of the shared last-level cache (LLC) to be exclusively reserved for some applications. This technique dramatically limits interactions between applications that are concurrently executing on a multi- core machine. We have provided efficient algorithms to co-schedule multiple applications on cache-partitioned systems and evaluations showing that they performed well [13], [6]. We are currently in the process of evaluating them on real machines.

7.17. Dynamic memory-aware task-tree scheduling

We have provided new efficient algorithms that can be used for sparse matrices factorizations under memory constraints. We provide speedup of 15 to 45% over existing strategies and we are working on an actual implementation in QR-MUMPS [14].

FLOWERS Project-Team

7. New Results

7.1. Computational Models Of Human Development and Cognition

7.1.1. Computational Models Of Information-Seeking and Curiosity-Driven Learning in Humans and Animals

Participants: Pierre-Yves Oudeyer [correspondant], William Schueller, Sebastien Forestier, Alvaro Ovalle.

This project involves a collaboration between the Flowers team, the Cognitive Neuroscience Lab of J. Gottlieb at Columbia Univ. (NY, US), and the developmental psychology lab of Celeste Kidd at Univ. Rochester, US, on the understanding and modeling of mechanisms of curiosity, attention and active intrinsically motivated exploration that until now have been little explored in neuroscience, machine learning and cognitive robotics.

It is organized around the study of the hypothesis that information gain (or control gain) could generate intrinsic reward in the brain (living or artificial), driving attention and exploration independently from material rewards, and allowing for autonomous lifelong acquisition of open repertoires of skills. The project combines expertise about attention and exploration in the brain and a strong methodological framework for conducting experimentations with monkeys, human adults (Gottlieb's lab) and children (Kidd's lab) together with computational modeling of curiosity/intrinsic motivation and learning in the Flowers team.

Such a collaboration paves the way towards a central objective, which is now a central strategic objective of the Flowers team: designing and conducting experiments in animals and humans informed by computational/mathematical theories of information seeking, and allowing to test the predictions of these computational theories.

7.1.1.1. Context

Curiosity can be understood as a family of mechanisms that evolved to allow agents to maximize their knowledge (or their control) of the useful properties of the world - i.e., the regularities that exist in the world - using active, targeted investigations. In other words, we view curiosity as a decision process that maximizes learning/competence progress (rather than minimizing uncertainty) and assigns value ("interest") to competing tasks based on their epistemic qualities - i.e., their estimated potential allow discovery and learning about the structure of the world.

Because a curiosity-based system acts in conditions of extreme uncertainty (when the distributions of events may be entirely unknown) there is in general no optimal solution to the question of which exploratory action to take [31], [152], [160]. Therefore we hypothesize that, rather than using a single optimization process as it has been the case in most previous theoretical work [120], curiosity is comprised of a family of mechanisms that include simple heuristics related to novelty/surprise and measures of learning progress over longer time scales [150] [98], [146]. These different components are related to the subject's epistemic state (knowledge and beliefs) and may be integrated with fluctuating weights that vary according to the task context. We will quantitatively characterize this dynamic, multi-dimensional system in the framework of Bayesian Reinforcement Learning, as described below.

Because of its reliance on epistemic currencies, curiosity is also very likely to be sensitive to individual differences in personality and cognitive functions. Humans show well-documented individual differences in curiosity and exploratory drives [137], [159], and rats show individual variation in learning styles and novelty seeking behaviors [115], but the basis of these differences is not understood. We postulate that an important component of this variation is related to differences in working memory capacity and executive control which, by affecting the encoding and retention of information, will impact the individual's assessment of learning, novelty and surprise and ultimately, the value they place on these factors [156], [168], [93], [174]. To start understanding these relationships, about which nothing is known, we will search for correlations between curiosity and measures of working memory and executive control in the population of children we test in our tasks, analyzed from the point of view of a computational model based on Bayesian reinforcement learning.

A final premise guiding our research is that essential elements of curiosity are shared by humans and non-human primates. Human beings have a superior capacity for abstract reasoning and building causal models, which is a prerequisite for sophisticated forms of curiosity such as scientific research. However, if the task is adequately simplified, essential elements of curiosity are also found in monkeys [137], [132] and, with adequate characterization, this species can become a useful model system for understanding the neurophysiological mechanisms.

7.1.1.2. Objectives

Our studies have several highly innovative aspects, both with respect to curiosity and to the traditional research field of each member team.

- Linking curiosity with quantitative theories of learning and decision making: While existing investigations examined curiosity in qualitative, descriptive terms, here we propose a novel approach that integrates quantitative behavioral and neuronal measures with computationally defined theories of Bayesian Reinforcement Learning and decision making.
- Linking curiosity in children and monkeys: While existing investigations examined curiosity in humans, here we propose a novel line of research that coordinates its study in humans and non-human primates. This will address key open questions about differences in curiosity between species, and allow access to its cellular mechanisms.
- Neurophysiology of intrinsic motivation: Whereas virtually all the animal studies of learning and decision making focus on operant tasks (where behavior is shaped by experimenter-determined primary rewards) our studies are among the very first to examine behaviors that are intrinsically motivated by the animals' own learning, beliefs or expectations.
- Neurophysiology of learning and attention: While multiple experiments have explored the singleneuron basis of visual attention in monkeys, all of these studies focused on vision and eye movement control. Our studies are the first to examine the links between attention and learning, which are recognized in psychophysical studies but have been neglected in physiological investigations.
- Computer science: biological basis for artificial exploration: While computer science has proposed and tested many algorithms that can guide intrinsically motivated exploration, our studies are the first to test the biological plausibility of these algorithms.
- Developmental psychology: linking curiosity with development: While it has long been appreciated that children learn selectively from some sources but not others, there has been no systematic investigation of the factors that engender curiosity, or how they depend on cognitive traits.

7.1.1.3. Current results

In particular, new works and results in 2017 include:

7.1.1.4. Experiments in Active Categorization

In 2017, we have been occupied by the implementation, running and analysis of the human adult experiment piloted the year before. A distinguishing feature of curiosity is that, rather than seeking to obtain information in a known task context (e.g., reading the menu in a restaurant) curiosity has to discover regularities whose existence is a priori unknown. This raises the question of how active learners become interested in specific items: how do agents decide which task to be interested in – i.e., allocate "study time" - given that the underlying rewards or patterns are sparse and unknown? A theoretical solution to this problem is suggested by the optimal learning literature, and proposes that allocation of resources may be based on the relative difficulty of competing tasks , or the learning progress (LP) expected from engaging a task. While these strategies can make equivalent predictions in certain simple situations (e.g., when learning curves are known and concave), LP-based mechanisms are superior in open-ended environments that contain unlearnable tasks. In such situations, LP-based strategies assign lower value to tasks where little progress is made and allow the learner to disengage from such tasks, while performance-based mechanisms, by assigning higher value to the lower-competence task, can push the learner to labor in vain. In the present experiment we asked whether humans possess, and use, metacognitive abilities to guide performance-based or LP-based exploration in two

contexts in which they could freely choose to learn about 4 competing tasks. Participants (n = 505, recruited via Amazon Mechanical Turk) were tested on a paradigm in which they could freely choose to engage with one of four different classification tasks. We are currently analyzing the results and working on a computational models of the underlying cognitive and motivational mechanisms.

7.1.2. Computational Models Of Tool Use and Speech Development: the Roles of Active Learning, Curiosity and Self-Organization

Participants: Pierre-Yves Oudeyer [correspondant], Sébastien Forestier.

7.1.2.1. Modeling Speech and Tool Use Development in Infants

A scientific challenge in developmental and social robotics is to model how autonomous organisms can develop and learn open repertoires of skills in high-dimensional sensorimotor spaces, given limited resources of time and energy. This challenge is important both from the fundamental and application perspectives. First, recent work in robotic modeling of development has shown that it could make decisive contributions to improve our understanding of development in human children, within cognitive sciences [120]. Second, these models are key for enabling future robots to learn new skills through lifelong natural interaction with human users, for example in assistive robotics [154].

In recent years, two strands of work have shown significant advances in the scientific community. On the one hand, algorithmic models of active learning and imitation learning combined with adequately designed properties of robotic bodies have allowed robots to learn how to control an initially unknown high-dimensional body (for example locomotion with a soft material body [3]). On the other hand, other algorithmic models have shown how several social learning mechanisms could allow robots to acquire elements of speech and language [105], allowing them to interact with humans. Yet, these two strands of models have so far mostly remained disconnected, where models of sensorimotor learning were too "low-level" to reach capabilities for language, and models of language acquisition assumed strong language specific machinery limiting their flexibility. Preliminary work has been showing that strong connections are underlying mechanisms of hierarchical sensorimotor learning, artificial curiosity, and language acquisition [54].

Recent robotic modeling work in this direction has shown how mechanisms of active curiosity-driven learning could progressively self-organize developmental stages of increasing complexity in vocal skills sharing many properties with the vocal development of infants [39]. Interestingly, these mechanisms were shown to be exactly the same as those that can allow a robot to discover other parts of its body, and how to interact with external physical objects [149].

In such current models, the vocal agents do not associate sounds to meaning, and do not link vocal production to other forms of action. In other models of language acquisition, one assumes that vocal production is mastered, and hand code the meta-knowledge that sounds should be associated to referents or actions [105]. But understanding what kind of algorithmic mechanisms can explain the smooth transition between the learning of vocal sound production and their use as tools to affect the world is still largely an open question.

The goal of this work is to elaborate and study computational models of curiosity-driven learning that allow flexible learning of skill hierarchies, in particular for learning how to use tools and how to engage in social interaction, following those presented in [51], [3], [45], [39]. The aim is to make steps towards addressing the fundamental question of how speech communication is acquired through embodied interaction, and how it is linked to tool discovery and learning.

We take two approaches to study those questions. One approach is to develop robotic models of infant development by looking at the developmental psychology literature about tool use and speech and trying to implement and test the psychologists' hypotheses about the learning mechanisms underlying infant development. Our second approach is to directly collaborate with developmental psychologists to analyze together the data of their experiments and develop other experimental setup that are well suited to answering modeling questions about the underlying exploration and learning mechanisms. We thus started to collaborate with Lauriane Rat-Fischer, a developmental psychologist working in Toulouse on the emergence of tool use in the first years of human life. We are currently analyzing together the behaviour of 22 month old infants in a tool

use task where the infants have to retrieve a toy put in the middle of a tube by inserting sticks into the tube and pushing the toy out. We are looking at the different actions of the infant with tools and toys but also its looking behaviour, towards the tool, toys or the experimenter, and we are trying to infer the goals and exploration strategies of the infant.

In our recent robotic modeling work, we showed that the Model Babbling learning architecture allows the development of tool use in a robotic setup, through several fundamental ideas. First, goal babbling is a powerful form of exploration to produce a diversity of effects by self-generating goals in a task space. Second, the possible movements of each object define a task space in which to choose goals, and the different task spaces form an object-based representation that facilitates prediction and generalization. Also, cross-learning between tasks updates all skills while exploring one in particular. A novel insight was that early development of tool use could happen without a combinatorial action planning mechanism: modular goal babbling in itself allowed the emergence of nested tool use behaviors.

This year we extended this architecture so that the agent can imitate caregiver's sounds in addition to exploring autonomously [78]. We hypothesized that these same algorithmic ingredients could allow a joint unified development of speech and tool use. Our learning agent is situated in a simulated environment where a vocal tract and a robotic arm are to be explored with the help of a caregiver. The environment is composed of three toys, one stick that can be used as a tool to move toys, and a caregiver moving around. The caregiver helps in two ways. If the agent touches a toy, the caregiver produces this toy's name, but otherwise produces a distractor word as if it was talking to another adult. If the agent produces a sound close to a toy's name, the caregiver moves this toy within agent reach 8.



Figure 8. Agent's robotic and vocal environment. Left: Agent's 3 DOF arm, controlled with 21 parameters, grabs toys with its hand, or uses the stick to reach toys. Caregiver brings a toy within reach if the agent says its name. Right: Agent's vocal environment representing sounds as trajectories in the two first formants space. Agent's simulated vocal tract produces sounds given 28 parameters. When agent touches a toy, caregiver says toy's name. Some sounds corresponding to random parameters are plotted in red, and some sounds produced when imitating caregiver's /uye/ word in blue (best imitation in bold, error 0.3).

We show that our learning architecture based on Model Babbling allows agents to learn how to 1) use the robotic arm to grab a toy or a stick, 2) use the stick as a tool to get a toy, 3) learn to produce toy names with the vocal tract, 4) use these vocal skills to get the caregiver to bring a specific toy within reach, and 5) choose the most relevant of those strategies to retrieve a toy that can be out-of-reach. Also, the grounded exploration of toys accelerates the learning of the production of accurate sounds for toy names once the caregiver is able to recognize them and react by bringing them within reach, with respect to distractor sounds without any

meaning in the environment. Our model is the first to allow the study of the early development of tool use and speech in a unified framework.

This model focuses on the role of one important form of body babbling where exploration is directed towards self-generated goals in free play, combined with imitation learning of a contingent caregiver. This model does not assume capabilities for complex sequencing and combinatorial planning which are often considered necessary for tool use. Yet, we show that the mechanisms in this model allow a learner to progressively discover how to grab objects with the hand, how to use objects as tools to reach further objects, how to produce vocal sounds, and how to leverage these vocal sounds to use a caregiver as a social tool to retrieve objects. Also, the discovery that certain sounds can be used as a social tool further guides vocal learning. This model predicts that infants learn to vocalize the name of toys in a natural play scenario faster than learning other words because they often choose goals related to those toys and engage caregiver's help by trying to vocalize those toys' names. We presented those results at the 39th Annual Conference of the Cognitive Science Society (CogSci 2017).

7.1.3. Computational Models Of Developmental Exploration Mechanisms in Vocal Babbling and Arm Reaching in Infants

Participants: Pierre-Yves Oudeyer [correspondant], Clement Moulin-Frier, Freek Stulp, Jules Brochard. 7.1.3.1. Proximodistal Exploration in Motor Learning as an Emergent Property of Optimization

To harness the complexity of their high-dimensional bodies during sensorimotor development, infants are guided by patterns of freezing and freeing of degrees of freedom. For instance, when learning to reach, infants free the degrees of freedom in their arm proximodis-tally, i.e. from joints that are closer to the body to those that are more distant. We formulated and studied computationally the hypothesis that such patterns can emerge spontaneously as the result of a family of stochastic optimization processes (evolution strategies with covariance-matrix adaptation), without an innate encoding of a maturational schedule. In particular, we made simulated experiments with an arm where a computational learner progressively acquires reaching skills through adaptive exploration, and we showed that a proximodistal organization appears spontaneously, which we denoted PDFF (ProximoDistal Freezing and Freeing of degrees of freedom). We also compared this emergent organization between different arm morphologies – from human-like to quite unnatural ones – to study the effect of different kinematic structures on the emergence of PDFF. This work was published in the journal *Developmental Science*[74].

7.1.3.2. Emergent Jaw Predominance in Vocal Development through Stochastic Optimization

Infant vocal babbling is strongly relying on jaw oscillations, especially at the stage of canonical babbling, which underlies the syllabic structure of world languages. We have proposed, modelled and analyzed an hypothesis to explain this predominance of the jaw in early babbling. This hypothesis states that general stochastic optimization principles, when applied to learning sensorimotor control, automatically generate ordered babbling stages with a predominant exploration of jaw movements in early stages, just like they generate proximo-distal organization of exploration in arm reaching as described in the paragraph above. In particular, such stochastic optimization principles predominantly explore jaw movement at the beginning of vocal learning, and when close to the rest position of the vocal tract, as it impacts the auditory effects more than other articulators. This work was published in the journal *IEEE Transactions on Cognitive and Developmental Systems*[73].

7.1.4. Models of Self-organization of lexical conventions: the role of Active Learning and Active Teaching in Naming Games

Participants: William Schueller [correspondant], Pierre-Yves Oudeyer.

How does language emerge, evolve and gets transmitted between individuals? What mechanisms underly the formation and evolution of linguistic conventions, and what are their dynamics? Computational linguistic studies have shown that local interactions within groups of individuals (e.g. humans or robots) can lead to self-organization of lexica associating semantic categories to words [165]. However, it still doesn't scale well to complex meaning spaces and a large number of possible word-meaning associations (or lexical conventions), suggesting high competition among those conventions.

In statistical machine learning and in developmental sciences, it has been argued that an active control of the complexity of learning situations can have a significant impact on the global dynamics of the learning process [120] [130] [139]. This approach has been mostly studied for single robotic agents learning sensorimotor affordances [150][40]. However active learning might represent an evolutionary advantage for language formation at the population level as well [54] [167].

Naming Games are a computational framework, elaborated to simulate the self-organization of lexical conventions in the form of a multi-agent model [166]. Through repeated local interactions between random couples of agents (designated *speaker* and *hearer*), shared conventions emerge. Interactions consist of uttering a word – or an abstract signal – referring to a topic, and evaluating the success or failure of communication.

However, in existing works processes involved in these interactions are typically random choices, especially the choice of a communication topic.

The introduction of active learning algorithms in these models produces significant improvement of the convergence process towards a shared vocabulary, with the speaker [49], [60] [109] or the hearer [61] actively controlling vocabulary growth.





7.1.4.1. Definition of a local measure of convergence for the Naming Game: Local Approximated Probability of Success (LAPS), using limited memory of past interactions

In the Naming Game, one measure is usually used to represent the state of convergence of the population: the success rate, or probability of success at a given time step. It increases over time, from 0 to 1. This measure is however global, and not accessible to individual agents; in which case it would have been a perfect candidate for a functional whose maximization would drive local behavior. Several other measures have been suggested, as one based on local information gain, or entropy reduction [60]. Those measures however are either defined in a very constrained case (without synonymy and homonymy, and fixed and known numbers of words and meanings), and their minimization can actually block the process of convergence – as their evolution is not easily predictable.

Instead, we defined a local approximation of the success rate. For this, we need a representation of the state of the population. This is done by constructing an *average vocabulary* representing the population, using a partial memory of the past interactions. This vocabulary is then used together with the agent's own vocabulary to compute a probability of success. A key element of this measure is the time scale associated to the memory: in fact, it allows not only to define a degree of certainty of a given association, but also a degree of uncertainty at a higher level (word or meaning). This measure is local (available to an agent through only its own knowledge) but its convergence to 100% is bound to global dynamics. In other words, we can use it as a functional to maximize at the local level to reinforce agreement at the population level.

7.1.4.2. Active Topic Choice: LAPS and Multi-Armed Bandits

Usually, the topic used in an interaction of the Naming Game is picked randomly. A first way of introducing active control of complexity growth is through the mechanism of topic choice: choosing it according to past memory. It allows each agent to balance reinforcement of known associations and invention of new ones, which can be seen as an exploitation vs. exploration problem. This can speed up convergence processes, and even lower significantly local and global complexity: for example in [60], [61], where heuristics based on the number of past successful interactions were used.

However, we can now define new strategies directly maximizing the LAPS measure. At each step, the agent picking a topic will choose one that yields maximum expected increase of the LAPS measure. However, this expected value being computationally really costly, we use a Multi-Armed Bandit algorithm. At the beginning, only one *machine* is available, the exploration machine. When used by the agent, its parameters are updated through Thompson Sampling algorithm, and a new machine is created with the exact same parameters, corresponding to the newly explored meaning. At any time, the number of machines available to the agent is then equal to the number of already known meanings, plus one (the exploration machine).

This strategy can speed up convergence the convergence process, but also diminishes significantly the global complexity - i.e. the maximum number of distinct word-meaning association present in the population. See figure 10.



Figure 10. Measures of convergence (global probability of success) and global complexity (number of distinct word-meaning association present in the population) for simulations using Random Topic Choice and MAB LAPS maximization Topic Choice. The active topic choice strategy yields faster convergence, with less complexity. Parameters used: 60 agents, 40 meanings, 40 words, time scale for LAPS 10 interactions.

7.1.4.3. Acceptance policy: Updating or not vocabulary based on memory of past interactions

Another way to control complexity in the Naming Game is to choose whether to trust or not other agents during a given interaction, by taking into account or not their own word-meaning associations. In previous work, a purely stochastic acceptance of new information has been studied [97]. However, accepting or not new information should depend on the memory provided by past interactions to be efficient. To do so, we use a local approximation of the global agreement as a functional to optimize at each interaction, based on recent information: the LAPS measure 7.1.4.1. We can show that for an appropriate time scale of this recent information, local complexity (amount of word-meaning association to be remembered) remains low, without impacting the duration of the global agreement process. The exact dependance on parameters (time scale, population size, meaning and word spaces size) is still to be explored.

7.1.4.4. Structured meaning spaces exploration

In the models we have considered so far, meanings were always in a finite number, and without any structure or relative importance. Also, the whole meaning space is accessible from the start. We studied a scenario where meanings are not all available from the beginning, but taken from a growing space: known meanings plus the *Adjacent Possible* [131] [172]. In practice, we consider a graph of meanings, and a starting meaning m_0 . The adjacent possible is the set of nodes connected to m_0 . Whenever a meaning from this set is explored, it is withdrawn from the adjacent possible, but all its neighbors not already known are added to it. In this case, Active Topic Choice helps to keep a quasi-linear pace of exploration, while agreeing on explored meanings. Random Topic Choice explores all available meanings before starting the agreement process: hence, on a big meaning space, possibly infinite, this is really inefficient in terms of communication success. See figure 11.



Figure 11. Comparison of Random and Active Topic Choice on a structured meaning space. The space used is a balanced tree, with initially accessible meaning being the root of the tree. On the left, evolution of global complexity (number of distinct word-meaning association present in the population): Active Topic Choice helps keeping a low complexity, with quasi linear growth, whereas Random Topic Choice first goes to a maximum way

higher than the final expected value. Parameters used: 10 agents, 100 meanings, 100 words. On the right, illustration of the status of a population in both cases, after half of the interactions needed to converge to global agreement. Nodes represent meanings, their size the number of agents having at least a word for them, and their color the level of agreement between all agents of the population for the given meaning. We can see that Active Topic Choice population has not talked about all meanings, but agrees on all the one that were used; whereas in the other case all meanings were used but almost no agreement is reached.

7.1.4.5. Interactive application for collaborative creation of a language: Experimenting how humans actively negotiate new linguistic conventions

How do humans agree and negotiate linguistic conventions? This question is at the root of the domain of experimental semiotics [118], which is the context of our experiment/application. Typically, the experiments of this field consist in making human subjects play a game where they have to learn how to interact/collaborate through a new unknown communication medium (such as abstract symbols). In recent years, such experiments allowed to see how new conventions could be formed and evolve in population of individuals, shading light on the origins and evolution of languages [133] [116].

We consider a version of the Naming Game [177] [140], focusing on the influence of active learning/teaching mechanisms on the global dynamics. In particular, agreement is reached sooner when agents actively choose the topic of each interaction [49], [60], [61].

Through this experiment, we confront existing topic choice algorithms to actual human behavior. Participants interact through the mediation of a controlled communication system – a web application – by choosing words to refer to objects. Similar experiments have been conducted in previous work to study the agreement dynamics on a name for a single picture [106]. Here, we make several pictures or interaction topics available, and quantify the extent to which participants actively choose topics in their interactions.

- Individual short experiment (implemented): each user interacts for about 3-4 min (<30 interactions) with a brand new population of 7 simulated agents. They take the role of one designated agent, and play the Naming Game as this agent. Each time they interact as speakers, they can select the topics of conversation from a set of 5 objects, and are offered 6 possible words to refer to them. Their choices influence the global emergence of a common lexical convention, reached when communications are successful. The goal is to maximize a score based on the number of successful interactions (among the 50 in total for each run). They can see a list of the past interactions, with chosen topic, chosen word, and whether the interaction was succesful or not. This experiment allows us to directly measure if there is a bias in the choice of topics, compared to random choice, based on memory of past interactions. Performance can then be compared to existing topic choice algorithms [49], [60], [61].
- Collective creation of a language and conceptual exploration (under development): Users interact with agents picked from a population which is kept for the whole duration of the experiment, common to all users. Meanings that can be used as topics are drawn from a bigger space than for the first experiment. Word space is a combination of a few basic available syllables (to avoid direct usage of known words). Users interact with a slowly increasing subset of this population, so that newcomers have the same level of influence within their own part of the experiment as people who interacted at the beginning of the day. Successfully communicating about certain meanings/objects unlocks new available meanings, and therefore we can observe the whole process of collective conceptual exploration. Linguistic conventions are set and learned/shared by users, through the interaction with simulated agents. Users never interact directly with each other, therefore no synchronization is needed. In other words, if one user decides not to finish the current interaction, it will not affect other users. We can measure in this scenario statistical properties of the language like frequency distribution, rate of exploration as well as degree of convergence.

The experiment – available at http://naming-game.bordeaux.inria.fr – was presented at the Kreyon Conference in Rome, in september 2017, during a talk and as part of interactive installation consisting in numerous scientific experiments. Insufficient data was collected to get significant results. To recruit more players and collect a large amount of data, we plan to use crowdsourcing platforms.

7.2. Autonomous Machine Learning and Applications to Developmental Robotics

7.2.1. Intrinsically Motivated Goal Exploration and Multi-Task Reinforcement Learning

Participants: Sébastien Forestier, Pierre-Yves Oudeyer [correspondant], Alexandre Péré, Olivier Sigaud, Pierre Manceron, Yoan Mollard.



Figure 12. Example of a game with the interface already existing. Play it here: http://naming-game.bordeaux.inria.fr

7.2.1.1. Intrinsically Motivated Exploration of Spaces of Parameterized Skills/Tasks and Application to Robot Tool Learning

A major challenge in robotics is to learn parametrized policies to solve multi-task reinforcement learning problems in high-dimensional continuous action and effect spaces. Of particular interest is the acquisition of inverse models which map a space of sensorimotor problems to a space of motor programs that solve them. For example, this could be a robot learning which movements of the arm and hand can push or throw an object in each of several target locations, or which arm movements allow to produce which displacements of several objects potentially interacting with each other, e.g. in the case of tool use. Specifically, acquiring such repertoires of skills through incremental exploration of the environment has been argued to be a key target for life-long developmental learning [96].

This year we developed a formal framework called "Unsupervised Multi-Goal Reinforcement Learning", as well as a formalization of intrinsically motivated goal exploration processes (IMGEPs), that is both more compact and more general than our previous models [89]. We experimented several implementations of these processes in a complex robotic setup with multiple objects 13, associated to multiple spaces of parameterized reinforcement learning problems, and where the robot can learn how to use certain objects as tools to manipulate other objects. We analyzed how curriculum learning is automated in this unsupervised multi-goal exploration process, and compared the trajectory of exploration and learning of these spaces of problems with the one generated by other mechanisms such as hand-designed learning curriculum, or exploration targeting a single space of problems, and random motor exploration. We showed that learning several spaces of diverse problems can be more efficient for learning complex skills than only trying to directly learn these complex skills. We illustrated the computational efficiency of IMGEPs as these robotic experiments use a simple memory-based low-level policy representations and search algorithm, enabling the whole system to learn online and incrementally on a Raspberry Pi 3.





Figure 13. Robotic setup. Left: a Poppy Torso robot (the learning agent) is mounted in front of two joysticks. Right: full setup: a Poppy Ergo robot (seen as a robotic toy) is controlled by the right joystick and can hit a tennis ball in the arena which changes some lights and sounds.

In order to run more scientific experiments in a shorter time, we scaled up this experimental setup to a platform of 6 identical Poppy Torso robots, each of them having the same environment to interact with. Every robot can run a different task with a specific algorithm and parameters each 14. In this setup Poppy Torso robots are requesting jobs to a dedicated computer acting as a job manager which monitors execution and distributes jobs to available robots. Moreover, each Poppy Torso can also perceives the motion of a second Poppy Ergo robot, than can be used, this time, as a distractor performing random motions to complicate the learning problem. 12 top cameras and 6 head cameras can dump video streams during experiments, in order to record video datasets. Data and videos are stored on-the-fly on 6 hard disks.



Figure 14. Platform of 6 robots with identical environment: joysticks, Poppy Ergo, ball in an arena, and a distractor. The central bar supports the 12 top cameras.

7.2.1.2. Unsupervised Deep Learning of Goal Spaces for Goal Intrinsically Motivated Goal Exploration

Intrinsically motivated goal exploration algorithms enable machines to discover repertoires of policies that produce a diversity of effects in complex environ- ments. These exploration algorithms have been shown to allow real world robots to acquire skills such as tool use in high-dimensional continuous state and action spaces. However, they have so far assumed that self-generated goals are sampled in a specifically engineered feature space, limiting their autonomy. We have proposed an approach using deep representation learning algorithms to learn an adequate goal space. This is a developmental 2-stage approach: first, in a perceptual learning stage, deep learning algorithms use passive raw sensor observations of world changes to learn

a corresponding latent space; then goal exploration happens in a second stage by sampling goals in this latent space. We made experiments with a simulated robot arm interacting with an object, and we show that exploration algorithms using such learned representations can closely match, and even sometimes improve, the performance obtained using engineered representations.

7.2.1.3. Combining deep reinforcement learning and curiosity-driven exploration

A major challenge of autonomous robot learning is to design efficient algorithms to learn sensorimotor skills in complex and high-dimensional continuous spaces. Deep reinforcement learning (RL) algorithms are natural candidates in this context, because they can be adapted to the problem of learning continuous control policies with low sample complexity. However, these algorithms, such as DDPG (Lillicrap et al., 2016) suffer from exploration issues in the context of sparse or deceptive reward signals.

In this project, we investigate how to integrate deep reinforcement learning algorithms with curiosity-driven exploration methods. A key idea consists in decorrelating the exploration stage from the policy learning stage by using a memory structure used in deep RL called a replay buffer. Curiosity-driven exploration algorithms, also called Goal Exploration Processes (GEPs) are used in a first stage to efficiently explore the state and action space of the problem, and the corresponding data is stored into a replay buffer. Then a DDPG learns a control policy from the content of this replay buffer.

The internship of Pierre Manceron has been dedicated to trying this methodology in practice. Pierre has combined GEPs obtained from the Explauto open-source library (Moulin-Frier et al., 2014) and his own implementation of DDPG, and benchmarked the combination using the openAI Gym toolkit (Duan et al., 2016).

Preliminary results have revealed some stability issues in DDPG, whereas encouraging results where obtained about the combination with GEPs. Beyond getting more robust results and publishing them, our next goal is to envision other ways to integrate deep RL with curiosity-driven exploration processes by using the tools of the former to more efficiently implement the latter.

7.2.2. Social Learning of Interactive Skills

Participants: Manuel Lopes [correspondant], Baptiste Busch, Yoan Mollard, Thibaut Munzer.

This work was made in collaboration with Marc Toussaint and Guilherme Maeda.

7.2.2.1. Preference learning on the execution of collaborative human-robot tasks

One important aspect of the human-robot collaboration is to be able to learn the user's preferences on the sequence of actions. By querying the user on the next action, when the uncertainty is high, the robot learns the user preferences (Q-function) to solve the task. From a planning point of view, this Q-function can then be integrated into the solver to select the user preferred route to solve a task when multiple choices are available. Therefore, this work aims at reducing the human cognitive load by:

- querying demonstrations only when the uncertainty is above a certain threshold,
- always choose the user preferred actions.

This work has been accepted for publication in the *International Conference on Robotics and Automation* (ICRA) 2017 and presented during the conference [80].

Interestingly, this also raises questions on the robot autonomy and its perception by the human coworker. By interacting with a user, the robot starts to learn the preferred actions and will takes initiative to perform them on the next assembly. The question is, how does the user perceives this initiative taking? To answer this question, we have conducted a user study to analyze the impact of robot initiative on the collaboration. Two conditions were considered:

- a semi-autonomous robot that learns and decides when to execute a supporting action,
- a support robot that has to be instructed of each action on a collaborative task.

We found that users prefer the semi-autonomous robot and that the behavior was closer to their expectations despite them being more afraid of it. We also found that even if users noticed the robot was learning in one case, they wanted more autonomy in both conditions. This research was published in the companion of the *Conference on Human-Robot Interaction* (HRI) 2017 and presented during the poster sessions of the conference [82].

7.2.2.2. Learning legible motions from interaction

In a human-robot collaboration context, understanding and anticipating the robot intentions ease the completion of a joint-task. Whereas previous work has sought to explicitly optimize the legibility of behavior, we investigate legibility as a property that arises automatically from general requirements on the efficiency and robustness of joint human-robot task completion.

Following our previous work on legibility of robot motions [64], we have conducted several user experiments to analyze the effects of the policy representation on the universality of the legibility.

This work lead to a submission of a journal article to the International Journal of Social Robotics (IJSR) [69] 7.2.2.3. *Postural optimization for an ergonomic human-robot interaction*

7.2.2.3. Postural optimization for an ergonomic numan-robot interaction

When we, humans, accomplish a task our body posture is (partially) constrained. For example, acting on an object constrains the pose of the hand relatively to the object, and the head faces the object we are acting upon. But due to the large number of degrees of freedom (DOF) of the human body, other body parts are unconstrained and several body postures are viable with respect to the task. However, not all of them are viable in terms of ergonomics. Using a personalized human model, observational postural assessment techniques can be automatized. Optimizing the model body posture is then the logical next step to find an ergonomically correct posture for the worker to accomplish a specific task.

To optimize the subject's model to achieve a specific task, we define an objective function that minimizes the efforts of the whole body posture, based on the Rapid Entire Body Assessment (REBA) technique [124]. The objective function also account for visibility of the target object and worker's laterality. We have also implemented an automatic assessment of the worker's body posture based on the REBA method.



Figure 15. Representation of the setup considered in the user study. The robot presents to the user a spherical ball in which multiple shapes can be inserted. Final pose of the object is calculated from the user posture at his current location. Body motions during the insertion are recorded using a suit made from OptiTrack markers.

Using a spherical object, carried by a Baxter humanoid robot as illustrated in Fig. 15, we mimic an industrial scenario where the robot helps the worker by positioning and orienting an object in which the worker has to insert specific shapes. In a user-study with forty participants, we compare three different robot's behaviors,

one of them being the result of the postural optimization of the subject's personalized model. By the mean of a survey session, and the online assessment of the subject's posture during the interaction, we prove that our method leads to a safer posture, and is perceived as more comfortable.

This work has been published to the International Conference on Intelligent Robots and Systems (IROS) [75] and was presented during the conference.

7.2.2.4. Planning ergonomic sequences of actions in human-robot interaction

Following our work on physical ergonomics [75], we have extended our method to include it in the Logic-Geometric Program (LGP) [171]. This method allows us to solve Task and Motion Planning (TAMP) problems simultaneously while optimizing for maximum ergonomics on the human side.

In a simulated experimented we prove that the solver is able to choose the logic actions (e.g. "the robot places the screwdriver on the table") that provides maximum ergonomics throughout the interaction. By the mean of an experiment on our Baxter robot, we also prove that optimizing ergonomics over the full sequences of actions, as opposed to the step-wise approach we were considering in [75] where ergonomics was optimized for single atomic actions, lead to a more ergonomic interaction.

As both human and robot agents are capable of performing the same task, this creates a need to communicate the planned sequence of actions efficiently to the human. Two problems are raised by this dynamic task allocation. First, the human need to understand the current action performed by the robot to anticipate and react if necessary. Second, the human must know beforehand or during the executions the actions that he or she has to perform and when to perform them.

time 1540

The robot, using its left arm, hands over the toolbox handle to the agent's right hand.

Figure 16. A representation of an action generated by the solver and displayed on a webpage to simplify the visualization. The agent can click on the image to start a video of the action. Arrows on the right corner are also clickable to navigate between the previous and the next actions.

To this purpose, we also introduce a graphical interface that displays the current action and the geometric of the scene as illustrated in figure 16. This graphical interface can be used offline to train the human on the steps required for the whole task. It can also be displayed online to show the current action to reduce the human cognitive load of understanding the action performed by the robot and understanding the action that he or her is expected to perform.

This work was submitted to the International Conference on Robotics and Automation (ICRA) and is currently under review.

7.2.2.5. Active incremental learning of robot movement primitives

A robot coworker acting as a third hand brings the challenge that its skills must be augmented and tailored as needed, over time. To this end, imitation learning can rely on the presence of the human coworker as a teacher. However, imitation learning has primarily addressed *how* to endow and refine robots with motor skills but not *when* the learning should take place. Reasoning when improvement is actually needed is, nevertheless, an essential and difficult problem to be solved. We propose an active learning algorithm that allows a robot to reason about the confidence of its movement primitives. It allows the robot to decide when a demonstration is required, making active requests to the human coworker depending on its confidence.

This capability also sheds light onto the problem of deciding how many demonstrations are needed to construct a probabilistic model when learning from demonstrations. Under active learning, the number of demonstrations is indicated by the robot, on-demand. The method can be used on a single demonstration, in a *one-shot learning* fashion. If the extrapolation is, however, beyond the scope of the existing demonstration—indicated by the uncertainty—an active request will be made. The proposed method also offers a principled way to train Dynamical Movement Primitives (DMPs) with contextualized demonstrations encoded by Gaussian Processes (GPs)— details can be found in [79].

Algorithm 1 GP-DMP($\{\mathbf{x}^d, t^d\}_{1:D}, \mathbf{x}_*, u_{trig}$)								
1:	$GP.train(\{\mathbf{x}^d, t^d\}_{1:D})$							
2:	$(\mu_{1:N}, \sigma_{1:N}^2) \leftarrow GP.predict(\mathbf{x}_*)$							
3:	if TRAJ_UNC($\sigma_{1:N}^2$) < u_{trig} then							
4:	$DMP.train(\mu_{1:N})$							
5:	$y \leftarrow DMP.reach_goal(\mathbf{x}_*)$							
6:	ROBOT_EXECUTION (y)							
7:	else							
8:	$\{\mathbf{x}^{D+1}, t^{D+1}\} \leftarrow \text{ACTIVE}_{\text{REQUEST}}$							
9:	go to line 1							

Figure 17. Pseudocode of the algorithm to add a demonstration only if required

The method is based on a combination of GPs and DMPs, where the former provides the confidence bounds in which the demonstrations are being extrapolated, and the latter accounts for prediction errors due to the nonlinearities of the function being approximated. Algorithm 17 shows a pseudocode of the proposed method.

Figure 18 shows the decrease of the uncertainty on the prediction of 10 trajectories as the number of demonstrations increase (the ellipses represent the projections of one standard deviation along the trajectory). Figure 18 (a)(b) and (c) show the predictions after one, two and five demonstrations, respectively. The blue color in (c) indicates that the robot has confidence it can execute the 10 predicted trajectories.

7.3. Representation Learning



Figure 18. Decrease of the trajectory uncertainties as demonstrations for objects are being requested. The ellipses show a single standard deviation. Uncertainties above the threshold are shown in red. The blue trajectories indicate that the robot is confident in executing the primitives. (a) The uncertainties for all contexts after one demonstration. (b) Predictions after two demonstrations. (c) Predictions after five demonstrations.

7.3.1. Cross-situational noun and adjective learning in an interactive scenario

Participants: David Filliat [correspondant], Yuxin Chen.

Future intelligent robots are expected to be able to adapt continuously to their environment. For this purpose, recognizing new objects and learning new words through interactive learning with humans is fundamental. Such setup results in ambiguous teaching data which humans have been shown to address using cross-situational learning, i.e. by analyzing common factors between multiple learning situations. Moreover, they have been shown to be more efficient when actively choosing the learning samples, e.g. which object they want to learn. Implementing such abilities on robots can be performed by latent-topic learning models such as Non-Negative Matrix Factorization or Latent Dirichlet Allocation. These cross-situational learning methods tackle referential and linguistic ambiguities, and can be associated with active learning strategies. We propose two such methods: the Maximum Reconstruction Error based Selection (MRES) and Confidence Base Exploration (CBE). We present extensive experiments using these two learning algorithms through a systematic analysis on the effects of these active learning strategies in contrast with random choice. In addition, we study the factors underlying the active learning by focusing on the use of sample repetition, one of the learning behaviors that have been shown to be important for humans. These results have been published in a journal paper [70]

7.3.2. State Representation Learning in the Context of Robotics

Participants: David Filliat [correspondant], Natalia Diaz Rodriguez, Timothee Lesort, Mathieu Seurin.

Our understanding of the world depends highly on our capacity to produce intuitive and simplified representations which can be manipulated and combined easily to solve problems. We worked on reproducing this simplification process using a neural network to build a low dimensional state representation of the world from images acquired by a robot. As in the approach from Jonschkowski [129], we learn in an unsupervised way using prior knowledge about the world as loss functions called robotic priors.

The robotic priors loss function impose constraint in a low dimension space. We call this space the representation space and it contains the underlying parameters of the robot environment. This constraint are physic related as the time coherence of the representation, the repeatability, the proportionality and causality of the actions inside the representation space. Imposing those constraint to sequences of images makes it possible to learn a mapping from image to our representation state space. We extend the previous approach to high dimension richer images to learn a 3D representation of the hand position of a robot from RGB images.

We propose a quantitative evaluation of the learned representation using nearest neighbors in the state space that allows to assess its quality and show both the potential and limitations of robotic priors in realistic environments. We augment image size, add distractors and domain randomization, all crucial components to achieve transfer learning to real robots.

Finally, we also contribute a new prior to improve the robustness of the representation. This prior takes profit of the initial state of the robot to bring together representation of different sequences. The applications of such low dimensional state representation range from easing reinforcement learning (RL) and knowledge transfer across tasks, to facilitating learning from raw data with more efficient and compact high level representations.







Our experiments [90] (see figure 19 for an illustration) compare results in different setup with state representation in 2 and 3D, with different amount of distractors. The results show that the robotic prior approach is able to extract high level representation such as the 3D position of an arm and organize it into a compact and coherent space of states in a challenging dataset.

7.3.3. Transfer Learning from Simulated to Real World Robotic Setups

Participants: Florian Golemo [correspondant], Pierre-Yves Oudeyer.

This work was made in collaboration with Adrien Ali Taiga and Aaron Courville. Reinforcement learning with function approximation has demonstrated remarkable performance in recent years. Prominent examples include playing Atari games from raw pixels, learning complex policies for continuous control, or surpassing human performance on the game of Go. However most of these successes were achieved in non-physical environments (simulations, video games, etc.). Learning complex policies on physical systems remains an

open challenge. Typical reinforcement learning methods require a lot of data which makes it unsuitable to learn a policy on a physical system like a robot, especially for dynamic tasks like throwing or catching a ball. One approach to this problem is to use simulation to learn control policies before applying them in the real world. This raises new problems as the discrepancies between simulation and real world environments ("reality gap") prevent policies trained in simulation from performing well when transfered to the real world. This is an instance of *domain adaption* where the input distribution of a model changes between training (in simulation) and testing (in real environment). The focus of this work is in settings where resetting the environment frequently in order to learn a policy directly in the real environment is highly impractical. In these settings the policy has to be learned entirely in simulation but is evaluated in the real environment, as *zero-shot transfer*.

In simulation there are differences in physical properties (like torques, link weights, noise, or friction) and in control of the agent, specifically joint control in robots. We propose to compensate for both of these source of issues with a generative model to bridge the gap between the source and target domain. By using data collected in the target domain through task-independent exploration we train our model to map state transitions from the source domain to state transition in the target domain. This allows us to improve the quality of our simulated robot by grounding its trajectories in realistic ones. With this learned transformation of simulated trajectories we are able to run an arbitrary RL algorithm on this augmented simulator and transfer the learned policy directly to the target task. We evaluate our approach in several OpenAI gym environments that were modified to allow for drastic torque and link length differences.

7.3.4. Measuring Uncertainty in Deep Learning Networks

Participants: Florian Golemo [correspondant], Manuel Lopes.

As precursor to the main objective of the IGLU project, we investigated methods that would enable deep neural networks to judge their knowledge about a domain.

Neural networks, especially deep ones, have been shown to be able to model arbitrarily complex problems, and thus offer powerful tools for machine learning. Yet they come with a significant flaw of not being inherently able to represent certainty of their predictions. By adding a measure of uncertainty to neural networks, this technology could be applied to autonomous exploration and open-ended learning tasks.

Thus the goal of this project was to find a method to measure how much knowledge a neural network has about about an unlabeled data item (measure of uncertainty), and to apply this new measure in an active learning context. The objective of the latter was to demonstrate the efficiency in handpicking interesting data, to optimally extend the system's own capabilities.

We were successful in finding a measure of uncertainty that would reliably distinguish data that the network has seen before, from data that was generally unfamiliar to the network. This measure was created by measuring the entropy of the network's last layer across a batch of stochastic samples generated by adding Poisson noise to the inputs.

The measure failed however to outperform random sampling in several active learning scenarios. Yarin Gal published related work as part of his dissertation [117] after this project was concluded. He elaborated that deep neural networks are very effective in canceling out input noise. The author suggested to use existing "Dropout" layers instead for stochastic sampling, but he reaches the same conclusion of using the last layer entropy as measure of uncertainty.

7.4. Applications in Robotic myoelectric prostheses

Participants: Pierre-Yves Oudeyer [correspondant], Manuel Lopes, Mathilde Couraud, Sebastien Mick, Aymar de Rugy, Daniel Cattaert, Florent Paclet.

Together with the Hybrid team at INCIA, CNRS, the Flowers team continued to work on establishing the foundations of a long-term project related to the design and study of myoelectric robotic prosthesis. The ultimate goal of this project is to enable an amputee to produce natural movements with a robotic prosthetic arm (open-source, cheap, easily reconfigurable, and that can learn the particularities/preferences of each user). This will be achieved by 1) using the natural mapping between neural (muscle) activity and limb movements in healthy users, 2) developing a low-cost, modular robotic prosthetic arm and 3) enabling the user and the prosthesis to co-adapt to each other, using machine learning and error signals from the brain, with incremental learning algorithms inspired from the field of developmental and human-robot interaction.

7.4.1. Model and experiments to optimize co-adaptation in a simplified myoelectric control system

To compensate for a limb lost in an amputation, myoelectric prostheses use surface electromyography (EMG) from the remaining muscles to control the prosthesis. Despite considerable progress, myoelectric controls remain markedly different from the way we normally control movements, and require intense user adaptation. To overcome this, our goal is to explore concurrent machine co-adaptation techniques that are developed in the field of brain-machine interface, and that are beginning to be used in myoelectric controls. We combined a simplified myoelectric control with a perturbation for which human adaptation is well characterized and modeled, in order to explore co-adaptation settings in a principled manner. First, we reproduced results obtained in a classical visuomotor rotation paradigm in our simplified myoelectric context, where we rotate the muscle pulling vectors used to reconstruct wrist force from EMG. Then, a model of human adaptation in response to directional error was used to simulate various co-adaptation settings, where perturbations and machine co-adaptation are both applied on muscle pulling vectors. These simulations established that a relatively low gain of machine co-adaptation that minimizes final errors generates slow and incomplete adaptation, while higher gains increase adaptation rate but also errors by amplifying noise. After experimental verification on real subjects, we tested a variable gain that cumulates the advantages of both, and implemented it with directionally tuned neurons similar to those used to model human adaptation. This enables machine co-adaptation to locally improve myoelectric control, and to absorb more challenging perturbations. Significance. The simplified context used here enabled to explore co-adaptation settings in both simulations and experiments, and to raise important considerations such as the need for a variable gain encoded locally. This work was published in the *Journal Of Neural Engineering* in [71]

7.4.2. Performance and Usability of Various Robotic Arm Control Modes from Human Force Signals

Elaborating an efficient and usable mapping between input commands and output movements is still a key challenge for the design of robotic arm prostheses. In order to address this issue, we developped and compared three different control modes, by assessing them in terms of performance as well as general usability. Using an isometric force transducer as the command device, these modes convert the force input signal into either a position or a velocity vector, whose magnitude is linearly or quadratically related to force input magnitude. With the robotic arm from the open source 3D-printed Poppy Humanoid platform simulating a mobile prosthesis, an experiment was carried out with eighteen able-bodied subjects performing a 3-D target-reaching task using each of the three modes. The subjects were given questionnaires to evaluate the quality of their experience with each mode, providing an assessment of their global usability in the context of the task. According to performance metrics and questionnaire results, velocity control modes were found to perform better than position control mode in terms of accuracy and quality of control as well as user satisfaction and comfort. Subjects also seemed to favor quadratic velocity control over linear (proportional) velocity control, even if these two modes did not clearly distinguish from one another when it comes to performance and usability assessment. These results highlight the need to take into account user experience as one of the key criteria for the design of control modes intended to operate limb prostheses. This work was published in the journal Frontiers in Neurorobotics [72].

7.5. Applications in Educational Technologies

7.5.1. Machine Learning for Adaptive Personalization in Intelligent Tutoring Systems

Participants: Manuel Lopes [correspondant], Pierre-Yves Oudeyer, Didier Roy, Alexandra Delmas, Benjamin Clement.

7.5.1.1. The Kidlearn project

Kidlearn is a research project studying how machine learning can be applied to intelligent tutoring systems. It aims at developing methodologies and software which adaptively personalize sequences of learning activities to the particularities of each individual student. Our systems aim at proposing to the student the right activity at the right time, maximizing concurrently his learning progress and its motivation. In addition to contributing to the efficiency of learning and motivation, the approach is also made to reduce the time needed to design ITS systems.

We present an approach to Intelligent Tutoring Systems which adaptively personalizes sequences of learning activities to maximize skills acquired by students, taking into account the limited time and motivational resources. At a given point in time, the system proposes to the students the activity which makes them progress faster. We introduce two algorithms that rely on the empirical estimation of the learning progress, **RiARiT** that uses information about the difficulty of each exercise and **ZPDES** that uses much less knowledge about the problem.

The system is based on the combination of three approaches. First, it leverages recent models of intrinsically motivated learning by transposing them to active teaching, relying on empirical estimation of learning progress provided by specific activities to particular students. Second, it uses state-of-the-art Multi-Arm Bandit (MAB) techniques to efficiently manage the exploration/exploitation challenge of this optimization process. Third, it leverages expert knowledge to constrain and bootstrap initial exploration of the MAB, while requiring only coarse guidance information of the expert and allowing the system to deal with didactic gaps in its knowledge. The system is evaluated in a scenario where 7-8 year old schoolchildren learn how to decompose numbers while manipulating money. Systematic experiments are presented with simulated students, followed by results of a user study across a population of 400 school children. [14]

7.5.1.2. Linear UCB for intelligent tutoring system

What we wanted to do was to use the feature space considering features of students, features of explanations and maybe feature about the exercises. We wanted use this feature space to guide the bandit algorithm to recommend explanations. The algorithms that we already developed cannot be used in this kind of problem, because the order and the value of the bandit values depends on the time and on the learning progress the actions give in time. And in this new experiment, we wanted to recommend feedback, base on the population results depending on student's and explanation's features. In this problem, there is no consideration about temporality and the long-term progression of the student is too hard to correlate with a particular explanation. So different algorithms have been studied in the literature to use contextual bandit to make recommendation over a population. The algorithm that we wanted to use and adapt to our purpose is the LinUCB algorithm from [136].

Different kind of simulation have been made to test this framework. These experiments were made by generating various population of student defined by binary features with a number of dimension between 5 and 20. Also a set of activity have been defined and depending of the features of the students, their result for each activity was right or wrong. The algorithm was trained by making the student work on random activities, this way, the algorithm would learn the correlation between the student feature space and the success/failure to activities. After that the tests was made by letting the algorithms to choose the activity to propose to the student depending of their features. A lot of different configurations have been tested. The algorithm showed good results for low dimension feature space (3 to 9 features) with 90% accuracy, but for high dimension feature space, the results dropped to 20-30% accuracy.

7.5.1.3. Experiment in class in 2018

An experiment will be held in mars 2018 about testing the kidlearn framework in classroom in Bordeaux Metropole. The goal is to test a new feature by giving the student the opportunity to have different kinds of

choice. This choice would be managed by a multi-armed bandit algorithm. We want to make an experiment with 600 student from bordeaux Metropole and we are currently discussing with the Local Education Authority to also test the application in other school in other departments.

7.5.1.4. The KidBreath project

To create learning contents linked to asthma to personalize it like mathematics activities in Kidlearn project [14] we used recommandation criterias in Therapeutic Education Program for asthma kids made by Health High Autority. Following an approach of participatory design [102], contents were validated by medical experts like health educators, pulmonoligists and pediatrics. Then, we conducted a workshop with forty kids aged 8 in order to iterate over the application interfaces and evaluate enjoy about it with observations. Finally, we realized a focus group with 5 asthma kids to validate the global comprehension of a part of the content. It revealed that children wanted more contents about the crisis treatment and how the asthma works in the human system (verbatims).

In a preliminary study, we experimented the participatory design process (PD) in the context of asthma elearning using KidBreath tool. We evaluated in two Year 4 classes its efficacy in a motivation way [175], usability [142], [157], disease knowledge [92] and interests of children by their system [119]. After two weeks a use, results showed, in acceptance with behaviors, high level of intrinsic motivation when using KidBreath, usability and enjoyment of edutainment activities. This pilot study tends to confirm to continue with this approach with asthma kids at home (study in progress). These results was presented in the 29st Conference in Computer-Human Interaction in Poitiers, France.

We presented Thesis project in some events this year, with one publication surbmitted and validated:

- 21ème Congrès de Pneumologie de la Langue Française, January 2017, Marseille, France (Poster)
- Journée IHM et IA, March 2017, Paris, France (oral presentation)
- Ma Thèse en 180 secondes, March 2017, Bordeaux, France (oral presentation),
- 29ème conférence francophone sur l'Interaction Homme-Machine, August 2017, Poitiers, France (article and oral presentation) [76],
- La nuit européenne des chercheurs, September 2017, Bordeaux, France (oral presentation).

7.5.2. Poppy Education: Designing and Evaluating Educational Robotics Kits

Participants: Pierre-Yves Oudeyer [correspondant], Didier Roy, Théo Segonds, Stéphanie Noirpoudre, Thibault Desprez, Damien Caselli, Aurélie Lopes, Kelian Schindowsky.

Poppy Education project aims to create, evaluate and disseminate all-inclusive pedagogical kits, open-source and low cost, for teaching computer science and robotics.

It is designed to help young people to take ownership with concepts and technologies of the digital world, and provide the tools they need to allow them to become actors of this world, with a considerable socio-economic potential. It is carried out in collaboration with teachers and several official french structures (French National Education, High schools, engineer schools, ...). For secondary education and higher education, scientific literacy centers, Fablabs.

Poppy Education is based on the robotic platform poppy (open-source platform for the creation, use and sharing of interactive 3D printed robots), including:

- Poppy Humanoid, a robust and complete robotics platform designed for genuine experiments in the real world and can be adapted to specific user needs.
- Poppy Torso, a variant of Poppy Humanoid that can be easily installed on any flat support.
- Ergo Jr, a robotic arm. Durable and inexpensive, it is perfect to be used in class. Python. Directly from a web browser, using Ipython notebooks (an interactive terminal, in a web interface for the Python Programming Language).
- Snap. The visual programming system Snap, which is a variant of Scratch. Its features allow a thorough introduction of information technology.

- C++, Java, Matlab, Ruby, Javascript, etc. thanks to a REST API that allows you to send commands and receive information from the robot with simple HTTP requests.
- Virtual robots (Poppy Humanoid, Torso and Ergo) can be simulated with the free simulator V-REP. It is possible in the classroom to work on the simulated model and then allow students to run their program on the physical robot.

7.5.2.1. Pedagogical experimentations : Design and experiment robots and the pedagogical activities in classroom.

This project is user centered design. The pedagogical tools of the project (robots and ressources) are being created directly with the users and evaluated in real life by experiments. So teachers and researchers co-creatie activities, test with students in class-room, exchange their uses and develop the platform as needed [81]. The activities were designed mainly with Snap! and Python. Most activities use Poppy Ergo Jr, but some use Poppy Torso (mostly in higher school because of its cost).

The pedagogical experiments in classroom carried out during the first year of the project notably allowed to create and experiment many robotic activities and to create pedagogical resources to taking in hand the robot. The main objective of the second year was to make all the activities and resources reusable (with description, documentation and illustration) easily and accessible while continuing the experiments and the diffusion of the robotic kits.



Figure 20. Experiment robots and pedagogical activities in classroom

• Pedagogical working group : in the second year, the teacher partners continued to use the robots in the classroom and to create and test new classroom activities. Four new schools (including 10 new teachers) from different backgrounds (middle-school and high school teachers) have been added to the group to add diversity. We organized some training to help them to discover and learn how to use the robotics platform. As well an engineer of the Poppy Education team went to visit the teachers in their school to see and to evaluate the pedagogical tools (robots and activities) in real contexts of use.

Five meetings have been organized during the year will all teachers partners and Poppy Education team to exchange about their projects with robots, to understand their need and to get some feedbacks from them. This experimentations are still helping us to understand better the educational needs, to create and improve the pedagogical tools.

You can see the videos of pedagogical robotics activities here:

 $https://www.youtube.com/playlist?list=PLdX8RO6QsgB7hM_7SQNLvyp2QjDAkkzLn$

7.5.2.2. Pedagogical documents and resources

• We continued to improve the documentation of the robotic platform Poppy (https://docs.poppyproject.org/en/) and the documentation has been translated into French (https://docs.poppy-project. org/fr/.

We configured a professional platform to manage the translation of the documentation (https:// crowdin.com/project/poppy-docs. This allow whoever wants to participate in the translation of the language of their choice.

• To complete the pedagogical booklet [48] that provides guided activities and small challenges to become familiar with Poppy Ergo Jr robot and the Programming language Snap! (https://drive.google.com/file/d/0B2jV8VX-IQHwTUxXZjF3OGxHVGM/view) we provided a list of Education projects. Educational projects have been written for each activity carried out and tested in class. So each projects have its own web page including resources allowing any other teacher to carry out the activity (description, pedagogical sheet, photos / videos, pupil's sheet, teacher's sheet with correction etc.).

You can see the activities here :

https://www.poppy-education.org/activites/activites-lycee The pedagogical activities are also available on the Poppy project forum where everyone is invited to comment and create new ones: https://forum.poppy-project.org/t/liste-dactivites-pedagogiques-avec-les-robots-poppy/2305



Figure 21. Open-source educational activities with Poppy robots are available on Poppy-Education.org

• A FAQ have been written with the most frequents questions to help the users: https://www.poppy-education.org/aide/

7.5.2.3. Diffusion

- A website have been created to present the project and to share all resources and activities. https://www.poppy-education.org/
- A press release was issued to announce the poppy education website release. https://www.inria.fr/actualite/mediacenter/poppy-education-la-robotique-pedagogique-s-enrichit-dun-nouveau-site?mediego_ruuid=09f6a0c0-2ab3-11e7-a75f-fd895fe51065

- To promote educational uses of the platform, we participated in events (conference, seminar etc.). We participated as well at some workshops to introduce students to robotics and programming. See the chapter "popularization" to know the whole list.
- We sent 6 newsletters https://us13.campaign-archive.com/home/?u=17b6815514db7361fc260e0ce&id=95e9e13ae2.
- We wrote blog articles to describe workshops / activities and give feedback from experiences. https://www.poppy-education.org/#

7.5.2.4. Symposium robotics

We organized a symposium robotics for the third year (http://dm1r.fr/colloque-robotique-education/) that present research results and feedback on the use of Poppy and Thymio robots in education (other robots have been discussed as well). Poppy Education team and the working group teachers helped with the organisation of the event and during the event (talk and workshops).

7.5.2.5. Evaluate the pedagogical kits

After experimenting and create tools with educational activities in class and for the class, it is now to evaluate qualitatively and quantitatively the impact of these tools. We must therefore assess, at first, if these tools offer good usability (i.e. effectiveness, efficiency, satisfaction). Then, in a second step, select items that can be influenced by the use of these tools. For example, students' representations of robotics, their motivation to perform this type of activity, or the evolution of their skills in these areas. In 2017 we conducted experiments to evaluate the usability of kits. We also collected data on students' perceptions of robotics.

• Population

Our sample is made up of 28 teachers and 146 students from the "New Aquitaine" region who completed survey (online) during the month of June 2017. Here, we study several groups of individuals: teachers and students. Among the students we are interested in those who practiced classroom activities with the Ergo Jr kit during the school year 2016 - 2017 (N = 68) (age = 16, S = 2, 44), 37 are from section "Computer Science and Digital Sciences" (BAC S option ISN), 12 of section "Computer and Digital Creation" (BAC S option ICN) and 18 of the middle School. His 68 students are then divided into different modalities according to the characteristics of the activities they have followed: they may have declared using the educational booklet provided in the kit (N = 13) or not (N = 55); have used other robotic kits (N = 16) have practiced less than 6 hours of activity with the robot (N = 30), between 6 and 25 hours (N = 22) or more than 25 hours (N = 16); having built the robot (N = 12); have used the visual programming language Snap! (N = 46), the language of Python textual programming (N = 21), both (N = 8) or none (N = 9), it should be noted that these two languages are directly accessible via the main interface of the robot.

• Evaluation of the tool

We have selected two standardized survey dealing with this issue: SUS (The Systeme Usability Scales) [103] and The AttrakDiff [135]. These two survey are complementary and allow to identify the design problems and to account for the perception of the user during the activities. The results of these are available in the article (in French) [77] publish in the conference Didapro (Lausanne Feb, 2018). Figures 22 and 23 show the averages of the 96 respondents (68 students + 28 teachers) for each of the 10 statements from the SUS and 28 pairs of antonyms to be scored on a scale of 1 to 5 and a 7-point scale, respectively.

• Evaluation of impact on learner

	-1	2 -	1	Ŷ	1	Ļ	2	2	x	\$
[1. Je pense que je vais utiliser ce kit fréquemment]	Non							Oui	-0,15	1,26
*[2. Je trouve ce kit inutilement complexe]	Oui						7	Non	1,01	1,09
[3. Je pense que ce kit est facile à utiliser]	Non				A			Oui	0,80	1,00
*[4.Je pense que j'aurai besoin de l'aide d'un technicien pour être capable d'utiliser ce kit]	Oui			4				Non	0,80	1,18
[5.J'ai trouvé que les différentes fonctions de ce kit ont été bien intégrées]	Non						>	Oui	0,80	1,00
*[6. Je pense qu'il y a trop d'incohérences dans ce kit.]	Oui							Non	0,94	0,96
[7. J'imagine que la plupart des gens seraient capables d'apprendre à utiliser ce kit très rapidement.]	Non							Oui	1,11	1,09
*[8 J'ai trouvé ce kit très gênant à utiliser.]	Oui							Non	0,51	1,12
[9. Je me sentais très en confiance en utilisant ce kit.]	Non					1		Oui	1,17	0,95
*[10. J'ai besoin d'apprendre beaucoup de choses avant de pouvoir utiliser ce kit]	Oui			Ν	N RAN			Non	0,56	1,11

Figure 22. Result of SUS survey



Figure 23. Result of AttrakDiff survey

One of the objectives of the integration of digital sciences in school is to allow students to have a better understanding of the technological tools that surround them daily (i.e. web, data, algorithm, connected object, etc.). So, we wanted to measure how the practice of activities with ErgoJr robot had changed this apprehension; especially towards robots. For that, we used a standardized survey: "attitude towards robot" *EuroBarometer 382* originally distributed in 2012 to more than 1000 people in each country of the European Union. On the one hand, we sought to establish whether there had been a change in response between 2012 and 2017, and secondly whether there was an impact on the responses of 2017 according to the participation, or not, in educational activities with ErgoJr robot. The analysis of the results is in progress and will be published in 2018.

7.5.2.6. Partnership on education projects

Ensam

The Arts and Métiers campus at Bordeaux-Talence in partnership with Inria wishes to contribute to its educational and scientific expertise to the development of new teaching methods and tools. The objective is to develop teaching sequences based on a project approach, relying on an attractive multidisciplinary technological system: the humanoid Inria Poppy robot.

The humanoid Inria Poppy robot offers an open platform capable of providing an unifying thread for the different subjects covered during the 3-years of the Bachelor training: mechanics, manufacturing (3D printing), electrical, mecha-tronics, computer sciences, design.

Last year students of "bachelor degree" (ENSAM-Talence) have designed, manufactured, assembled and programmed 4 different solutions to replace the fixed hand of Poppy by a gripper device:

https://www.youtube.com/watch?v=DZjGaJk2fQk. For the second year, students of "bachelor degree" have designed Wheels for Poppy Torso.

https://www.poppy-education.org/2017/06/19/des-roues-pour-poppy-torso-2eme-edition-du-projet-etudiant-de-lensam





Figure 24.

• Poppy entre dans la danse (Poppy enters the dance)

The project "Poppy enters the dance" (Canope 33) took place for the second year. It uses the humanoid robot Poppy. This robot is able to move and experience the dance. The purpose of this project is to allow children to understand the interactions between science and choreography, to play with the random and programmable, to experience movement in dialogue with the machine. At the beginning of the project they attended two days of training on the humanoid robot (Inria - Poppy Education). During the project, they met the choreographer Eric Minh Cuong Castaing and the engineer Segonds Theo (Inria - Poppy Education).

You can see a description and an overview of the project here: https://www.youtube.com/watch?v=XfxXaq899kY

7.5.3. IniRobot: Educational Robotics in Primary Schools

Participants: Didier Roy [correspondant], Pierre-Yves Oudeyer.

IniRobot (a project done in collaboration with EPFL/Mobsya) aims to create, evaluate and disseminate a pedagogical kit which uses Thymio robot, open-source and low cost, for teaching computer science and robotics.

IniRobot Project consists to produce and diffuse a pedagogical kit for teachers and animators, to help to train them directly or by the way of external structures. The aim of the kit is to initiate children to computer science and robotics. The kit provides a micro-world for learning, and takes an enquiry-based educational approach, where kids are led to construct their understanding through practicing an active investigation methodology within teams. It is based on the use of the Thymio II robotic platform. More details about this projects were published in RIE 2015 [50], which presents the detailed pedagogical objectives and a first measure of results showing that children acquired several robotics-related concepts. See also http://www.inirobot.fr.

Deployment: After 30 months of activity, IniRobot is used by about 1800 adults and 20 000 children in 72 cities of France. Example of action in university: MEEF teacher training for the hope of Aquitaine. Example of action in school: training of all Gironde Pedagogical ICT Advisors, covering nearly 1000 schools. Example of action in the extracurricular time: training 82 facilitators TAP cities of Talence, Pessac, Lille, ..., CDC Gates of inter-seas. Example of national action: Training of the digital mediators of the 8 Inria centers.

7.5.3.1. Partnership

The project is carried out in main collaboration with the LSRO Laboratory from EPFL (Lausanne) and others collaborations with French National Education/Rectorat d'Aquitaine, with Canopé Educational Network, with ESPE (teacher's school) Aquitaine, ESPE Martinique, ESPE Poitiers, National Directorate of Digital Education

7.5.3.2. Created pedagogical documents and resources

- Inirobot pedagogical kit [24]: This pedagogical booklet provides activities scenarized as missions to do. An update of Inirobot pedagogical kit : https://dm1r.inria.fr/uploads/default/original/1X/70037bdd5c290e48c7ec4cb4f26f0e426a4b4cf6.pdf Another pedagogical booklet has been also created by three pedagogical advisers for primary school, with pedagogical instructions and aims, under ou supervision. new pedagogical kit is available, "Inirobot Scolaire, Langages et robotique", which extends Inirobot in a full primary school approach. http://tice33.ac-bordeaux.fr/Ecolien/ASTEP/tabid/5953/language/fr-FR/Default.aspx
- Inirobot website and forum https://dm1r.inria.fr/inirobot or http://www.inirobot.fr With this website, teachers, animators and general public can download documents, exchange about their use of inirobot's kit.
- Publication about Inirobot and Poppy Education La robotique éducative au service du développement de la pensée informatique Didier Roy (2017) dans "L'informatique et le numérique dans la classe (Presses universitaires de Namur)" http://pun.be/fr/livre/?GCOI=99993100805880 Inirobot et Poppy education, deux dispositifs robotiques open source pour l'enseignement des Sciences du Numérique Didier Roy (2017) Revue Sésamath 55 http://revue.sesamath.net/spip.php?article922 Poppy Education : un dispositif robotique open source pour l'enseignement de l'informatique et de la robotique Stéphanie Noirpoudre, Didier Roy, Thibault Desprez, Théo Segonds, Damien Caselli, Pierre-Yves Oudeyer (2017) EIAH 2017, Strasbourg, France [81] Thymio nella didattica in Francia e in Svizzera, Morgane Chevalier, Gordana Gerber, Didier Roy (2017) Pedagogika, Anno XXI, n1 Gennaio/Febbraio/Marzo http://www.pedagogia.it/ Une brève histoire de la robotique, d'Electric Dog à Poppy, Revue du Palais de la Découverte (2017) : http://www.palais-decouverte.fr/ fr/ressources/revue-decouverte/n-412-septembre-octobre-2017/articles/

7.5.3.3. Scientific mediation

Inirobot is very popular and often presented in events (conferences, workshops, ...) by us and by others.

7.5.3.4. Symposium robotics

With Poppy Education, Inirobot is a main line in our colloquium "Robotics and Education" (http://dmlr.fr/)

7.5.3.5. Spread of Inirobot activities

Inirobot activities are inside several projects : Dossier 123 codez from Main à la Pâte Fundation, Classcode project, ...

7.5.3.6. Future MOOC Thymio

A new project is coming, MOOC Thymio, in collaboration with Inria Learning Lab and EPFL (Lausanne, Switzerland), on FUN platform and edX EPFL Platform.) To teach how to use Thymio robot in education.
MANAO Project-Team

7. New Results

7.1. Analysis and Simulation

7.1.1. A Two-Scale Microfacet Reflectance Model Combining Reflection and Diffraction,,

Adequate reflectance models are essential for the production of photorealistic images. Microfacet reflectance models predict the appearance of a material at the macroscopic level based on microscopic surface details. They provide a good match with measured reflectance in some cases, but not always. This discrepancy between the behavior predicted by microfacet models and the observed behavior has puzzled researchers for a long time. In these papers [14], [24], [19], we show that diffraction effects in the micro-geometry provide a plausible explanation. We describe a two-scale reflectance model (cf. Figure 8), separating between geometry details much larger than wavelength and those of size comparable to wavelength. The former model results in the standard Cook-Torrance model. The latter model is responsible for diffraction effects. Diffraction effects at the smaller scale are convolved by the micro-geometry normal distribution. The resulting two-scale model provides a very good approximation to measured reflectances.



Figure 8. Material reflectance properties are caused by small variations in surface geometry. We separate these surface variations into micro-geometry, of size larger than the wavelength of visible light, and nano-geometry, of size comparable to the wavelength. The latter produces diffraction effects, with wavelength-dependent effects. The former corresponds to the classical Cook-Torrance lobe. We explain how these two levels interact and show that combined together, they reproduce measured materials faithfully, including subtle color shifts.

7.1.2. A Practical Extension to Microfacet Theory for the Modeling of Varying Iridescence

Thin film iridescence permits to reproduce the appearance of leather. However, this theory requires spectral rendering engines (such as Maxwell Render) to correctly integrate the change of appearance with respect to viewpoint (known as goniochromatism). This is due to aliasing in the spectral domain as real-time renderers only work with three components (RGB) for the entire range of visible light. In this work [11], we show how to anti-alias a thin-film model, how to incorporate it in microfacet theory, and how to integrate it in a real-time rendering engine. This widens the range of reproducible appearances with microfacet models (cf. Figure 9).



Figure 9. Material appearance such as that of leather is usually reproduced with microfacet models in computer graphics. A more realistic result is achieved by adding a thin-film coating that produces iridescent colors. We replace the classic Fresnel reflectance term with a new Airy reflectance term that accounts for iridescence due to thin-film interference. Our main contribution consists in an analytical integration of the high-frequency spectral oscillations exhibited by Airy reflectance, which is essential for practical rendering in RGB. When the scene is rotated, goniochromatic effects such as subtle purple colors may be observed at grazing angles.

7.2. From Acquisition to Display

7.2.1. Diffraction effects detection for HDR image-based measurements,

Modern imaging techniques have proved to be very efficient to recover a scene with high dynamic range (HDR) values. However, this high dynamic range can introduce star-burst patterns around highlights arising from the diffraction of the camera aperture. The spatial extent of this effect can be very wide and alters pixels values, which, in a measurement context, are not reliable anymore. To address this problem, we introduce [21], [15] a novel algorithm that, utilizing a closed-form PSF, predicts where the diffraction will affect the pixels of an HDR image, making it possible to discard them from the measurement. Our approach gives better results (cf. Figure 10) than common deconvolution techniques and the uncertainty values (convolution kernel and noise) of the algorithm output are recovered.

7.2.2. A low-cost multitouch spherical display: hardware and software design,

Spherical mulitouch displays offer exciting possibilities but are still costly. In this work [17], we first describe hardware and software considerations to build a more affordable one, with off-the-shelf optical components and 3D printed elements. We exploit the technology of laser-beam steering projectors and use optical tracking for multitouch. Besides, although spherical displays become more and more pervasive, the design of interactive content for these displays still remains difficult as it requires most developers to get familiar with specific tools for managing the output and input. We thus present [18] a framework for developing applications for multitouch spherical displays that makes it possible to create interactive content by programming standard GUI applications, as for example interactive web pages. The principal idea is to adapt the window output and interaction input of classical GUIs outside the application. To this end, our framework consists of two standalone applications where the first one captures the window output and changes the projection via GPU shaders, and the second one adapts the input with a Node.js server and sends interaction and mouse events. In this way, the same application runs on a standard desktop, and on the spherical display. Advantages of our approach include fast prototyping, and the fact that masses of developers can create applications for spherical displays just as if it were, for example, classical web applications. We believe that our framework will contribute to making spherical displays even more pervasive in the future.



Figure 10. Results of the algorithm applied on real HDR images for various camera configurations, with input parameters $D_b = 10$ and $\rho = 5\%$. The wavelengths used for each color channel are $[\lambda_R, \lambda_G, \lambda_B] = [600nm, 540nm, 470nm]$. The segmentation images show the discarded pixels (red), the valid pixels (green), and the under-exposed ones (black). If the HDR images exhibits obvious star shaped patterns, the algorithm detects it, and they are finally removed. Such result is qualitative in nature, because there is no reference HDR image without diffraction. False predictions are present in the first two cases (l), where the diffraction prediction seems rotated from the real one. This problem emerges from the misfit of the lens diaphragm, as discussed in subsection 7.7.1.

7.3. Rendering, Visualization and Illustration

7.3.1. Example-Based Expressive Animation of 2D Rigid Bodies



Figure 11. Given a set of frames F^S coming from reference 2D rigid body source animations, corresponding hand-animated exemplars E, and a new target animation F^T , the synthesis algorithm relates physical parameters in F^S and F^T to produce the output stylized sequence F^O that resembles F^E .

We have presented [12] a novel approach to facilitate the creation of stylized 2D rigid body animations. Our approach can handle multiple rigid objects following complex physically-simulated trajectories with collisions, while retaining a unique artistic style directly specified by the user. Starting with an existing target animation (e.g., produced by a physical simulation engine) an artist interactively draws over a sparse set of frames, and the desired appearance and motion stylization is automatically propagated to the rest of the sequence (fig. 11). The stylization process may also be performed in an off-line batch process from a small set of drawn sequences. To achieve these goals, we combine parametric deformation synthesis that generalizes and reuses hand-drawn exemplars, with non-parametric techniques that enhance the hand-drawn appearance of the synthesized sequence. We demonstrate the potential of our method on various complex rigid body animations which are created with an expressive hand-drawn look using notably less manual interventions as compared to traditional techniques.

7.3.2. Edge- and Substrate-based Effects for Watercolor Stylization,



Figure 12. Our methods allow new and improved edge- and substrate-based effects for watercolor stylization: edge darkening (red), gaps (blue), overlaps (green) and dry-brush (yellow). Still Life, model by Dylan Sisson ©Pixar Animation Studios.

We investigated [20] characteristic edge-and substrate-based effects for watercolor stylization. These two fundamental elements of painted art play a significant role in traditional watercolors and highly influence the pigment's behavior and application. Yet a detailed consideration of these specific elements for the stylization of 3D scenes has not been attempted before. Through this investigation, we contributed to the field by presenting ways to emulate two novel effects: dry-brush and gaps & overlaps. By doing so, we also found ways to improve upon well-studied watercolor effects such as edge-darkening and substrate granulation. Finally, we integrated controllable external lighting influences over the watercolorized result, together with other previously researched watercolor effects. These effects are combined through a direct stylization pipeline [22] to produce sophisticated watercolor imagery (fig. 12), which retains spatial coherence in object-space and is locally controllable in real-time.

7.3.3. Specular Motion and 3D Shape Estimation

Dynamic visual information facilitates three-dimensional shape recognition. It is still unclear, however, whether the motion information generated by moving specularities across a surface is congruent to that available from optic flow produced by a matte-textured shape. Whereas the latter is directly linked to the first-order properties of the shape and its motion relative to the observer, the specular flow, the image flow generated by a specular object, is less sensitive to the object's motion and is tightly related to second-order properties of the shape. We therefore hypothesize [13] that the perceived bumpiness (a perceptual attribute related to curvature magnitude) is more stable to changes in the type of motion in specular objects compared with their matte-textured counterparts. Results from two two-interval forced-choice experiments in which observers judged the perceived bumpiness of perturbed spherelike objects support this idea and provide an additional layer of evidence for the capacity of the visual system to exploit image information for shape inference.

7.3.4. The Perception of Hazy Gloss

Most previous work on gloss perception has examined the strength and sharpness of specular reflections in simple bidirectional reflectance distribution functions (BRDFs) having a single specular component. However, BRDFs can be substantially more complex and it is interesting to ask how many additional perceptual dimensions there could be in the visual representation of surface reflectance qualities. To address this, we tested [16] materials with two specular components that elicit an impression of hazy gloss. Stimuli were renderings of irregularly shaped objects under environment illumination, with either a single specular BRDF component, or two such components, with the same total specular reflectance but different sharpness parameters, yielding both sharp and blurry highlights simultaneously. Differently shaped objects were presented side by side in matching, discrimination, and rating tasks. Our results show that observers mainly attend to the sharpest reflections in matching tasks, but they can indeed discriminate between single-component and two-component specular materials in discrimination and rating tasks. The results reveal an additional perceptual dimension of gloss-beyond strength and sharpness-akin to "haze gloss". However, neither the physical measurements of Hunter and Harold nor the kurtosis of the specular term predict perception in our tasks. We suggest the visual system may use a decomposition of specular reflections in the perception of hazy gloss, and we compare two possible candidates: a physical representation made of two gloss components, and an alternative representation made of a central gloss component and a surrounding halo component.

POTIOC Project-Team

7. New Results

7.1. HOBIT

Participants: David Furio, Benoit Coulais, Martin Hachet

Practical work in optics learning allows supporting the construction of knowledge, in particular when the concept to be learned remains diffuse. To overcome the limitations of the current experimental setups, we have designed a hybrid system that combines physical interaction and numerical simulation. This system relies on 3D-printed replicas of optical elements, which are augmented with pedagogical information (see Figure 3). In a first step, we have focused on the well-known Michelson interferometer experiment, widely studied in under graduate programs of Science. A 3-months user study with 101 students and 6 teachers showed that, beyond the practical aspects offered by this system, such an approach enhances the technical and scientific learning compared to a standard Michelson interferometer experiment. This work has been published at CHI 2017 [24], and the paper obtained a Best Paper - Honorable Mention Award.

Currently, we are developping a second version of HOBIT. This new version will let us simulate and augment multiple experiments related with optics, like polarization or Young's interferometer.



Figure 3. HOBIT: Hybrid Optical Bench for Innovative Teaching

7.2. Inner Garden

Participants: Joan Sol Roo, Renaud Gervais, Jeremy Frey, Martin Hachet

Digital technology has completely integrated our daily lives; we use it for entertainment, productivity and our social lives. However, the potential of leveraging technology to improve its users' overall happiness and life satisfaction is still largely untapped. Mindfulness, the act of paying a deliberate and non-judgmental attention to the present moment, has been shown to have a positive impact on a person's subjective well-being. With this in mind we created Inner Garden, an ambient mixed reality installation, inspired by a zen garden, comprised of an augmented sandbox along with a virtual reality modality to support mindful experiences (Figure 4 . By shaping the sand, the user creates a living miniature world that is projected back onto the sand. Moreover, using a VR headset, she can take a moment to herself by actually going inside her own garden to meditate. The natural elements of the garden are connected to real-time physiological measurements, such as breathing, helping staying focused on the body. We evaluated the system through a first user study and consulted meditation teachers, who envisioned the use of the garden in their teaching, especially for novice practitioners. The reception of the system seems to indicate that technology can, when designed carefully, both engage the users and foster well-being.

This work has been published at CHI 2017 [32], and the paper obtained a Best Paper - Honorable Mention Award.



Figure 4. Inner Garden, an ambient mixed reality installation to support mindful experiences

7.3. Art and Science

Participants: Clémentine Petit, Maxime Agor, Martin Hachet

Potioc collaborates with artists for Art and Science projects. These projects are supported by a dedicated program at Idex - Université de Bordeaux.

The first one, Kilometre 2.0, is a joint project with Cécile Léna https://www.lenadazy.fr who is a visual artist and scenographer. We have augmented physical mockups with digital objects (see Figure 5). More concretely, our system detects train tickets being manipulated above a miniature scenery. Depending on their locations, dedicated movies are projected exactly on them, in the 3D space. Other sound and smoke effects are also generated. This artistic and interactive setup has been showed during the FACTS festival in November 18.



Figure 5. Spatial augmented reality in an artistic installation

The second project is conducted in collaboration with Antoine Clée from Le Cirque Inachevé http://www. lecirqueinacheve.fr. Our objective is to explore new forms of juggling where the balls are not constrained by gravity anymore. More precisely, the balls are held by nano-quadcopters (drones) as illustrated in Figure 6. The juggler controls these drones by way of tracked gloves and associated interaction techniques. We presented this work as a poster presentation at IHM 2017 [16]. It has also been demonstrated during a live performance at OARA in Bordeaux on November 23rd, part of the FACTS festival.



Figure 6. A juggler interacting with flying balls

7.4. New version of Teegi and its pedagogical potential

Participants: Jeremy Frey, Fabien Lotte and Martin Hachet

Cerebral activity is an intangible physiological process that is difficult to apprehend, especially for children. To overcome this difficulty, Teegi was designed as a new type of educational support. This tangible interface enables children to discover the relationship between brain activity and the functions of the human body.

This year, we have designed a new version of Teegi (see Figure 7. It is 3D printed, and embeds a Raspberry Pi 3 and NiMh batteries (autonomy of approximately 2 hours). A python script on the Raspberry Pi handles the 402 LEDs (Adafruit Neopixel) covering the "head", which are connected to its GPIO pins. For a smoother display, the light of the LEDs is diffused by a 3mm thick cap made of acrylic glass. Two 8-by-8 white LEDs matrices picture the eyes. The script also commands the servomotors placed in the hands and feet, 4 Dynamixel XL320.

We used this new version of Teegi as a case of study for developing a multi-methods research approach to estimate the pedagogical potential of a tangible interface used in a real-life educational context. Using this methodology, we conducted a user study (N=29) that highlighted the strengths of this interface, both in terms of its usability and its impact on learning. Moreover, results revealed possible improvements to further increase pedagogical effectiveness. This type of interface, as well as the evaluation method that we propose, contribute to extending our knowledge concerning the pedagogical use of new interactive tools at school.

This work was published at IHM 2017 [22], and the accompanying demo won the best demo award. Teegi was also demonstrated at CHI 2017 [23].

7.5. One Reality

Participants: Joan Sol Roo and Martin Hachet





This project explores the combination of Physical and Virtual Reality through the usage of Mixed Reality. Early explorations involved the usage of Spatial Augmented Reality in combination with Virtual Reality, two technologies with complementary characteristics that evolved separately in the past. Spatial Augmented Reality (SAR) augments the environment using projectors or screens, without the need of user instrumentation. By keeping a single unified frame of reference, it supports social interaction and natural perception of the space, but the augmentation is limited by physical constraints (e.g., it requires a surface to display information). Immersive head mounted displays on the other hand are not limited by the physical properties of the environment, yet they isolate the user from their environment. We have proposed a unified frame of reference for both SAR and immersive displays, where the users can select the visualization that is best suited for a given task (Figure 8). This enables both asymmetric collaboration between users, and back-and-forths for a single user. These explorations were followed by the combination of additional modalities, in an incremental fashion. This way, one or more users can chose the desired modalities, and immerse themselves as much as the task requires. As a result, the virtual world can be framed in relationship with the physical one.

A preliminary version of this work was presented at 3DUI 2017 [30] where we obtained a best paper - honorable mention- award. An extended version of this work was then presented at UIST 2017 [33].

In order for such systems to succeed, it is required that users are able to create unified mental models out of heterogeneous representations. To better understand how humans perceive hybrid systems as the one described above, we conducted two studies. They focused on the users' performance on heterogeneous systems (using Spatial Augmented Reality and immersive Virtual Reality displays), and combining viewpoints (egocentric and exocentric). The results show robust estimation capabilities across conditions. This work has been (conditionally) accepted at CHI 18.

7.6. Collaboration in VR

Participants: Damien Clergeaud and Pascal Guitton

The aerospace industry is no longer composed of local and individual businesses. Due to the complexity of the products (their size, the number of components, the variety of systems and regulation constraints), the design of an aircraft or a launcher involves a considerable number of engineers with various fields of expertise. Furthermore, aerospace companies often have industrial facilities all over the world. In such a complex setting, it is necessary to build virtual experiments that can be shared between different remote sites. Specific problems then arise, particularly in terms of the perception of other immersed users and of interaction tasks involving several immersed users.



Figure 8. One Reality combines the real and virtual worlds

We work with Airbus Group in order to design efficient collaborative interaction methods. These collaborative sessions allow multiple sites to be connected within the same virtual experiment and enable experts from different fields to be immersed simultaneously. For instance, if a problem occurs during the final stages of a launcher assembly, it may be necessary to bring together experts on different sites who were involved in previous steps (initial design, manufacturing processes). In the context of this collaboration, we working on various projects:

- Design of basic communication tools for the aerospace context.
- Pano: a 360° visualization system that's facilitate the communication in the case of guiding someone else [19],
- Design of an Annotation System for taking notes in VR [18].



Figure 9. An immersed user has to perform a virtual task in a complex environment. In order to help the user to be fully aware of the VE, another immersed operator may guide him using a Through-The-Lens metaphor.

7.7. Tangible interaction and augmented reality for collaborative learning

Participants: Philippe Giraudeau and Martin Hachet

Part of the e-Fran project e-Tac, we explore approaches based on the hybridization of physical and digital content for mind-mapping activities at schools. Based on the literature in the fields of cognitive science and HCI, we have designed a mixed-reality (MR) interface called Reality-Map (Figure 10. We conducted a pilot study with 11 participants suggesting that learning and manipulating information about the brain and their cognitive functions could be improved by the use of such a MR interface compared to a traditional WIMP interface [45]. We are now extending this approach with the partners of the project to design and develop a new pedgogical tool that will be evaluated in classrooms.



Figure 10. Digital and physical objects combined for collaborative learning

7.8. A model of Mental-Imagery BCI

Participants: Camille Jeunet and Fabien Lotte

Mental-Imagery based Brain-Computer Interfaces (MI-BCIs) enable users to control applications using their brain activity alone, by realising mental imagery tasks. Although promising, MI-BCIs remain barely used outside laboratories, notably due to the difficulties users encounter when attempting to control them. We claim that understanding and improving the user training process could greatly improve users' MI-BCI control abilities. Yet, to better understand the training process, we need a model of the factors impacting MI-BCI performance. In other words, we need to understand which traits and states impact MI-BCI performance, how these factors interact and how to influence them to improve this performance. Such a model would enable us to design adapted and adaptive training protocols, to guide europhysiological analyses or design informed classifiers, among others. In this paper we propose a theoretical model of MI-BCI tasks, which is the first step towards the design of this full cognitive and computational model. This work was published in the International BCI conference [46].

7.9. PEANUT - Personalized Emotional Agent for Neurotechnology User Training

Participants: Léa Pillette, Camille Jeunet and Fabien Lotte



Figure 11. A conceptual model of Mental-Imagery BCI user training

Mental-Imagery based Brain-Computer Interfaces (MI-BCI) are neurotechnologies enabling users to control applications using their brain activity alone. Although promising, they are barely used outside laboratories because they are poorly reliable, partly due to inappropriate training protocols. Indeed, it has been shown that tense and non-autonomous users, that is to say those who require the greatest social presence and emotional support, struggle to use MI-BCI. Yet, the importance of such support during MI-BCI training is neglected. Therefore we designed and tested PEANUT, the first Learning Companion providing social presence and emotional support dedicated to the improvement of MI-BCI user-training. PEANUT was designed based on the literature, data analyses and user-studies. Promising results revealed that participants accompanied by PEANUT found the MI-BCI system significantly more usable. This work was published in the International BCI conference [29].



Figure 12. The PEANUT learning companion for BCI user training

7.10. The Impact of Flow on BCI user training

Participants: Jelena Mladenovic, Jérémy Frey, Manon Bonnet-Save, Fabien Lotte

Major issues in Brain Computer Interfaces (BCIs) include low usability and poor user performance. This paper tackles them by ensuring the users to be in a state of immersion, control and motivation, called state of flow. Indeed, in various disciplines, being in the state of flow was shown to improve performances and learning. Hence, we intended to draw BCI users in a flow state to improve both their subjective experience and their performances. In a Motor Imagery BCI game, we manipulated flow in two ways: 1) by adapting the task difficulty and 2) by using background music. Results showed that the difficulty adaptation induced a higher flow state, however music had no effect. There was a positive correlation between subjective flow scores and offline performance, although the flow factors had no effect (adaptation) or negative effect (music) on online performance. Overall, favoring the flow state seems a promising approach for enhancing users' satisfaction, although its complexity requires more thorough investigations. This work was published at the international BCI conference [27]



Figure 13. Adaptive BCI user training to maximize the flow state

7.11. New Performance metrics to study BCI user training

Participants: Fabien Lotte and Camille Jeunet

While promising for many applications, Electroencephalography (EEG)-based Brain-Computer Interfaces (BCIs) are still scarcely used outside laboratories, due to a poor reliability. It is thus necessary to study and fix this reliability issue. Doing so requires to use appropriate reliability metrics to quantify both signal processing and user learning performances. So far, Classification Accuracy (CA) is the typical metric used for both aspects. However, we argue in this paper that CA is a poor metric to study how well users are learning to use the BCI. Indeed CA is notably unspecific, discrete, training data and classifier dependent, and as such may not always reflect successful EEG pattern self-modulation by the user. We thus propose new performance metrics to specifically measure how distinct and stable the EEG patterns produced by the user are. By reanalyzing EEG data with these metrics, we indeed confirm that CA may hide some learning effects or hide the user inability to self-modulate a given EEG pattern. This was published at the international BCI Conference [25].

7.12. Joint EEG-fMRI Neurofeedback training

Participants: Fabien Lotte

Neurofeedback is a promising tool for brain rehabilitation and peak performance training. Neurofeedback approaches usually rely on a single brain imaging modality such as EEG or fMRI. Combining these modalities for neurofeedback training could allow to provide richer information to the subject and could thus enable him/her to achieve faster and more specific self-regulation. Yet unimodal and multimodal neurofeedback have never been compared before. In the present work, we introduce a simultaneous EEG-fMRI experimental protocol in which participants performed a motor-imagery task in unimodal and bimodal NF conditions. With this protocol we were able to compare for the first time the effects of unimodal EEG-neurofeedback and fMRI-neurofeedback versus bimodal EEG-fMRI-neurofeedback by looking both at EEG and fMRI activations. We also propose a new feedback metaphor for bimodal EEG-fMRI-neurofeedback that integrates both EEG and fMRI signal in a single bi-dimensional feedback (a ball moving in 2D). Such a feedback is intended to relieve the cognitive load of the subject by presenting the bimodal neurofeedback task as a single regulation task instead of two. Additionally, this integrated feedback metaphor gives flexibility on defining a bimodal neurofeedback target. Participants were able to regulate activity in their motor regions in all NF conditions. Moreover, motor activations as revealed by offline fMRI analysis were stronger during EEG-fMRIneurofeedback than during EEG-neurofeedback. This result suggests that EEG-fMRI-neurofeedback could be more specific or more engaging than EEG-neurofeedback. Our results also suggest that during EEG-fMRIneurofeedback, participants tended to regulate more the modality that was harder to control. Taken together our results shed first light on the specific mechanisms of bimodal EEG-fMRI-neurofeedback and on its addedvalue as compared to unimodal EEG-neurofeedback and fMRI-neurofeedback.

This work in collaboration with Inria teams Hybrid, Visage and Athena, was published in the journal Frontiers in Neuroscience [14].

7.13. Robust EEG spatial filters for single trial regression

Participants: Fabien Lotte

In the field of Brain-Computer Interfaces (BCI), robust methods for the decoding of continuous brain states are of great interest as new application fields are arising. When capturing brain activity by an electroencephalogram (EEG), the Source Power Comodulation (SPoC) algorithm allows to compute spatial filters for the decoding of a continuous variable. However, dealing with high-dimensional EEG data that suffer from low signal-to-noise ratio, the method reveals instabilities for small training data sets and is prone to overfitting. In this paper, we introduce a framework for applying Tikhonov regularization to the SPoC approach in order to restrict the solution space of filters. Our findings show that an additional trace normalization of the included covariance matrices is a necessary prerequisite to tune the sensitivity of the resulting algorithm. In an offline analysis with data from N=18 subjects, the introduced trace normalized and Tihonov regularized SPoC variant (NTR-SPoC) outperforms the standard SPoC method for the majority of individuals. With this proof-of-concept study, a generalizable regularization framework for SPoC has been established which allows to implement a variety of different regularization strategies in the future. This work in collaboration with Freiburg University, Germany, was published at the international BCI Conference [26].

7.14. Optimal motor tasks for Sensorimotor BCI calibration

Participants: Fabien Lotte

SensoriMotor Rhythm (SMR)-based Brain-Computer Interfaces (BCI) are among the most used ElectroEncephaloGraphy (EEG) BCI systems. However, such systems have low performance and many of their users are "non-responders". There is thus a need to understand the limitations of current SMR-BCI and to improve them. Many of them use machine learning. They are typically calibrated on EEG signals collected while the users are performing Motor Imagery (MI), i.e., imagining limb movements. Once calibrated, they also use MI as control strategy. However, for many first time users of SMR-BCI, performing MI is new and difficult, and they may be unable to perform clear MI. Thus, using MI for calibration may result in suboptimal EEG features and corresponding real-time feedback. Therefore, we aim at elucidating whether MI tasks are the best motor tasks to use for calibration and control in SMR-BCI. To do so, we collected EEG signals from subjects

instructed to perform four different motor tasks and a rest task, for multiple trials. In particular, subjects have to 1) execute real feet movements; 2) imagine feet movements (walking); 3) observe feet movements (walking), in a first person view and 4) observe feet movements while imagining them at the same time. Preliminary results revealed that for some subjects, calibrating EEG spatial filters on real motor movements can lead to better performances with an MI-BCI than calibrating them on MI tasks. This thus warrant further investigation into the calibartion tasks in SMR-BCI. This preliminary work, in collaboration with RIKEN Brain Science Institute in Japan, was presented as a poster at RTFIN 2017 [48].

7.15. A review of Rapid Serial Visualization Protocol-based BCI

Participants: Fabien Lotte

Rapid serial visual presentation (RSVP) combined with the detection of event related brain responses facilitates the selection of relevant information contained in a stream of images presented rapidly to a human. Event related potentials (ERPs), measured non-invasively with electroencephalography (EEG), can be associated with infrequent target stimuli(images) in groups of images, potentially providing an interface for human-machine symbiosis, where humans can interact and interface with a computer without moving and which may offer faster image sorting than scenarios where humans are expected to physically react when a target image is detected. Certain features of the human visual system impact on the success of the RSVP paradigm. Pre-attentive processing supports the identification of target information 100ms following information presentation. This paper presents a comprehensive review and evaluation of research in the broad field of RSVP-based brain-computer interfaces (BCIs). Applications that use RSVP-based BCIs are classified based on the operation mode whilst protocol design considerations are critiqued. Guidelines for using the RSVP-based BCI paradigms are defined and discussed, with a view to further standardization of methods and experimental evidence gathering to support the use of RSVP-based BCIs in practice. This review in collaboration with Ulster University, UK, was published in Journal of Neural Engineering [13].

7.16. Active Inference-based design of adaptive P300-Speller BCIs

Participants: Jelena Mladenovic, Fabien Lotte

Recent developments in computational neuroscience gave rise to an efficient generic framework to implement both optimal perceptual (Bayesian) inference and choice behaviour. This framework named Active Inference rests on minimizing free energy or surprise [3]. We suggest it could be used to implement efficient adaptive Brain-Computer Interfaces (BCIs). We briefly illustrate it on a simulated P300-speller task. This work in collaboration with Inserm Lyon, was published in the first Neuro Adaptive Technology conference [28].

7.17. BCI Handbook

Participants: Fabien Lotte

Together with Chang S Nam (USA) and Anton Nijholt (Netherlands/Malaysia), we edited an Handbook of BCI technologies [41]. This handbook is a valuable resource to anyone involved with improvement of people's lives by replacing, restoring, supplementing and improving motor action, and understanding the neural bases of such functions. While there are several other resources available, there is no handbook such as this one. This handbook addresses the recent and rapid changes in the field of braincomputer interfaces (BCIs). Due to these changes interest in BCI has grown enormously, including interest from computer science researchers with a background in computational intelligence, human-computer interaction, and researchers in entertainment technology.

7.18. Augmented Reality Maps for Visually Impaired People

Participants: Anke Brock

VISTE builds on the previous development of the GEOTHNK platform (Kavouras et al., 2016). The VISTE framework and associated resources and tools focus on collaborative learning of spatial concepts and skills for sighted and VI students to foster inclusion within mainstream education. VISTE will empower students with VI to acquire spatial skills through specially designed learning activities as well as through an augmented reality prototype. At Inria Bordeaux, we have designed and implemented an augmented reality prototype that can be used as spatial thinking training tool in special education schools. It makes use of the PapARt technology, an OpenSource augmented reality framework⁰. Current low-tech Orientation & Mobility (O&M) tools for visually impaired people, e.g. tactile maps, possess limitations. Interactive accessible maps have been developed to overcome these ⁰. However, most of them are limited to exploration of existing maps, and have remained in laboratories. Using a participatory design approach, we have worked closely with 15 visually impaired students and 3 O&M instructors over 6 months. We iteratively designed and developed an augmented reality map destined at use in O&M classes in special education centers. This prototype combines projection, audio output and use of tactile tokens, and thus allows both map exploration and construction by low vision and blind people. Our user study demonstrated that all students were able to successfull fy use the prototype, and showed a high user satisfaction. A second phase with 22 international special education teachers allowed us to gain more qualitative insights. This work shows that augmented reality has potential for improving the access to education for visually impaired people. A publication about this map prototype and the user study has been (conditionally) accepted at CHI'18⁰ Learn more about VISTE project : http://visteproject.eu/https:// team.inria.fr/potioc/viste-empowering-spatial-thinking-of-students-with-visual-impairment/

7.19. Accessibility of e-learning systems

Participants: Pierre-Antoine Cinquin and Pascal Guitton

New digital teaching systems such as MOOCs are taking an increasingly important place in current teaching practices. Unfortunately, accessibility for people with disabilities is often forgotten, which excludes them, particularly those with cognitive impairments for whom accessibility standards are fare from being established. This is truly unfortunate as the interest of using these specialized practices for this audience is scientifically proven.

To overcome these limitations, we propose new design principles based on knowledge in the areas of accessibility (Ability-based Design and Universal Design), digital pedagogy (Instruction Design with functionalities that reduce the cognitive load : navigation by concept, slowing of the flow...), specialized pedagogy (Universal Design for Learning, eg, automatic note-taking, and Self Determination Theory, e.g., configuration of the interface according to users needs and preferences) and psychoducational interventions (eg, support the joint teacher-learner attention), but also through a participatory design approach involving students with disabilities and experts in the field of disability. From these framework, we have designed interaction features which have been implemented in a specific MOOC player called Aïana. Moreover, we have produced a MOOC on digital accessibility which is published on the national MOOC platform (FUN) using Aïana (2 sessions in 2016 and 2017) https://mooc-francophone.com/cours/mooc-accessibilite-numerique/.

⁰Laviole & Hachet, 2012

⁰Ducasse et al., 2018

⁰Jeremy Albouys-Perrois, Jeremy Laviole, Carine Briant, and Anke Brock. Towards a Multisensory Augmented Reality Map for Blind and Low Vision People: a Participatory Design Approach. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). ACM, New York, USA



Figure 14. The Aïana MOOC player.