RESEARCH CENTER
Bordeaux - Sud-Ouest

FIELD

Activity Report 2014

Section New Results

Edition: 2015-03-24

LFANT Project-Team

5. New Results

5.1. Highlights of the Year

Aurel Page has defended his PhD thesis on *Méthodes explicites pour les groupes arithmétiques* [12] in July 2014. Nicolas Mascot has defended his PhD thesis on *Computing modular Galois representations* [11], in July 2014.

5.2. Class groups and other invariants of number fields

Participants: Karim Belabas, Jean-Paul Cerri, Pierre Lezowski.

In [21], P. Lezowski describes the explicit computation of the Euclidean minimum of a number field. It has been published in Mathematics of Computation.

Ohno and Nakagawa have proved, relations between the counting functions of certain cubic fields. These relations may be viewed as complements to the Scholz reflection principle, and Ohno and Nakagawa deduced them as consequences of 'extra functional equations' involving the Shintani zeta functions associated to the prehomogeneous vector space of binary cubic forms. In [26], Henri Cohen, Simon Rubinstein-Salzedo and Frank Thorne generalize their result by proving a similar identity relating certain degree fields with Galois groups D and F respectively, for any odd prime, and in particular we give another proof of the Ohno–Nakagawa relation without appealing to binary cubic forms.

The article [16] by H. Cohen and F. Thorne, H. Cohen on Dirichlet series associated to cubic fields with given resolvent has been published. This article gives an explicit formula for the Dirichlet series $\sum_K |\Delta(K)|^{-s}$, where the sum is over isomorphism classes of all cubic fields whose quadratic resolvent field is isomorphic to a fixed quadratic field k.

This work is extended in [15] where H. Cohen give efficient numerical methods for counting exactly the number of D_ℓ number fields of degree ℓ with given quadratic resolvent, for calculating the constants occurring in their asymptotic expansions, and give tables for typical cases.

5.3. Number and function fields

Participants: Jean-Marc Couveignes, Karim Belabas.

In the article [29], J. Brau study the growth of the Galois invariants of the p-Selmer group of an elliptic curve in a degree p Galois extension. He shows that this growth is determined by certain local cohomology groups and determine necessary and sufficient conditions for these groups to be trivial.

In the article [30] written with J. Nathan, J. Brau study the modular curve $X'(6)$ of level 6 defined over $\mathbb{Q}$ whose $\mathbb{Q}$-rational points correspond to j-invariants of elliptic curves E over $\mathbb{Q}$ for which $\mathbb{Q}(E[2])$ is a subfield of $\mathbb{Q}(E[3])$. They characterize the j-invariants of elliptic curves with this property by exhibiting an explicit model of $X'(6)$. $X'(6)(\mathbb{Q})$ gives an infinite family of examples of elliptic curves with non-abelian "entanglement fields," which is relevant to the systematic study of correction factors of various conjectural constants for elliptic curves over $\mathbb{Q}$.

5.4. Quaternion algebras

Participants: Jean-Paul Cerri, Pierre Lezowski, Aurel Page.

In the article [14] written with J. Chaubert, J.-P. Cerri and P. Lezowski study totally indefinite Euclidean quaternion fields over a number field K, that is to say where no infinite place is ramified. Relying on some generalisation of Hasse–Schilling–Maaß Norm Theorem, they prove that the Euclidean property of K implies the Euclidean property of any totally indefinite Euclidean quaternion field over K. Conversely, they provide the complete list of norm-Euclidean and totally indefinite quaternion fields over an imaginary quadratic number field. In particular, the article exhibits a totally indefinite and norm-Euclidean quaternion field over a non-Euclidean number field. This provides an answer to a question by Eichler. The proofs are both theoretic and algorithmic. The article has been published in Acta Arithmetica.

Deciding whether an ideal of a number field is principal and finding a generator is a fundamental problem with many applications in computational number theory. In the article [25] gives a an algorithm for indefinite quaternion algebras by reducing the decision problem to that in the underlying number field. It also gives an heuristically subexponential algorithm for finding a generator.

5.5. Complex multiplication and modularity

Participants: Jean-Marc Couveignes, Andreas Enge, Nicolas Mascot, Enea Milio, Aurel Page, Damien Robert.

A. Enge and E. Thomé describe in [20] a quasi-linear algorithm for computing Igusa class polynomials of Jacobians of genus 2 curves via complex floating-point approximations of their roots. After providing an explicit treatment of the computations in quartic CM fields and their Galois closures, they pursue an approach due to Dupont for evaluating θ-constants in quasi-linear time using Newton iterations on the Borchardt mean. They report on experiments with the implementation CMH and present an example with class number 20016.

In [34] E. Milio explains how to generalise the work of Régis Dupont for computing modular polynomials in dimension 2 to invariants derived from theta constants. Modular polynomials have many applications. In particular, they could speed up the CRT-algorithm to compute class fields of degree 4 CM-fields which would lead to faster algorithms to construct cryptographically secure Jacobians of hyperelliptic curves. They are also used to compute graphs of isogenies. This paper presents how to compute modular polynomials and the polynomials computed and then it proves some of their properties.

With F. Morain, A. Enge has determined exhaustively under which conditions “generalised Weber functions”, that is, simple quotients of η functions of not necessarily prime transformation level and not necessarily of genus 1, yield class invariants [19]. The result is a new infinite family of generators for ring class fields, usable to determine complex multiplication curves. They examine in detail which lower powers of the functions are applicable, thus saving a factor of up to 12 in the size of the class polynomials, and describe the cases in which the polynomials have integral rational instead of integral quadratic coefficients.

N. Mascot has continued his work on computing Galois representations attached to Jacobians of modular curves. He has given tables of modular Galois representations in [33] obtained using the algorithm of [39]. He has computed Galois representations modulo primes up to 31 for the first time. In particular, he has computed the representations attached to a newform with non-rational (but of course algebraic) coefficients, which had never been done before. These computations take place in the Jacobians of modular curves of genus up to 26.

5.6. Elliptic curve and Abelian varieties cryptology

Participants: Jean-Marc Couveignes, Andreas Enge, Damien Robert.

In [27] J.-M. Couveignes and T. Ezome show how to efficiently evaluate functions, including Weil functions and canonical Theta functions, on Jacobian varieties and their quotients. They deduce a quasi-optimal algorithm to compute (l, l) isogenies between Jacobians of genus two curves, using a compact representation and differential characterisation of isogenies in this context. This work has been submitted to the LMS Journal of Computation and Mathematics.

The paper [18] by J.-M. Couveignes and R. Lercier describing the problem of parameterisations by radicals of low genus algebraic curves has been accepted in *Advances in mathematics of communications*.

In [31] D. Lubicz and D. Robert explain how to improve the arithmetic of Abelian and Kummer varieties. The speed of the arithmetic is a crucial factor in the performance of abelian varieties based cryptosystem. Depending on the cryptographic application, the speed record holder are elliptic curves (in the Edwards model) or the Kummer surface of an hyperelliptic curves of genus 2 (in the level 2 theta model). One drawback of the Kummer surface is that only scalar multiplications are available, which may be a drawback in certain cryptographic protocols. The previous known models to work on the Jacobian rather than the Kummer surface (Mumford coordinates or theta model of level 4) are too slow and not competitive with Elliptic Curves. This paper explains how to use geometric properties (like projective normality) to speed up the arithmetic. In particular it introduces a novel addition algorithm on Kummer varieties (compatible additions), and use it to enhance multi-exponentiations in Kummer varieties and to obtain new models of abelian surfaces where the scalar multiplication is as fast as on the Kummer surface.

In [32] (which has been accepted at LMS Journal of Computation and Mathematics), D. Lubicz and D. Robert explain how to compute isogenies between abelian varieties given algebraic equation of the kernel. The previous algorithms to compute isogenies between abelian varieties needed the coordinates of generators of the kernel. One drawback was that even if the kernel is rational, these generators may live in extension of large degree, especially for Abelian varieties defined over a number field rather than a finite field. This paper combines the use of formal coordinates together with a normalisation alongs linear subspaces of the kernel rather than the whole kernel to derive an algorithm which is quasi-optimal if the degree of the isogeny is ℓ^g, for ℓ congruent to 1 modulo 4.

This article expands the article [17] by D. Cosset and D. Robert about the computation of (ℓ, ℓ)-isogenies in dimension 2 which has been published in Mathematics of Computation.

5.7. Pairings

Participants: Andreas Enge, Damien Robert.

The article [22] by D. Lubicz and D. Robert explaining how to compute optimal pairings on abelian varieties described by their theta models has been accepted for publication at Journal of Symbolic Computation.

In [24], A. Enge and J. Milan report on the APIP implementation of cryptographic pairings on elliptic curves in PARI/GP. For security levels equivalent to the different AES flavours, they exhibit suitable curves in parametric families and show that optimal ate and twisted ate pairings exist and can be efficiently evaluated. They provide a correct description of Miller's algorithm for signed binary expansions such as the NAF and extend a recent variant due to Boxall et al. to addition-subtraction chains. They analyse and compare several algorithms proposed in the literature for the final exponentiation. Finally, they give recommendations on which curve and pairing to choose at each security level.

BACCHUS Team

5. New Results

5.1. Penalisation methods using unstructured meshes

Participants: Heloise Beaugendre [Corresponding member], Cécile Dobrzynski, Leo Nouveau, Quentin Viville.

In Computational Fluid Dynamics the interest on embedded boundary methods for Navier-Stokes equations increases because they simplify the meshing issue, the simulation of multi-physics flows and the coupling of fluid-solid interactions in situation of large motions or deformations. Nevertheless an accurate treatment of the wall boundary conditions remains an issue of these methods. In this work we develop an immersed boundary method for unstructured meshes based on a penalization technique and we use mesh adaption to improve the accuracy of the method close to the boundary. The idea is to combine the strength of mesh adaptation, that is to provide an accurate flow description especially when dealing with wall boundary conditions, to the simplicity of embedded grids techniques, that is to simplify the meshing issue and the wall boundary treatment when combined with a penalization term to enforce boundary conditions. The bodies are described using a level-set method and are embedded in an unstructured grid. Once a first numerical solution is computed mesh adaptation based on two criteria the level-set and the quality of the solution is performed. The full paper has been published in the Journal of Computational Physics in January 2014.
External contributors. This work has benefitted from the collaboration with the University of Zurich, and in particular with R. Abgrall.

5.2. Mesh adaptation by continuous deformation

Participants: Luca Arpaia, Mario Ricchiuto [Corresponding member].

As discussed in section 3.3 Meshes and scalable discrete data structures an accurate resolution of time dependent flows requires a dynamic mesh adaptation procedure which is quite complex and costly, especially when combined with parallel distributed memory implementations. To alleviate this cost, and still allow mesh adaptation for time dependent problems we have started to look into adaptation techniques which do not involve any re-meshing. In particular, we have studied methods based on continuous mesh deformation. These methods require, at each time step, the solution of a PDE for the mesh as well as for the flow variables. This year we have settled several fundamental questions related to the basic formulation of the method, and its coupling with either implicit or explicit time discretisation methods of the flow variables. Initial applications to free surface flows have been considered showing the generality and potential of our results [39].

5.3. Non-hydrostatic modelling of free surface flows

Participants: Stevan Bellec, Mathieu Colin [Corresponding member], Andrea Filippini, Maria Kazolea, Mario Ricchiuto.

This year we have made a lot of progress in the understanding of the properties of Boussinesq-type models for near shore applications. In particular, we have performed a systematic analysis of the nonlinear behaviour of these models in the surf-zone, and in particular of their shoaling properties. These properties influence fundamentally the wave breaking process, and thus the impact of the wave on coastal structures. We have clearly identified two families of physical behaviours, associated to a similar formal structure of the equations. This result has been presented in [30], [45], and the full study is currently in revision on the Coastal Engineering journal.

In parallel, we have continued the study of the implementation of wave breaking models, comparing several physical criteria for the detection of the beginning and end of the breaking process. So far, we have only tested the so-called hybrid approach in which the hyperbolic Shallow Water equations are used in breaking regions, and the energy dissipation of breaking waves is modelled by the dissipation of mathematical entropy in shock waves. The work performed complements the initial study performed by M. Kazolea in her PhD and also proposes new physical detection criteria [28], [44] (a full paper is in preparation).

Furthermore, we have began a systematic study on the existence of particular solutions (such as solitary waves for example) to the different Boussinesq-type models in view of having efficient materials to determinate the efficiency of our numerical schemes and to perform preliminary simulations.

The last important theoretical brick we added this year is the study of fully discrete asymptotic models, obtained by pre-discretizing the two-dimensional incompressible free surface Euler equations with a finite element method, and then by performing an asymptotic development (in terms of the classical nonlinearity and dispersion parameters). We have thus obtained a discrete model which, although consistent with a known continuous Boussinesq system, represents a surprisingly improved discret eversion of these equations, hardly obtainable by classical discretisation choices.

Besides the modelling effort, we have also started woking on real applications. In particular, we have worked on case studies involving harbour dynamics and river hydraulics. In the first case, M. Kazolea has performed a systematic study of the contribution of harbour resonance in the excitation of the Venetian harbor basin of Chania, during typical winder storms. Concerning river hydraulics, we have performed a parametric study of the appearance of tidal bores in estuaries, with parameters given by the tide non-linearity (amplitude), and the friction in the river. Both works will be presented at the next world congress of the International Association for Hydro-Environment Engineering.
External contributors. This work has benefitted from the collaboration with the EPOC lab in Bordeaux, and in particular with P. Bonneton.

5.4. Two-phase flow numerical simulation with real-gas effects and occurrence of rarefaction shock waves

Participants: Maria Giovanna Rodio, Pietro Marco Congedo [Corresponding member].

A discrete equation method (DEM) for the simulation of compressible multiphase flows including real-gas effects has been developed. A reduced five equation model is obtained starting from the semi-discrete numerical approximation of the two-phase model. A simple procedure is then proposed for using a more complex equation of state, thus improving the quality of the numerical prediction. Classical test-cases well-known in literature are performed featuring a strong importance of thermodynamic complexity for a good prediction of temperature evolution. Finally, a computational study on the occurrence of rarefaction shock waves (RSW) in a two-phase shock tube is presented, with dense vapors of complex organic fluids. Since previous studies have shown that a RSW is relatively weak in a single-phase (vapor) configuration, its occurrence and intensity are investigated considering the influence of the initial volume fraction, initial conditions and the thermodynamic model [11]. A transition modelling has been also introduced for considering heat and mass transfer terms. In this way, metastable states have been simulated in cavitating flows Finally, a semi-intrusive stochastic technique has been formulated for taking into account uncertainties in the simulation of metastable states.
External contributors. This work has benefitted from the collaboration with the University of Zurich, and in particular with R. Abgrall.

5.5. Formulation of stochastic methods for CFD

Participants: Gianluca Geraci, Kunkun Tang, Francesca Fusi, Pietro Marco Congedo [Corresponding member].

A novel adaptive strategy for stochastic problems has been developed, inspired from the classical Harten's framework. The proposed algorithm allows building, in a very general manner, stochastic numerical schemes starting from a whatever type of deterministic schemes and handling a large class of problems, from unsteady to discontinuous solutions. Its formulations permits to recover the same results concerning the interpolation theory of the classical multiresolution approach, but with an extension to uncertainty quantification problems. The present strategy permits to build numerical scheme with a higher accuracy with respect to other classical uncertainty quantification techniques, but with a strong reduction of the numerical cost and memory requirements. Moreover, the flexibility of the proposed approach allows to employ any kind of probability density function, even discontinuous and time varying, without introducing further complications in the algorithm. The advantages of the present strategy are demonstrated by performing several numerical problems where different forms of uncertainty distributions are taken into account, such as discontinuous and unsteady custom-defined probability density functions. In addition to algebraic and ordinary differential equations, numerical results for the challenging 1D Kraichnan–Orszag are reported in terms of accuracy and convergence. Finally, a two degree-of-freedom aeroelastic model for a subsonic case is presented. Though quite simple, the model allows recovering some physical key aspect, on the fluid/structure interaction, thanks to the quasi-steady aerodynamic approximation employed. The injection of an uncertainty is chosen in order to obtain a complete parameterization of the mass matrix. All the numerical results are compared with respect to classical Monte Carlo solution and with a non-intrusive Polynomial Chaos method [3].

Moreover, in [15], an anchored ANOVA method is proposed to decompose statistical moments. Compared to standard ANOVA with mutually orthogonal components, anchored ANOVA, with arbitrary anchor point, loses orthogonality if employing the same measure. However, an advantage consists in the considerably reduced number of deterministic solver's computations, which renders uncertainty quantification of real engineering problems much easier. Different from existing methods, covariance decomposition of output variance is used in this paper to take account of interactions between non-orthogonal components, yielding an exact variance expansion, and thus, with a suitable numerical integration method, provides a strategy that converges. This convergence is verified by studying academic tests. In particular, sensitivity problem of existing method to anchor point is analyzed via Ishigami case, and we point out covariance decomposition survives from it. Covariance-based sensitivity indices (SI) are also used, compared to variance-based SI. Furthermore, we emphasize covariance decomposition can be generalized in a straightforward way to decompose high order moments. For academic problems, results show the method converges to exact solution regarding both skewness and kurtosis. Finally, the proposed method is applied on a realistic case, i.e. estimating chemical reactions uncertainties in a hypersonic flow around a space vehicle during an atmospheric reentry.

External contributors. This work has benefitted from the collaboration with the University of Zurich, and in particular with R. Abgrall.

CAGIRE Team

6. New Results

6.1. DNS of a Taylor Green vortex

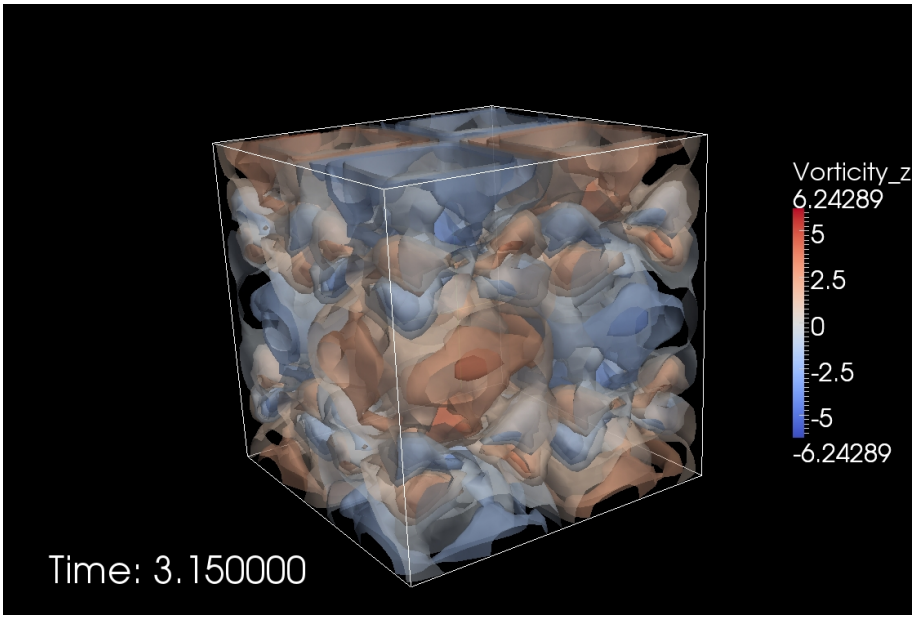

Figure 3. AeroSol simulation of a Taylor-Green vortex: snapshot of one component of the vorticity.

In 2014, we finished the validation of Navier-Stokes discretization with the discontinuous Galerkin method in the Aerosol library. The result of Figure 3 is a first validation in turbulence conditions. The Taylor-Green vortex case is part of the C3 (i.e. "difficult") test cases of the high order CFD workshop, see https://www.grc.nasa.gov/hiocfd/.

6.2. Low Mach number flows simulations issues

Our activity for developing schemes suitable for the simulation of low Mach number flows considers the two main techniques developed initially for dealing with either zero Mach number flows (pressure-velocity coupling) or compressible flows (density based approach). For both approaches, we concentrated this year on the specific difficulties related to unsteady flows simulations. For the methodology adressing the pressure-velocity coupling with a low-order discretization technique, we introduced an inertia term in the AUSM+ -up scheme. The resulting scheme, called AUSM-IT (IT for Inertia Term), was designed as an extension of the AUSM+ -up scheme allowing for full Mach number range calculations of unsteady flows including acoustic features. In line with the continuous asymptotic analysis, the AUSM-IT scheme satisfies the conservation of the discrete linear acoustic energy at first order in the low Mach number limit. Its capability to properly handle low Mach number unsteady flows, that may include acoustic waves or discontinuities was numerically illustrated [7].

As far as density based approach are concerned, an analysis of explicit RKDG schemes have been performed for unstationary acoustic waves propagating in a low Mach number flow. Classical cures of the unaccuracy of upwind schemes at low Mach number consist in using centered flux on the pressure. By a two scale asymptotic expansion of the scheme, we proved that this cure is a dead end for resolving unstationary acoustic waves, because it leads to a non dissipative scheme for the wave equations. We developed a dissipative term that can both stabilize the stationary incompressible equations, and the system of acoustic waves. The results with this new type of scheme have been presented in [8].

6.3. Improving the flexibility of turbulence models for industrial applications

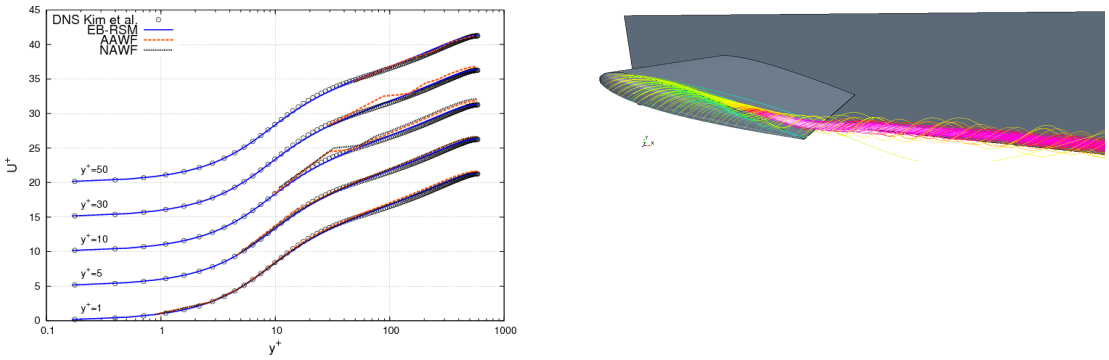

Figure 4. Left: Computation (Code_Saturne) of turbulent channel flow at 3 Reynolds numbers. Comparison with reference DNS of the results given by the EB-RSM integrated down to the wall (ItW, fine mesh) and the EB-RSM with analytical adaptive wall function (AAWF, 3 meshes). Right: EB-RSM computation (STARCCM+ code) of the wing-tip vortex generated by the flow around a NACA 0012 at 10 deg incidence. Visualisation of the streamlines colored with the streamwise vorticity.

In collaboration with industrial partners (EDF and CD-Adapco) developing CFD codes (code_Saturne and STARCCM+, respectively), we are working on the flexibility and robustness of the EB-RSM, an advanced Reynolds-stress turbulence model. Indeed, the two main problems that slow down the spreading of the use of such low-Reynolds number models (i.e., integrating the equations down to solid boundaries) in the industry are the impossiblity to control the near-wall mesh quality in the whole domain of a complex industrial application and the occurence of numerical instabilities due to spurious relaminarizations in some configurations.

In order to address the first issue, we are working, in particular in the frame of the PhD thesis of J.-F. Wald, on the development of adaptive wall functions, i.e., non-homogeneous Dirichlet boundary conditions for the turbulent variables dependant on the size of the cell adjacent to the wall. These wall functions are based on the physical properties of turbulence in the different layers of the near-wall region (asymptotic behaviour in the viscous sublayer and log law in the equilibrium layer), such a way that the flow is correctly reproduced whatever the near-wall refinement of the mesh. Fig. 4 (left) shows that the reproduction of the mean velocity profile in turbulent channel flows obtained using a typical, industrial mesh ($y^+ = 50$) remains very close to the grid-converged solution.

The second issue, the numerical instabilities due to local, spurious relaminarization of the model, can be addressed by investigating the solutions of the dynamical system formed by the model equations in homogeneous situations. Equilibrium solution are intersections of the nullclines (the locus of steady solutions for individual equations) and the stability properties of these fixed points can be visualized using trajectories in the phase space. By investigating the dependance of these stability properties on the parameters of the model, it is possible to eliminate undesired stable fixed points and thus to avoid the appearance of spurious

laminarization. Fig. 4 (right) shows the fully turbulent solution obtained with the modified model in a case where the original model exhibited a severe, unphysical relaminarization of the wing-tip vortex.

6.4. Assessment of the discontinuous Galerkin methods on curved meshes

The internship of Hamza Belkhayat-Zougari was concerned with the handling of high order curved meshes in the Aerosol library. During his internship, we developed new analytical solutions of the Laplace and of the Navier-Stokes equations on curved domains for emphasizing the limitation at second order of high order methods on straight meshes, and for assessing the right order on high order meshes. Example of order obtained on straight and curved meshes can be found on Figure 5 .

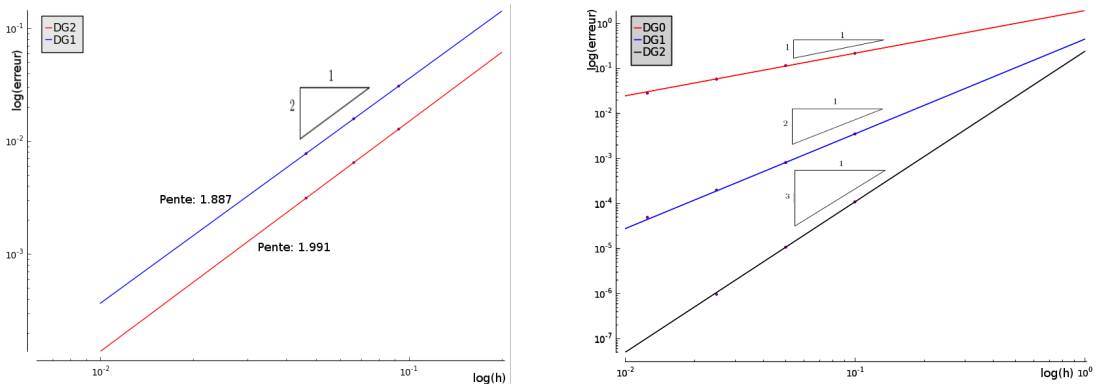

Figure 5. Convergence on a ring for the Laplace equation. Left: high order method on a straight mesh is limited to two. Right: third order accuracy can be recovered by using a second order mesh.

CQFD Project-Team

6. New Results

6.1. Highlights of the Year

Creation of the Associate Team Inria: CDSS (2014-2016) with the University of Sao Paulo, Brasil.

6.2. Approximate Kalman–Bucy filter for continuous-time semi-Markov jump linear systems

Participants: Benoîte de Saporta, Eduardo Costa.

We propose a new numerical approximation of the Kalman–Bucy filter for semi-Markov jump linear systems. This approximation is based on the selection of typical trajectories of the driving semi-Markov chain of the process by using an optimal quantization technique. The main advantage of this approach is that it makes pre-computations possible. We derive a Lipschitz property for the solution of the Riccati equation and a general result on the convergence of perturbed solutions of semi-Markov switching Riccati equations when the perturbation comes from the driving semi-Markov chain. Based on these results, we prove the convergence of our approximation scheme in a general infinite countable state space framework and derive an error bound in terms of the quantization error and time discretization step. We employ the proposed filter in a magnetic levitation example with markovian failures and compare its performance with both the Kalman–Bucy filter and the Markovian linear minimum mean squares estimator. This work was presented at the international conference [37] and is submitted to an international journal [50].

6.3. Modeling and optimization of a launcher integration process

Participants: Benoîte de Saporta, François Dufour, Christophe Nivot.

We are interested in the optimization of a launcher integration process. It comprises several steps from the production of the subassemblies to the final launch. The four subassemblies go through various types of operations such as preparation, integration, control and storage. These operations are split up into three workshops. Due to possible breakdowns or staff issues, the time spent in each workshop is supposed random. So is the time needed to deliver the subassemblies, for similar reasons including e.g. shipping delays. We also have to deal with constraints related to the architecture of the assembly process itself. Indeed, we have to take into account waiting policies between workshops. The workshops may work in parallel but can be blocked if their output is not transferred to the next workshop in line. Storage capacity of output products is limited.

Our goal is finding the best rates of delivery of the subassemblies, the best choice of architecture (regarding stock capacities) and the best times when to stop and restart the workshops to be able to carry out twelve launches a year according to a predetermined schedule at minimal cost. To solve this problem, we choose a mathematical model particularly suitable for optimization with randomness: Markov decision processes (MDPs).

We have implemented a numerical simulator of the process based on the MDP model. It provides the fullest information possible on the process at any time. The simulator has first been validated with deterministic histories. Random histories have then been run with exponentially distributed delivery times for the subassemblies and several families of random laws for the time spent in each workshop. Using Monte Carlo simulations, we obtain the distribution of the launch times. Preliminary optimization results allow choosing stock capacities and delivery rates that satisfy the launch schedule. Work is still in progress concerning cost minimization. It was presented at Airbus internal PhD seminar in November 2014.

6.4. Numerical approximation for optimal stopping of MDP under partial observation

Participants: Benoîte de Saporta, François Dufour, Christophe Nivot.

We consider the optimal stopping problem for a continuous finite-dimensional state space Markov chain under partial observation. Our aim is to build a numerical approximation of the value function. To do so, we first translate the problem into the Partially Observed Markov Decision Process (POMDP) framework. Then, we define the equivalent fully observed Markov Decision Process (MDP) on an infinite dimensional state space. Finally, we proposed a discretization scheme based on the discretization of an underlying measure to obtain a finite dimensional problem and a discretization of the resulting state space to obtain a fully discrete model that is numerically tractable. We prove the convergence of the approximation procedure. This work is still in progress and was presented at the workshop [31]

6.5. Classification of EEG signals by evolutionary algorithm

Participants: Marie Chavent, Pierrick Legrand, Leonardo Trujillo.

The goal of this work is to predict the state of alertness of an individual by analyzing the brain activity through electroencephalographic data (EEG) captured with 58 electrodes. Alertness is characterized here as a binary variable that can be in a "normal" or "relaxed" state. We collected data from 44 subjects before and after a relaxation practice, giving a total of 88 records. After a pre-processing step and data validation, we analyzed each record and discriminate the alertness states using our proposed "slope criterion". Afterwards, several common methods for supervised classification (k nearest neighbors, decision trees (CART), random forests, PLS and discriminant sparse PLS) were applied as predictors for the state of alertness of each subject. The proposed "slope criterion" was further refined using a genetic algorithm to select the most important EEG electrodes in terms of classification accuracy. Results show that the proposed strategy derives accurate predictive models of alertness.

This work has been published in a book chapter [45].

6.6. Probabilistic low-rank matrix completion with adaptive spectral regularization algorithms

Participants: Marie Chavent, Adrien Todeschini.

We propose a novel class of algorithms for low rank matrix completion. Our approach builds on novel penalty functions on the singular values of the low rank matrix. By exploiting a mixture model representation of this penalty, we show that a suitably chosen set of latent variables enables to derive an EM algorithm to obtain a Maximum A Posteriori estimate of the completed low rank matrix. The resulting algorithm is an iterative soft-thresholded algorithm which iteratively adapts the shrinkage coefficients associated to the singular values.

This work is in collaboration with Francois Caron from University of Oxford. It has been presented in the national conference of the French Statistical Society of Statistics [41]

6.7. Variable selection to construct indicators of quality of life for data structured in groups

Participants: Marie Chavent, Amaury Labenne, Jérôme Saracco.

The analysis and measurement of quality of life may be made via two complementary approaches. The first one, based on survey of individuals, concerns the analysis of levels of life satisfaction. We focus here on the second one, based on national data, which analyses living conditions of people. The aim is to create composite indices of living conditions. According to authors, the components of quality of life are related to different themes (groups of variables): Family conditions", Employment", Housing",... For this purpose, dimension reduction methods are particularly suitable. Multiple Factor Analysis (MFA) is a method designed to handle data structured into groups of quantitative variables. In our study, each theme is composed of a group of quantitative and/or categorical variables. Since our data are naturally structured in groups of variables, we develop an extension of MFA for mixed data type, called MFAmix. Thus the principal components from MFAmix are our composite indices for measuring quality of life. However, the creation of these indices raises two questions. How many principal components keep to create indices? How select a limited number of variables to get similar indices for easier interpretation? We propose answers to these questions in this communication.

This work is in collaboration with Vanessa Kuentz from Irstea. It has been presented in the french meeting of the R users (Rencontres R) [40] and in the international conference COMPSTAT 2014 [36].

6.8. Efficiency of simulation in monotone hyper-stable queueing networks

Participants: Jonatha Anselmi, Bruno Gaujal.

We consider Jackson queueing networks with finite buffer constraints (JQN) and analyze the efficiency of sampling from their stationary distribution. In the context of exact sampling, the monotonicity structure of JQNs ensures that such efficiency is of the order of the 'coupling time' (or meeting time) of two extremal sample paths. In the context of approximate sampling, it is given by the 'mixing time'. Under a condition on the drift of the stochastic process underlying a JQN, which we call hyper-stability, in our main result we show that the coupling time is polynomial in both the number of queues and buffer sizes. Then, we use this result to show that the mixing time of JQNs behaves similarly up to a given precision threshold. Our proof relies on a recursive formula relating the coupling times of trajectories that start from network states having 'distance one', and it can be used to analyze the coupling and mixing times of other Markovian networks, provided that they are monotone. An illustrative example is shown in the context of JQNs with blocking mechanisms. This work has been published in an international journal; see [11].

6.9. Control of parallel non-observable queues: asymptotic equivalence and optimality of periodic policies

Participants: Jonatha Anselmi, Bruno Gaujal, Tommaso Nesti.

We consider a queueing system composed of a dispatcher that routes deterministically jobs to a set of non-observable queues working in parallel. In this setting, the fundamental problem is which policy should the dispatcher implement to minimize the stationary mean waiting time of the incoming jobs. We present a structural property that holds in the classic scaling of the system where the network demand (arrival rate of jobs) grows proportionally with the number of queues. Assume that each queue of type r is replicated k times and consider the set of policies that are periodic with period $k\sum_r p_r$ and such that exactly p_r jobs are sent in a period to each queue of type r. When $k \to \infty$, our main result shows that all the policies in this set are equivalent, in the sense that they yield the same mean stationary waiting time, and optimal, in the sense that no other policy having the same aggregate arrival rate to all queues of a given type can do better in minimizing the stationary mean waiting time. This property holds in a strong probabilistic sense. Furthermore, the limiting mean waiting time achieved by our policies is a convex function of the arrival rate in each queue, which facilitates the development of a further optimization aimed at solving the fundamental problem above for large systems. This work has been accepted for publication in the international journal "Stochastic Systems", the flagship journal of the INFORMS Applied Probability Society; see [46].

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

Participants: Jonatha Anselmi, Danilo Ardagna, Jonh C.s. Lui, Adam Wierman, Yunjian Xu, Zichao Yang.

This work 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. This work has been submitted to an international journal. A preliminary has been published in [10].

6.11. Generalized Nash Equilibria for Platform-as-a-Service Clouds

Participants: Jonatha Anselmi, Danilo Ardagna, Mauro Passacantando.

Cloud computing is an emerging technology that allows to access computing resources on a pay-per-use basis. The main challenges in this area are the cient performance management and the energy costs minimization. In this work we model the service provisioning problem of Cloud Platform-as-a-Service systems as a Generalized Nash Equilibrium Problem and show that a potential function for the game exists. Moreover, we prove that the social optimum problem is convex and we derive some properties of social optima from the corresponding Karush-Kuhn-Tucker system. Next, we propose a distributed solution algorithm based on the best response dynamics and we prove its convergence to generalized Nash equilibria. Finally, we numerically evaluate equilibria in terms of their efficiency with respect to the social optimum of the Cloud by varying our algorithm initial solution. Numerical results show that our algorithm is scalable and very efficient and thus can be adopted for the run-time management of very large scale systems. This work has been published in an international journal; see [12].

6.12. Stochastic approximations of constrained discounted Markov decision processes

Participants: Francois Dufour, Tomas Prieto-Rumeau.

We consider a discrete-time constrained Markov decision process under the discounted cost optimality criterion. The state and action spaces are assumed to be Borel spaces, while the cost and constraint functions might be unbounded. We are interested in approximating numerically the optimal discounted constrained cost. To this end, we suppose that the transition kernel of the Markov decision process is absolutely continuous with respect to some probability measure μ. Then, by solving the linear programming formulation of a constrained control problem related to the empirical probability measure μ_n of μ, we obtain the corresponding approximation of the optimal constrained cost. We derive a concentration inequality which gives bounds on the probability that the estimation error is larger than some given constant. This bound is shown to decrease exponentially in n. Our theoretical results are illustrated with a numerical application based on a stochastic version of the Beverton-Holt population model. This work has been published in Journal of Mathematical Analysis and applications: [27].

6.13. Non-Parametric Estimation of the Conditional Distribution of the Interjumping Times for Piecewise-Deterministic Markov Processes

Participants: Romain Azais, Francois Dufour, Anne Gegout-Petit.

We study a nonparametric method for estimating the conditional density associated to the jump rate of a piecewise-deterministic Markov process. In our framework, the estimation needs only one observation of the process within a long time interval. Our method relies on a generalization of Aalen?s multiplicative intensity model. We prove the uniform consistency of our estimator, under some reasonable assumptions related to the primitive characteristics of the process. A simulation study illustrates the behavior of our estimator. This work has been published in Scandinavian Journal of Statistics: [15].

6.14. Approximation of average cost Markov decision processes using empirical distributions and concentration inequalities

Participants: Francois Dufour, Tomas Prieto-Rumeau.

We consider a discrete-time Markov decision process with Borel state and action spaces, and possibly unbounded cost function. We assume that the Markov transition kernel is absolutely continuous with respect to some probability measure μ. By replacing this probability measure with its empirical distribution μ_n for a sample of size n, we obtain a finite state space control problem, which is used to provide an approximation of the optimal value and an optimal policy of the original control model. We impose Lipschitz continuity properties on the control model and its associated density functions. We measure the accuracy of the approximation of the optimal value and an optimal policy by means of a non-asymptotic concentration inequality based on the 1–Wasserstein distance between μ and μ_n. Obtaining numerically the solution of the approximating control model is discussed and an application to an inventory management problem is presented. This work has been published in Stochastics An International Journal of Probability and Stochastic Processes: [26].

6.15. Piecewise Deterministic Markov Processes based approach applied to an offshore oil production system

Participants: Huilong Zhang, Fares Innal, François Dufour, Yves Dutuit.

This work is keeping with the topic of two papers which treated dynamic reliability problems and were presented in previous conferences. Its aim is to confirm the potentialities of a method which combines the high modeling ability of the piecewise deterministic processes and the great computing power inherent to the Monte Carlo simulation. This method is now applied to a simplified but realistic offshore oil production system which is a hybrid system combining continuous-time and discrete-time dynamics. The results thus obtained have been compared with those given by an ad hoc Petri net model for comparison and validation purposes. This work has been published in an international journal; see [29].

6.16. Optimal Trajectories for Underwater Vehicles by Quantization and Stochastic control

Participants: Huilong Zhang, Benoîte de Saporta, François Dufour, Dann Laneuville, Adrien Nègre.

We propose in this paper a numerical method which computes the trajectory of a vehicle subject to some mission objectives. The method is applied to a submarine whose goal is to best detect one or several targets (we consider signal attenuation due to acoustic propagation) or/and to minimize its own detection range perceived by the other targets. Our approach is based on dynamic programming of a finite horizon Markov decision process. The position and the velocity of the targets are supposed to be known only up to a random estimation error, as a Kalman type filter is used to estimate these quantities from the measurements given by the on board sonar. We also take into account the information on the environment through a sound propagation code. A quantization method is applied to fully discretize the problem and solve it numerically. This work is still in progress and was presented at the international conference [39].

6.17. Multi-Objective Design and Maintenance Optimization of the Heated Hold-Up Tank Modeled by Piecewise Deterministic Markov Processes

Participants: Huilong Zhang, Yanfu Li.

We propose a numerical method for the optimal design and maintenance for the heated hold-up tank system. A multi-objective problem is framed to consider simultaneously the objectives of maximizing the operation profit and maximizing the reliability. The system consists of a tank containing a fluid whose level is controlled by three components: two inlet pumps and one outlet valve. A thermal power source heats up the fluid. The failure rates of the components depend on the temperature, the position of the three components monitors the liquid level in the tank and the liquid level determines the temperature. We model the system by a piecewise deterministic Markov process. To find the solution of the optimal maintenance interval, the non-dominated sorting genetic algorithm-II (NSGA-II) is used. This work is still in progress and was presented at the international conference [42].

6.18. Conditional quantile estimation through optimal quantization

Participants: Isabelle Charlier, Jérôme Saracco.

We use quantization to construct a nonparametric estimator of conditional quantiles of a scalar response Y given a d-dimensional vector of covariates X. First we focus on the population level and show how optimal quantization of X, which consists in discretizing X by projecting it on an appropriate grid of N points, allows to approximate conditional quantiles of Y given X. We show that this approximation is arbitrarily good as N goes to infinity and provide a rate of convergence for the approximation error. Then we turn to the sample case and define an estimator of conditional quantiles based on quantization ideas. We prove that this estimator is consistent for its fixed-N population counterpart. The results are illustrated on a numerical example. This work is in collaboration with Davy Paindaveine from Université Libre de Bruxelles. It has been presented in the national conference of the French Statistical Society of Statistics [35] and in the international conference on computational statistics [34].

6.19. Conditional quantile estimator based on optimal quantization: from theory to practice

Participants: Isabelle Charlier, Jérôme Saracco.

[21] recently introduced a promising nonparametric estimator of conditional quantiles based on optimal quantization, but almost exclusively focused on its theoretical properties. We now discuss its practical implementation (by proposing in particular a method to properly select the corresponding smoothing parameter, namely the number of quantizers) and (ii) we investigate how its finite-sample performances compare with those or classical kernel of nearest-neighbor competitors. Monte Carlo studies show that the quantization-based estimator competes well in all cases (in terms of mean squared errors) and tends to dominate its competitors as soon as the covariate is not uniformly distributed over its support. We also apply our approach to a real data set. While most of the paper focuses on the case of a univariate covariate, we also briefly discuss the multivariate case and provide an illustration for bivariate regressors. This work is in collaboration with Davy Paindaveine from Université Libre de Bruxelles. It has been presented in the national conference of the French Statistical Society of Statistics [35] and in the international conference on computational statistics [34].

6.20. QuantifQuantile : an R package for performing quantile regression trough optimal quantization

Participants: Isabelle Charlier, Jérôme Saracco.

Quantile regression allows to assess the impact of some covariate X on a response Y . An important application is the construction of reference curves and conditional prediction intervals for Y . Recently, [21] developed a new nonparametric quantile regression method based on the concept of optimal quantization. We now describe an R package, called QuantifQuantile, that allows to perform quantization-based quantile regression. We describe the various functions of the package and provide examples. This work is in collaboration with Davy Paindaveine from Université Libre de Bruxelles. It has been presented in the national conference on the R software [43].

6.21. Transcriptome profile analysis reveals specific signatures of pollutants in Atlantic eels

Participant: Jérôme Saracco.

Identifying specific effects of contaminants in a multi-stress field context remain a challenge in ecotoxicology. In this context, "omics" technologies, by allowing the simultaneous measurement of numerous biological endpoints, could help unravel the in situ toxicity of contaminants. In this study, wild Atlantic eels were sampled in 8 sites presenting a broad contamination gradient in France and Canada. The global hepatic transcriptome of animals was determined by RNA-Seq. In parallel, the contamination level of fish to 8 metals and 25 organic pollutants was determined. Factor analysis for multiple testing was used to identify genes that are most likely to be related to a single factor. Among the variables analyzed, arsenic (As), cadmium (Cd), lindane (γ-HCH) and the hepato-somatic index (HSI) were found to be the main factors affecting eel's transcriptome. Genes associated with As exposure were involved in the mechanisms that have been described during As vasculotoxicity in mammals. Genes correlated with Cd were involved in cell cycle and energy metabolism. For γ-HCH, genes were involved in lipolysis and cell growth. Genes associated with HSI were involved in protein, lipid and iron metabolisms. Our study proposes specific gene signatures of pollutants and their impacts in fish exposed to multi-stress conditions.

This work is in collaboration with G. Durrieu from Vannes University and R. Coudret. It will be published in Ecotoxicology [17].

6.22. Comparaison of kernel density estimators with assumption on number of modes : application on environmental monitoring data

Participant: Jérôme Saracco.

A data-driven bandwidth choice for a kernel density estimator called critical bandwidth is investigated. This procedure allows the estimation to have as many modes as assumed for the density to estimate. Both Gaussian and uniform kernels are considered. For the Gaussian kernel, asymptotic results are given. For the uniform kernel, an argument against these properties is mentioned. These theoretical results are illustrated with a simulation study that compares the kernel estimators that rely on critical bandwidth with another one that uses a plug-in method to select its bandwidth. An estimator that consists in estimates of density contour clusters and takes assumptions on number of modes into account is also considered. Finally, the methodology is illustrated using environment monitoring data.

This work is in collaboration with G. Durrieu from Vannes University and R. Coudret. It will be published in Communication in Statistics - Simulation and Computation [28].

6.23. A new sliced inverse regression method for multivariate response

Participant: Jérôme Saracco.

A semiparametric regression model of a q-dimensional multivariate response y on a p-dimensional covariate x is considered. A new approach is proposed based on sliced inverse regression (SIR) for estimating the effective dimension reduction (EDR) space without requiring a prespecified parametric model. The convergence at rate square root of n of the estimated EDR space is shown. The choice of the dimension of the EDR space is discussed. Moreover, a way to cluster components of y related to the same EDR space is provided. Thus, the proposed multivariate SIR method can be used properly on each cluster instead of blindly applying it on all components of y. The numerical performances of multivariate SIR are illustrated on a simulation study. Applications to a remote sensing dataset and to the Minneapolis elementary schools data are also provided. Although the proposed methodology relies on SIR, it opens the door for new regression approaches with a multivariate response.

This work is in collaboration with S. Girard from Inria MISTIS team and R. Coudret. It is published in CSDA [23].

6.24. An introduction to dimension reduction in nonparametric kernel regression

Participant: Jérôme Saracco.

Nonparametric regression is a powerful tool to estimate nonlinear relations between some predictors and a response variable. However, when the number of predictors is high, nonparametric estimators may suffer from the curse of dimensionality. In this chapter, we show how a dimension reduction method (namely Sliced Inverse Regression) can be combined with nonparametric kernel regression to overcome this drawback. The methods are illustrated both on simulated datasets as well as on an astronomy dataset using the R software [51].

This work is in collaboration with S. Girard from Inria MISTIS team .

6.25. Hidden Markov Model for the detection of a degraded operating mode of optronic equipment

Participant: Jérôme Saracco.

As part of optimizing the reliability, Thales Optronics now includes systems that examine the state of its equipment. The aim of this paper is to use hidden Markov Model to detect as soon as possible a change of state of optronic equipment in order to propose maintenance before failure. For this, we carefully observe the dynamic of a variable called "cool down time" and noted Tmf, which reflects the state of the cooling system. Indeed, the Tmf is an indirect observation of the hidden state of the system. This one is modelled by a Markov chain and the Tmf is a noisy function of it. Thanks to filtering equations, we obtain results on the probability that an appliance is in degraded state at time t, knowing the history of the Tmf until this moment. We have evaluated the numerical behavior of our approach on simulated data. Then we have applied this methodology on our real data and we have checked that the results are consistent with the reality. This method can be implemented in a HUMS (Health and Usage Monitoring System). This simple example of HUMS would allow the Thales Optronics Company to improve its maintenance system. This company will be able to recall appliances which are estimated to be in degraded state and do not control to soon those estimated in stable state.

This work is in collaboration with A. Gegout-Petit from Lorraine University. It is published in Journal de la SFdS [19].

6.26. On the asymptotic behavior of the Nadaraya-Watson estimator associated with the recursive SIR method

Participant: Jérôme Saracco.

We investigate the asymptotic behavior of the Nadaraya-Watson estimator for the estimation of the regression function in a semiparametric regression model. On the one hand, we make use of the recursive version of the sliced inverse regression method for the estimation of the unknown parameter of the model. On the other hand, we implement a recursive Nadaraya-Watson procedure for the estimation of the regression function which takes into account the previous estimation of the parameter of the semiparametric regression model. We establish the almost sure convergence as well as the asymptotic normality for our Nadaraya-Watson estimator. We also illustrate our semiparametric estimation procedure on simulated data.

This work is in collaboration with B. Bercu from Bordeaux University and T.M.N Nguyen. It is published in Statistics [20].

6.27. Evolving Genetic Programming Classifiers with Novelty Search

Participants: Enrique Naredo, Leonardo Trujillo, Pierrick Legrand.

Novelty Search (NS) is a unique approach towards search and optimization, where an explicit objective function is replaced by a measure of solution novelty to provide the selective pressure in an artificial evolutionary system. However, NS has been mostly used in evolutionary robotics, while it's applicability to classic machine learning problems has been mostly unexplored. This work presents a NS-based Genetic Programming (GP) algorithm for supervised classification, with the following noteworthy contributions. It is shown that NS can solve real-world classification tasks, validated over several commonly used benchmarks. These results are made possible by using a domain-specific behavioral descriptor, closely related to the concept of semantics in GP. Moreover, two new variants of the NS algorithm are proposed, Probabilistic NS (PNS) and a variant of Minimum Criterion NS (MCNS). The former models the behavior of each solution as a random vector, eliminating all the NS parameters and reducing the computational overhead of the traditional NS algorithm; the latter uses a standard objective function to constrain the search and bias the process towards high performance solutions. The paper also discusses the effects of NS on an important GP phenomenon, bloat. In particular, results indicate that some variants of the NS approach can have a beneficial effect on the search process by curtailing code growth. See [52].

6.28. Detecting mental states of alertness with genetic algorithm variable selection

Participants: Laurent Vezard, Pierrick Legrand, Marie Chavent, Frédérique Faïta, Léonardo Trujillo.

The objective of the present work is to develop a method that is able to automatically determine mental states of vigilance; i.e., a person's state of alertness. Such a task is relevant to diverse domains, where a person is expected or required to be in a particular state of mind. For instance, pilots and medical staff are expected to be in a highly alert state and the proposed method could help to detect possible deviations from this expected state. This work poses a binary classification problem where the goal is to distinguish between a "relaxed" state and a baseline state ("normal") from the study of electroencephalographic signals (EEG) collected with a small number of electrodes. The EEG of 58 subjects in the two alertness states (116 records) were collected via a cap with 58 electrodes. After a data validation step, 19 subjects were retained for further analysis. A genetic algorithm was used to select a subset of electrodes. Common spatial pattern (CSP) coupled to linear discriminant analysis (LDA) was used to build a decision rule and thus predict the alertness of the subjects. Different subset sizes were investigated and the best compromise between the number of selected electrodes and the quality of the solution was obtained by considering 9 electrodes. Even if the present approach is costly in computation time (GA search), it allows to construct a decision rule that provides an accurate and fast prediction of the alertness state of an unseen individual. See [45], [54].

6.29. A comparison of fitness-case sampling methods for Symbolic Regression

Participants: Yuliana Martinez, Léonardo Trujillo, Enrique Naredo, Pierrick Legrand.

The canonical approach towards fitness evaluation in Genetic Programming (GP), is to use a static training set to determine fitness, based on a cost function (root-mean-squared error) averaged over all cases. However, motivated by different goals, researchers have recently proposed several techniques that focus selective pressure on a subset of fitnesscases at each generation. These approaches can be described as fitnesscase sampling techniques, where the training set is sampled, in someway, to determine fitness. This paper shows a comprehensive evaluation of some sampling methods using benchmark problems and real-world problems. The algorithms considered here are Interleaved Sampling, Random Interleaved Sampling, Lexicase Selection and a new sampling technique is proposed called Keep-Worst Interleaved Sampling (KW-IS). The algorithms are extensively evaluated based on test performance, overfitting and bloat. Results suggest that sampling techniques can improve performance based on testing error, bloat and overfitting compared to standard GP. Some of the best results were achieved by Lexicase Selection and Keep Worse-Interleaved Sampling which obtained good results in overfitting and bloat effect. Results also show that on these problems overfitting correlates strongly with bloating and exhibits a good compromise among the considered performance measures.

6.30. Geometric Semantic Genetic Programming with Local Search

Participants: Emigdio Z. Flores, Léonardo Trujillo, Leonardo Vanneshi, Sara Silva, Pierrick Legrand.

Since its introduction, Geometric Semantic Genetic Programming (GSGP) has aroused the interest of numerous researchers and several studies have demonstrated that GSGP is able to effectively optimize training data by means of small variation steps, that also have the effect of limiting overfitting. In order to speed up the search process, in this paper we propose a system that integrates a local search strategy into GSGP (called GSGP-LS). Furthermore, we present a hybrid approach, that combines GSGP and GSGP-LS, aimed at exploiting both the optimization speed of GSGP-LS and the ability to limit overfitting of GSGP. The experimental results we present, performed on a set of complex real-life applications, show that GSGP-LS achieves the best training fitness while converging very quickly, but severely overfits; GSGP converges very slowly, but is basically not affected by overfitting. The best overall results were achieved with the hybrid approach, allowing the search to converge quickly, while also exhibiting a noteworthy ability to limit overfitting. These results are encouraging, and suggest that future GSGP algorithms should focus on finding the correct balance between the greedy optimization of a local search strategy and the more robust geometric semantic operators.

GEOSTAT Project-Team

6. New Results

6.1. Highlights of the Year

Paper **Spanning the Scales of Granular Materials through Microscopic Force Imaging** by N. Brodu *et al.* accepted in **Nature Communications** (will appear in 2015).
BEST PAPER AWARD :
[36] **IEEE TENSYMP 2014**. B. XU, S. BINCZAK, S. JACQUIR, O. PONT, H. YAHIA.

6.2. Super-resolution for Earth Observation data

Participants: Hussein Yahia, Joël Sudre, Oriol Pont, Véronique Garçon, Dharmendra Singh.

References: [17], [30], [28], [38], [29].
With partners at LEGOS and in the framework of the OPTIC associated team (7.4.1), we are developping the novel super-resolution approaches for Universe Sciences data. New results are obtained for ocean dynamics, partial pressures pCO_2 between the ocean and the atmosphere, and data fusion.

6.3. Fast and Accurate Texture Recognition with Multilayer Convolution and Multifractal Analysis

Participants: Hicham Badri, Hussein Yahia, Khalid Daoudi.

Reference: [25].
A fast and accurate texture recognition system is presented. It consists in extracting locally and globally invariant representations of a given texture image. The mapping from the locally to the globally invariant representation is based on a scale-invariant method via the calculation of singularity exponents. The final descriptor is extracted from the distribution of these exponents and leads to a more accurate descriptor compared to the popular box-counting method. We also propose to use a combination of the generative PCA classifier together with multi-class SVM as well as a synthetic training strategy. Experiments show that the proposed solution outperforms existing methods on three challenging public benchmark datasets, while being computationally efficient.

6.4. Fast Image Edge-Aware Processing

Participants: Hicham Badri, Hussein Yahia, Driss Aboutajdine.

Reference: Article *Fast Edge-Aware Processing via First Order Proximal Approximation* by H. Badri, H. Yahia, D. Aboutajdine, accepted with minor revision in **IEEE Transactions on Visualization & Computer Graphics**, will be in HAL in 2015.
We present a framework for fast edge-aware processing of images and videos. This is an extension of our previous SIGGRAPH Asia 2013 paper. The proposed approach uses non-convex sparsity on the gradients of the latent smooth image to better preserve sharp edges. We develop tools based on first order proximal estimation for fast processing. We also propose fast and efficient numerical solutions based on separable filters estimation, which enables our method to perform fast high-quality smoothing on large-scale images. Extensive experiments show that the proposed method produces high-quality smoothing compared to state-of-the-art methods, while being fast and simple to implement.

6.5. Cardiac arrhythmia induced by mild hypothermia in vitro – a pitchfork bifurcation type process

Participants: Binbin Xu, Oriol Pont.

Reference: [20].
The neurological damage after cardiac arrest constitutes a big challenge of hospital discharge. The mild therapeutic hypothermia (MTH) (34°C - 32°C) has shown its benefit to reduce this type of damage. However, it can have many adverse effects, among which the cardiac-arrhythmia-generation-a-posteriori (CAGP) can represent up to 34%. Our study with a cardiac culture in vitro showed that at 35°C the CAGP can be induced. The process of MTH can be represented by a Pitchfork bifurcation, which could explain the different ratio of arrhythmia among the adverse effects after this therapy. This nonlinear dynamics suggests that a variable speed of cooling / rewarming, especially when passing 35°C, would help to decrease the ratio of post-hypothermia arrhythmia and then improve the hospital output.

6.6. Characterizing the dynamics of cardiac arrhythmia

The dynamics of cardiac arrhythmia is quite complex. Better understanding its mechanism can help to improve the treatment. In vitro cultures of cardiac cells which has similar parameters as cell of human's heart represent valuable tool and model to study this issue.

6.6.1. by Complexity Analysis

Participants: Binbin Xu, Oriol Pont.

References: [36], [39].
Stochastic approaches provide a type of methods to characterize cardiac arrhythmia, aimed at quantifying the statistical properties of the time series. Complexity analysis such as Approximate Entropy (ApEn) and Sample Entropy (SampEn), are particularly useful to analyze time series in electro-cardiology in which the signals are characterized by their high regularity in normal condition in contrast to irregularity in pathological cases. It is shown that ApEn and SampEn can not only serve as a discrimination index, but also provide another parameter which showed doubling phenomenon. It proves in other terms that bifurcation happens in case arrhythmia. See figure 1 .

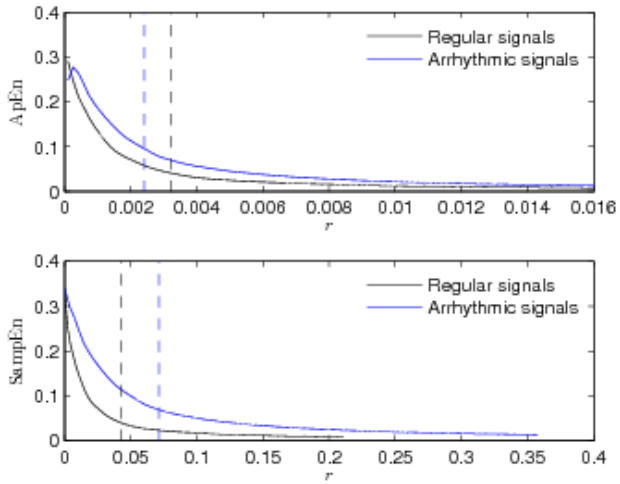

Figure 1. Illustration of ApEn / SampEn analysis for normal and arrhythmic electrical field potential.

6.6.2. by Phase Space Reconstruction

Participants: Oriol Pont, Binbin Xu.

References: [19], [40], [22].
Phase space reconstructions of electrical field potential signals in normal and arrhythmic cases are performed by characterizing the nonlinearity of these signals. The phase space reconstructions highlight attractors, whose dimension reveals that they are strange, depicting a deterministic dynamics of chaotic nature in the in vitro model. The electrical activity of the heart consists of nonlinear interactions emerging as a complex system. Electrocardiographic imaging provides a full spatiotemporal picture of the electric potential on the human epicardium. Rhythm reflects the connection topology of the pacemaker cells driving it. Hence, characterizing the attractors as nonlinear, effective dynamics can capture the key parameters without imposing any particular microscopic model on the empirical signals. A dynamic phase-space reconstruction from an appropriate embedding can be made robust and numerically stable with the methods developed in the team. With these, we have been able to show how both the phase-space descriptors and those of the a priori unrelated singularity analysis are able to highlight the arrhythmogenic areas on cases of atrial fibrillation. See figure 2 .

Figure 2. Illustration of phase space analysis of normal and arrhythmic electrical field potential.

6.7. The origin of the myth FitzHugh-Nagumo model

Participants: Binbin Xu, Oriol Pont.

Reference: [37].
History became legend. Legend became myth. Derived from the pioneer ionic Hodgkin-Huxley model and due to its simplicity and richness from a point view of nonlinear dynamics, the FitzHugh-Nagumo model (FHN) is one of the most successful simplified neuron / cardiac cell model. 60 years later, there exist many variations of this model whose parameters (ε, γ and α) are often used in biased conditions. The related results would be questionable. This study showed that α controls the global dynamics of FHN. $\alpha > 0$, the cell is in refractory mode and does not respond to external stimulation; if $\alpha < 0$, the cell is excitable. ε controls the main morphology of the action potential generated. γ influences barely AP, it showed linear relationship with the period and duration of AP. Though it can be freely chosen for excitable cell, but smaller values are recommended.

6.8. Pathological Speech Analysis

Participants: Khalid Daoudi, Vahid Khanagha, Blaise Bertrac, Safa Mrad, Ashwini Jaya Kumar.

References: [14], [13], [26], [27].

We applied our recent results in nonlinear speech analysis to the filed of pathological speech detection and classification. We presented new insights in the task of normal-vs-pathological voice classification using the widely used Kayelemetrics database. In particular, we showed that hat one single parameter, derived from matching pursuit decomposition of speech, allow perfect discrimination between normal and dysphonic voices of these database. This result raises some important questions on the way this task is generally addressed. Using our GCI detection algorithm, we also proposed new definitions of standard voice perturbation measures (jitter, shimmer...) which lead to significantly higher classification accuracy. Our new measures have the strong advantage to avoid the usual periodicity and linearity assumptions. On the other hand, we started investigating the task of discrimination between Parkinson's and healthy voices. Our phonetic segmentation algorithm has potentially the ability to detect vowel onset and offset regions which have different structures in Parkinson's voices that in healthy ones. This preliminary result is promising and we are continuing research in this direction.

6.9. Statistics and detection of most unpredictable points in data sets

Participants: Nicolas Brodu, Hussein Yahia, Suman-Kumar Maji.

References: [21], [16].
The assumption that local regularity amounts to predictability can be challenged, depending on the model that one may use to make predictions. A statistical framework, "computational mechanics", has been explicitly designed over the past 30 years, that precisely formalizes notions of causality and predictability within discrete data sets. Patterns with similar causal influence on the data are clustered in equivalence classes. Taken together, these classes form a Markovian automaton by definition, since no extra information is needed from other classes to (statistically) predict the influence of a group of patterns on the rest of the data set. These automata are defined at the lowest data description scale, but it has been suggested that sub-automata (thus clusters at larger scales) form an ideal coarse-graining of the system in terms of predictability (thus also descriptive power). The theory is also deeply rooted in statistical physics, offering a unique perspective on how macroscopic variables could be derived from a microscopic description of a studied system. Preliminary results are promising and show that, for example, edges may be detected in images with a precursor continuous implementation of the theory extension under construction. In order to make more progress, advanced statistical and computational developments are necessary to carry this work. In order to facilitate this development, N. Brodu has submitted a Marie-Curie outgoing fellowship that, if accepted, would allow to partner with Australian leaders on statistics and data processing (University of Melbourne, department of Mathematics).

6.10. Image Reconstruction from Highly Corrupted Gradients

Participants: Hicham Badri, Hussein Yahia, Driss Aboutajdine.

Reference: [23].
Surface-from-Gradients (SfG) is an important step in many imaging applications. It consists in reconstructing an image/surface from corrupted gradient fields, which results in an ill-posed problem. We propose to use sparsity to regularize the problem. We use sparsity in the gradient field together with a robust norm on the data-fitting term (CVPR 2014). In a work in porgress, we make use of a non-local regularization that manipulates non-local similar patches of the corrupted gradient and forcing them to be low-rank. The two approaches significantly outperform previous optimization-based SfG methods on both synthetic and real data.

6.11. Local/Non-Local Noisy Image Deconvolution

Participants: Hicham Badri, Hussein Yahia.

Reference: [24].

Image deconvolution is a standard step in many imaging applications. Sparse local regularization has shown to be fast but tends to over-smoothing images. On the other hand, non-local priors that manipulate similar patches produce better results but tend to be much slower. In this paper, we combine both local and non-local methods in one framework to offer both good quality image reconstruction and computational efficiency in the presence of noise. By studying the non-local singular values of the image patches, we show that the non-local patches tend to be much similar in the blurred version of the image. We thus use low-rank estimation to first estimate a blurred but noise-free image. Secondly, we show that this denoising step introduces outliers in the deconvotion model and propose anefficient optimization method to tackle this problem. Experiments show that the proposed method poduces comparable results to non-local methods while being more computationally efficient.

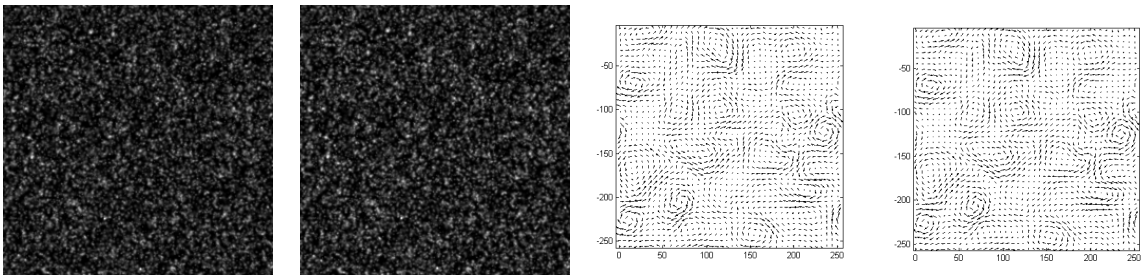

Figure 3. Motion estimation using the proposed method. From left to right: image sequences (2 images, at t and $t+1$ respectively) the ground-truth and the estimated flow (errors, from left to right : MSE=0.063, AAE=3.562, EPE=0.100).

6.12. Detection and dynamics of coastal upwelling

Participants: Ayoub Tamim, Khalid Daoudi, Hussein Yahia, Joël Sudre, Driss Aboutajdine.

References: [18], [34], [35], [33].
An unsupervised classification method is developed for the coarse segmentation of Moroccan coastal upwelling using the Sea Surface Temperature (SST) satellite images. The algorithm is used to provide a seasonal variability of upwelling activity in the southern Moroccan Atlantic coast using 70 Sea Surface Temperature (SST) images of the years 2007 and 2008. The performance of the proposed methodology has been validated by an oceanographer, showing its effectiveness for automatic delimitation of Moroccan upwelling region. We have also explored the applicability of the Fuzzy c-means (FCM) clustering, using an adaptive cluster merging, for the problem of detecting the Moroccan coastal upwelling areas in SST satellite images.

6.13. Nonlinear signal processing for adaptive optics

Participants: Suman-Kumar Maji, Hussein Yahia, Thierry Fusco.

Reference: [31].
The work developped by PhD student Suman Kumar Maji on nonlinear approaches to phase reconstruction in adaptive optics has been presented at the SPIE Astronomical Telescopes + Instrumentation, one of the great events in the field.

6.14. Turbulent Flow Estimation

Participants: Hicham Badri, Hussein Yahia.

We use singularity exponents (SE) to regularize the problem of turbulent flow estimation under the assumption that the brightness constancy constraint holds also for (SE). We also use weighted filtering (Lucas–Kanade's solution) and sparsity on the data-fitting term to improve robusteness to outliers. The proposed motion estimation is built on a Gaussian pyramid and uses the theory of warping for a better estimation of large displacements. Experiements on synthetic data show that the proposed method outperforms sophisticated methods while being simple. See figure 3 .

6.15. Adaptive Transfer Real Image Restoration

Participants: Hicham Badri, Hussein Yahia.

Image restoration is a very challenging task in low-level vision and is extensively used in many imaging applications. Sparsity in various forms (dictionary learning, low-rank estimation,...) has shown to be the key for succesful image denoising. However, the standard noise model used to validate the results is mainly Gaussian and uniform, with known standard deviation. Unfortunately, these assumptions do not hold for real camera noise. Instead of using sparsity to model the singular values of non-local clean similar patches, we use a learning model that trains a mapping between the noisy and ground-truth clean singular values. The training is performed on real camera noise, contrary to previous methods. Experiments show that the proposed method signficantly outperforms previous denoising works on real non-uniform noise and does not require estimating the standard deviation of the corruption. See figure 4 .

Figure 4. Image restoration demonstration on a severely corrupted image. The proposed method leads to a much better resotration quality compared to the standard BM3D method. From left to right: Ground-Truth, Noisy image, BM3D (20.46 dB), Proposed (22.25 dB).

6.16. Augmented Lagrangian for Fast Multi-Sparse Optimization

Participants: Hicham Badri, Hussein Yahia, Khalid Daoudi.

Sparsity has become one of the most important notions in many imaging applications. We address in this paper the problem of multi-sparse optimization, when the energy to minimize contains multiple sparse terms instead of a single one. We show that applying off-the-shelf proximal-based solvers such as ADMM results in a high computational cost due to the complexity of the resulting sub-problems in the case of multi-sparsity. We propose an efficient extension of ADMM for multi-sparse optimization, we study its convergence and complexity and show how it can be applied to computer vision problems. Experiments show that the proposed solver is not only computationally efficient, but also leads quickly to higher-quality results compared to the popular half-quadratic solver.

6.17. On the Fly Hybrid Video Denoising

Participants: Hicham Badri, Hussein Yahia.

Video denoising is a standard pre-processing step in many imaging applications. Non-local methods such as the BM3D method adapted to videos have shown to produce good quality results, but these methods require multiple frames to produce a temporally coherent result, especially when the amount of noise is high. On the other hand, using a hybrid camera, we can get clean images of the scene. However, these images suffer from low-temporal coherence. We present a new approach to video denoising which consists in learning a mapping between the clean images and their corresponding noisy frames and propagate denoising to intermediate frames. To improve temporal coherency, we use a fast method method to sparsify the temporal gradient. Experiments on high-resolution videos show that the proposed method produces good quality on the fly video denoising while being computationally efficient.

MC2 Team

6. New Results

6.1. Highlights of the Year

- **Models for gliomas**
  Glioblastoma multiforme (GBM) causes significant neurological morbidity and short survival times. Brain invasion by GBM is associated with poor prognosis. Recent clinical trials of bevacizumab in newly-diagnosed GBM found no beneficial effects on overall survival times; however, the baseline health-related quality of life and performance status were maintained longer in the bevacizumab group and the glucocorticoid requirement was lower. In a recent work in collaboration with UAB, we have constructed a clinical-scale model of GBM whose predictions uncover a new pattern of recurrence in 11/70 bevacizumab-treated patients. The findings support an exception to the Folkman hypothesis: GBM grows in the absence of angiogenesis by a cycle of proliferation and brain invasion that expands necrosis. Furthermore, necrosis is positively correlated with brain invasion in 26 newly-diagnosed GBM. The unintuitive results explain the unusual clinical effects of bevacizumab and suggest new hypotheses on the dynamic clinical effects of migration by active transport, a mechanism of hypoxia-driven brain invasion.
- **Electroporation modeling** (M. Leguebe, C. Poignard)
  Based on the new discovery of the team of Vectorolgy and anti-cancerous therapies on the membrane lipid oxidation during the pulse delivery, we have provided a model of cell permeabilization that makes it possible to explain the process of electroporation : pore formation during the pulse and surface diffusion of altered lipids after the pulse. Our model explains the long-term effect of electroporation (the permeable state of the membrane lasts a few minutes after the pulse delivery). A 3D-code in C++ has been implemented during the PhD thesis of M. Leguèbe. The team MC2 is now part of the European Lab EBAM on electroporation modeling. An international workshop on Electroporation and Biophysical Therapies was held in Bordeaux the 15th and 16th December.
- Simulation of **multiphysic fluid-structure impacts in 3D**. See http://www.math.u-bordeaux1.fr/ adebrauer/ for astinishing videos.

6.2. Cancer modeling

- **Patient specific simulation for lung metastases**
  The calibration process has tremendously improved by a deep study of the model and its parameter space. Work is ongoing to validate the whole process on a retrospective study of 30 patients. A prototype is being built for our collaborators at Institut Bergonié to use in their clinical routine. The same strategy has been applied to meningiomas in the last year of the post-doc of Julie Joie within the IRL MONICA with a retrospectuve study on 10 patients.
- Modelling of the response to targeted therapies for liver metastasis of a gist : 2 clinical cases with a long term longitudinal follow-up with CT-scans. We are able to fit the volume of the lesion but also the the texture of the image, that is the ratio between necrotic tissues and proliferative ones. See [82].
- Tumor growth model for ductal carcinoma: from in situ phase to stroma invasion. See [71].
- Permeable and conducting states of membrane submitted to electric pulse: non-linear PDE model, 2D and 3D code in C++.
- Free boundary value model for invadopodia and migration of cell developed in collaboration with Osaka University and Tokyo University of Sciences.
- Endothelial cell migration on polymers: agent based model. Paper accepted in DCDS-B.

- A. Peretti started her PhD on the modeling of the heterogenity on renal cancer.
- Benjamin Taton started a post-doc on the modeling of the renal function through perfusion MRI. B. Taton is a MD.
- Th. Michel obtained some mathematical properties on the system of PDEs used for the modelling of GIST metastases.
- **Models for preclinical studies**
  - Mathematical **ODE models of tumor volume kinetics in mice** (collaboration with the Center of Cancer and Systems Biology, Boston, USA and J. Ebos, Roswell Park Cancer Institute, Buffalo, USA).
    Rational and quantitative evaluation of the predictive and descriptive power of the majority of the classical ODE models for tumor growth against data from two distinct experimental systems [57]. One of the major finding was the huge improvement of the predictive properties when using the population *a priori* information on the distribution of the parameters.
  - Mathematical model for data of **preclinical metastatic burden dynamics** and clinical data of metastatic relapse probability of **breast cancer** (collaboration with J. Ebos, Roswell Park Cancer Institute, Buffalo, USA).
    Validation of the descriptive and predictive ability of a simple and minimally parameterized model. The major finding resulting from the modeling analysis was the quantification of the impact of surgery on survival improvement (highly nonlinear), which suggests a threshold primary tumor size for efficacy of the surgery in terms of preventing metastatic recurrence. A publication is in preparation.
  - Effect of **anti-cancer therapies** in preclinical experiments
    * Evaluation of several models (several already published but also new ones) for the effect of **anti-angiogenic drugs**[0] on tumor growth, based on statistical parameter estimation methods on experimental data (collaboration with J. Ebos, Roswell Park Cancer Institute, Buffalo, USA). The main finding was one model that was able to both describe the effect of the drug (Sunitinib) and predict the effect when changing the scheduling. See [66].
    * Effect of the **sequence of administration between cytotoxic and anti-angiogenic drugs** (collaboration with J. Ciccolini and D. Barbolosi, SMARTc, Inserm, Marseille, Fr). See [84].
- **Theoretical cancer biology**
  - Theories of metastatic initiation (collaboration with A. Bikfalvi, LAMC, Inserm and the RMSB, CNRS in Bordeaux, Fr).
    Confrontation of theories and experimental data challenged the classical view of metastatic establishment and growth and suggested that tumors could merge in initial phases. Quantitative impact of the merging was studied using a dedicated and properly calibrated spatial model.
  - Tumor-tumor distant interactions (collaboration with the Center of Cancer and Systems Biology, Boston, USA).
    Statistical and modeling analysis of experimental data for two tumors implanted in one organism.

[0] recent anti-cancer drugs that target the tumor vasculature rather than the cancer cells themselves

6.3. Newtonian fluid flows simulations and their analysis

- Development of a high-order (third order in time and space) level-set method which allow to compute consistently the curvature of the interface even for long times (L. Weynans, F. Luddens and M. Bergmann)
- Development of a sharp cartesian method for the simulation of incompressible flows with high density ratios, like air-water interfaces. This method is inspired from the second-order cartesian method for elliptic problems with immersed interfaces developed in Cisternino-Weynans [69]
- Study of the convergence in 1D and 2D of the method developed in Cisternino-Weynans [69]

REALOPT Project-Team

6. New Results

6.1. Highlights of the Year

- Olivier Beaumont and Lionel Eyraud-Dubois have received the HiPC best paper award for their work on resource allocation for large scale virtualized platforms with reliability guarantees. They provided a formulation based on a thorough analysis of a real life usage trace, and a very efficient two-step allocation algorithm.
- The team organized the annual conference of the French Operations Research Society ROADEF14 in Feb 2014.
- An Inria Innovation Lab has been created between Realopt and Ertus Consulting.
- The SAMBA associated team project with Brazil was renewed for 3 years including new collaborators from Chili.
- François Vanderbeck was invited as a plenary speaker at the conference OPTIMIZATION 2014, in Portugal [19].

6.2. Automation and combination of linear-programming based stabilization techniques in column generation

We reviewed in [88] stabilization techniques that can improve in practice the convergence of a column generation algorithm. Proximal methods based on penalising the deviation from the incumbent dual solution have become standards of the domain. However, the analysis of such methods is important to understand the mechanism on which they rely, to appreciate the difference between methods, and to derive intelligent schemes to adjust their parameters. As stabilization procedures for column generation can be viewed as cutting plane strategies in the dual problem, the link with cutting plane separation strategies can be exploited to enlarge the scope of methods and to refine their analysis. In [24], [40], we focus on stabilization schemes that rely solely on a linear programming (LP) solver for the column generation master program. This restrictive scope captures the most common implementations where one uses an LP solver to handle the master program. For dual price smoothing techniques, we analyse the link with the in-out separation strategy and we 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. Also, the dynamic adjustments, compared to a static setting, allows to adapt the parameters to the phase of the algorithm. We provide extensive test reports that highlight the comparative performances of such scheme and validate our self-adjusting parameter scheme. Furthermore, our results show that using smoothing in combination with penalty function yields a cumulative effect on convergence speed-ups [35]. We have also consider other stabilization strategies inspired form algorithmic strategies have been designed to accelerate convergence of cutting plane algorithms in mixed integer programming. In [37], we show that the "Multi-Point Separation" strategy translates into a column generation stabilization technique that consists in restricting the dual solution to be in the convex hull of the selected multi-point set. We have also considered other stabilization strategies inspired from algorithmic strategies that have been designed to accelerate convergence of cutting plane algorithms in mixed integer programming. In [37], we show that the "Multi-Point Separation" strategy translates into a column generation stabilization technique that consists in restricting the dual solution to be in the convex hull of the selected multi-point set.

6.3. Multi-Stage Column generation strategies

In [39], we propose another mechanism to improve the performance of column generation algorithms. We study the application of branch-and-price approaches to the automatic version of the Software Clustering Problem. To tackle this problem, we apply the Dantzig-Wolfe decomposition to a formulation from literature. Given this, we present two Column Generation (CG) approaches to solve the linear programming relaxation of the resulting reformulation: the standard CG approach, and a new approach, which we call Staged Column Generation (SCG). Also, we propose a modification to the pricing subproblem that allows to add multiple columns at each iteration of the CG. We test our algorithms in a set of 45 instances from the literature. The proposed approaches were able to improve the literature results solving all these instances to optimality. Furthermore, the SCG approach presented a considerable performance improvement regarding computational time, number of iterations and generated columns when compared with the standard CG as the size of the instances grows.

6.4. Aggregation techniques to reduce the size of column generation models

We proposed an aggregation method to reduce the size of column generation (CG) models for a class of set-covering problems in which the feasible subsets depend on a resource constraint. The aggregation relies on a correlation between the resource consumption of the elements and the corresponding optimal dual values. The aggregated model obtained allows to find good quality lower bounds more rapidly than the original CG algorithm. The speedup is due to less primal and dual variables in the master, and to an aggregated pricing sub-problem. To guarantee optimaly, we designed an algorithm that iteratively refines the aggregation until the CG optimum is reached. Computational results prove the usefulness of our methods.

6.5. Dual-feasible functions

Dual-feasible functions have been used in the past to compute fast lower bounds and valid inequalities for different combinatorial optimization and integer programming problems. Until now, all the dual-feasible functions proposed in the literature were 1-dimensional functions, and were defined only for positive arguments. In [12] we extended the principles of dual-feasible functions to the m-dimensional case by introducing the concept of vector packing dual-feasible function. We explored the theoretical properties of these functions in depth, and we proposed general schemes for generating some instances of these functions. Additionally, we proposed and analyzed different new families of vector packing dual-feasible functions. All the proposed approaches were tested extensively using benchmark instances of the 2-dimensional vector packing problem. Our computational results showed that these functions can approximate very efficiently the best lower bounds for this problem. In a second paper, currently submitted to a journal, we show that extending these functions to negative arguments raises many issues. Additionally, we describe different construction principles to obtain dual-feasible functions with domain and range $\mathbb{R}$. Specific instances obtained from these principles are proposed and analyzed.

6.6. Resource Allocation and Scheduling in Large Scale Distributed Platforms.

We have considered several problems arising in the context of large scale platforms, that are characterized by their heterogeneity, the difficulty of predicting performance and the risk failures. In [13], we concentrate on heterogeneity issues in collective communication schemes where the goal is to broadcast a message to a set of nodes. In particular, we consider a realistic model in the context of large scale distributed platforms where some nodes may lie behind NATs or firewalls and may be therefore unable to forward the message between them. In [21], [20], we consider resource allocation problems that arise in large scale data centers. In [20], we analyze the main characteristics of the services in a huge trace corresponding to an actual data center and that has been released recently by google. In the same context, in [21], we concentrate on issues related to fault tolerance by over subscribing services in order to guarantee quality of service in a failure prone environment. At last, the difficulty to predict the actual performance of resources made it very popular to rely on dynamic scheduling algorithms where scheduling decisions are made at runtime. In [22], we analyze the performance of such a dynamic scheduling algorithm in terms on number of induced communications for outer product and matrix multiplication kernels.

6.7. Employee timetabling with time varying demand

We addressed a multi-activity tour scheduling problem with time varying demand. The objective is to compute a schedule for a fixed roster in order to minimize the over-coverage and the under-coverage of different parallel activity demands along a planning horizon. Numerous complicating constraints are present in our problem: all employees are different and can perform several different activities during the same day-shift, lunch breaks and pauses are flexible, demand is given for 15 minutes periods. To the best of our knowledge, the work in [29] is the first attempt to combine days-off scheduling, shift scheduling, shift assignment, activity assignment, pause and lunch break assignment. To solve this problem, we developed several methods: a compact linear Mixed Integer Programming model, a branch-and-price like approach with a nested dynamic program to solve heuristically the subproblems, a diving heuristic, and a greedy heuristic based on our subproblem solver. The computational results, based on both real cases and instances derived from real cases, demonstrate that our methods are able to provide good quality solutions in a short computing time. Our algorithms are now embedded in a commercial software, which is already in use in a mini-mart company.

6.8. Time-dependent formulations for routing problems

The paper [16] presents a new formulation for the Time-Dependent Travelling Salesman Problem (TDTSP). We start by reviewing well known natural formulations with some emphasis on the formulation by Picard and Queyranne (1978). The main feature of this formulation is that it uses, as a subproblem, an exact description of the n-circuit problem. Then, we present a new formulation that uses more variables and is based on using, for each node, a stronger subproblem, namely a n-circuit subproblem with the additional constraint that the corresponding node is not repeated in the circuit. Although the new model has more variables and constraints than the original PQ model, the results given from our computational experiments show that the linear programming relaxation of the new model gives, for many of the instances tested, gaps that are close to zero. Thus, the new model is worth investigating for solving TDTSP instances. We have also provided a complete characterization of the feasible set of the corresponding linear programming relaxation in the space of the variables of the PQ model. This characterization permits us to suggest alternative methods of using the proposed formulations.

A well-known formulation for the unit-demand capacitated vehicle routing problem uses a single commodity flow system to represent the delivery of the items. The vehicle capacity is modeled by imposing a maximum capacity on the arcs used by the flow. In [30], we used a time-dependent formulation for the problem to derive, by projection, tighter bounding inequalities on the arcs. The first experiments show that these new inequalities permit to improve significantly the linear relaxation bound of the single commodity flow formulation. We are currently studying separation algorithms in order to generate dynamically these new inequalities.

6.9. Vehicle routing for dial-a-ride problems

Static and deterministic vehicle routing problems cannot be used in many real-life systems, as input data are not reliable and revealed over time. In [11], we study a pickup and delivery problem with time windows accounting for maximum ride time constraints – the so-called dial-a-ride problem – in its static and dynamic variant, and we make specific proposal on robust optimization models for this problem. To solve the static model, we develop a branch-and-price approach that handles ride time constraints in the process of generating feasible vehicle routes in the course of the optimization procedure. The work is focussed on the pricing problem solver and acceleration techniques for the branch-and-price approach. Our numerical results show that the method is competitive compared to existing approaches that are based on branch-and-cut. In the dynamic context, where some input data are revealed or modified over time, we apply our branch-and-price algorithm for re-optimization in a rolling horizon approach.

6.10. A MILP approach to minimize the number of late jobs with and without machine availability constraints

The study in [14] investigates scheduling problems that occur when the weighted number of late jobs that are subject to deterministic machine availability constraints have to be minimized. These problems can be modeled as a more general job selection problem. Cases with resumable, non-resumable, and semi-resumable jobs as well as cases without availability constraints are investigated. The proposed efficient mixed integer linear programming approach includes possible improvements to the model, notably specialized lifted knapsack cover cuts. The method proves to be competitive compared with existing dedicated methods: numerical experiments on randomly generated instances show that all 350-job instances of the test bed are closed for the well-known problem $1|r_i|\sum w_i U_i$. For all investigated problem types, 98.4% of 500-job instances can be solved to optimality within one hour.

6.11. Two phase solution for an intelligent moving target search problem based on a 0–1 linear model

We developed a generic discrete model for the moving, intelligent target problem. Our objective is to maximise the probability of detection of the moving target with respect to target and searcher's constraints. The solution method proposed in [15] is composed of two stages. The first one aims at providing a large-scale strategy based on an Integer Linear Program approach. As a direct solution of this problem is not practically possible, we use a decomposition of the problem into a searcher's strategy on one side, and the target's strategy on the other side. A good strategy for the searcher is determined using a sliding window procedure. Concerning the target, our approach consists in simulating some of the target's possible strategies and considering each of these possibilities as an independent and deterministic entity. The second stage is dedicated to adjusting the large-scale strategy provided by stage 1. Numerical results are presented so as to assess the impact of our approach.

6.12. Computing the Chromatic index and clique numberof special graphs

In our paper [17] on the strong chromatic index of planar graphs with large girth, we prove that every planar graph with maximum degree Δ (let Δ be an integer) and girth at least $10\Delta + 46$ is strong $(2\Delta - 1)$-edge-colorable, that is best possible (in terms of number of colors) as soon as G contains two adjacent vertices of degree Δ. This improves the best previous result when $\Delta \geq 6$. In [18] we show how one can compute the clique number of a-perfect graphs in polynomial time. A main result of combinatorial optimization is that clique and chromatic number of a perfect graph are computable in polynomial time (Grötschel, Lovasz and Schrijver 1981). This result relies on polyhedral characterizations of perfect graphs involving the stable set polytope of the graph, a linear relaxation defined by clique constraints, and a semi-definite relaxation, the Theta-body of the graph. A natural question is whether the algorithmic results for perfect graphs can be extended to graph classes with similar polyhedral properties. In [18] we consider a superclass of perfect graphs, the a-perfect graphs, whose stable set polytope is given by constraints associated with generalized cliques. We show that for such graphs the clique number can be computed in polynomial time as well. The result strongly relies upon Fulkersons's antiblocking theory for polyhedra and Lovasz's Theta function.

CARMEN Team

6. New Results

6.1. Highlights of the Year

- New associated team EPICARD (principal investigator N. zemzemi, Y. Coudière and J. Henry). The aim of of this associated team for the first year is to overcome the technical difficulties that we pointed out during the year 2014 in inverse problem for the heart.
- June 2014: Based on a peer-reviewed proposal, the Grand équipement national de calcul intensif (GENCI) has attributed us 3 million core-hours on the national high-performance computing system Turing, to be used in the year 2014.
- December 2014: Based on a peer-reviewed proposal, the Grand équipement national de calcul intensif (GENCI) has attributed us 3.5 million core-hours on the national high-performance computing machines Turing, Curie, and Occigen, to be used in the year 2015.
- LIRYC will fund a 2-year postdoctoral position on simulation of Brugada syndrome, a rare ECG anomaly predictive of sudden cardiac death in young, apparently healthy subjects. This work will be performed in tight collaboration with clinicians at the Haut-Lévèque hospital

6.2. Inverse problem

We tested our method using synthetic data generated with a highly realistic forward model. Propagating action potentials were generated using a monodomain reaction-diffusion model with a Ten Tusscher 2006 membrane model. An anisotropic human heart model at 0.2-mm resolution was used for this purpose. Torso potentials were then computed from the simulated transmembrane currents using a finite-difference torso model at 1-mm resolution with intracavitary blood, anisotropic myocardium, lungs, and an anisotropic skeletal muscle layer. We simulated 20 cases: 5 single stimuli, 1 dual stimulus and 14 re-entry simulations. From the simulated torso potentials a 200-channel body surface map recording was extracted and used to test the inverse methods. Inverse solutions in terms of epicardial potentials were computed both with MFS and with our .new optimal control approach. With our algorithms, we were able to construct the electrical potential on the heart surface with a very good accuracy in terms of correlation coefficient. Thus, we could accurately reconstruct the activation pattern.

6.3. CEPS

- Integration of some ionic models into CEPS (N. Zemzemi and F. Caro). Those developments will allow us now to use CEPS for the benchmark named *Second N-version Cardiac Electrophysiology Benchmark Specification actual developments*, see (benchmarck) for more details.
- Development of usefull tools for the code (compilation in order to check the depandancies, validation and coverage of the code). 16 test cases are now implementated in CEPS. Those test cases concern unit test case as test for small resolution of linear system (for the FE P1 implemented into the code) and validatation test case as the heat PDE problem. Tests verify also the parallel implementation.
- At this time, the development of the bi-domain model in CEPS is in progress in CEPS with N. Zemzemi.
- First integration of the new model of S. Labarthe initiated during his PhD with L. Colin. This task needs improvment for validation in terms of development.

6.4. Numerical Scheme

Y. Coudière, C. Pierre and R. Turpault wrote some new high order FV schemes. The goal of this study is a future implementation in CEPS.

6.5. Mathematical Model

M. Potse, P.E. Becue and F. Caro wrote a new model for numerical simulations for cardiac electrophysiology at the microscopic scale. We interfer with the LIRYC in order to describe, as much as possible, the interactions between the extra-cellular medium and the intra-cellular medium.

MAGIQUE-3D Project-Team

6. New Results

6.1. Inverse Problems

6.1.1. Complex-frequency domain Full Waveform Inversion

Participants: Florian Faucher, Maarten V. de Hoop, Henri Calandra.

We study the seismic inverse problem for the (complex) frequency-domain elastic isotropic wave equation; in particular the recovery of the Lamé parameters and density. We employ a Full Waveform Inversion where the iterative minimization is based on a gradient descent. The elastic inverse problem shows a Lipschitz-type stability where the Fréchet derivative has a strictly positive 'lower bound'. This bound is connected to the stability constant and can be approximated using the Gauss-Newton Hessian. The successive stability estimates provide a control of the convergence and decide the parameters of inversion. We develop a multi-level approach based on a structured domain partitioning of the sub-surface. The coefficients (Lamé parameters and density) are assumed to be piecewise constant functions following the domain partitioning, which is naturally defined with the successive stability estimates to maintain the radius of convergence, while refinement provides resolution. It allows us to start with minimal prior information for the coefficients and the algorithm is perfectly suitable for complex frequency. We have carried out numerical experiments in two and three dimensions; those results have been presented during the following conferences in 2014: [48], [49].

6.1.2. Imaging of complex media with elastic wave equations

Participants: Jérôme Luquel, Hélène Barucq, Henri Calandra, Julien Diaz.

Even if RTM has enjoyed the tremendous progresses of scientific computing, its performances can still be improved, in particular when applied to strong heterogeneous media. In this case, images have been mainly obtained by using direct arrivals of acoustic waves and the transition to elastic waves including multiples is not obvious essentially because elastic waves equations are still more computationally consuming. The accuracy of numerical wave fields is obviously of great importance. We have thus chosen to consider high-order Discontinuous Galerkin Methods which are known to be well-adapted to provide accurate solutions based upon parallel computing. Now one of the main drawback of RTM is the need of storing a huge quantity of information which is redhibitory when using elastic waves. For that purpose, we apply the Griewank algorithm following Symes' ideas for the acoustic RTM. The idea is to find a compromise between the number of wave equations to solve and the number of numerical waves that we have to store. This is the so-called Optimal Checkpointing. By reducing the occupancy of the memory, RTM should be efficient even when using elastic waves. By this way, one may wonder if considering elastic waves including multiples in order to improve images of heterogeneous media is a valid option. It must involve a careful numerical analysis including the evaluation of the impact of the imaging condition. It is thus necessary to derive accurate imaging conditions, which could take advantage of all the information contained in the wavefield. For acoustic media, Claerbout proposed an imaging condition which is widely used and turns out to be sufficient to accurately reproduce interfaces. But Claerbout conditions do not take wave conversions into account and it is not clear wether conversions do or do not contain interesting information to get accurate images of heterogeneous media.

Since P-wave and S-wave interact with each other, it might be relevant to use an imaging condition including these interactions. In fact, this has been done successfully by J.Tromp and C. Morency for seismology applications based upon the inversion of the global Earth. Their approach is based upon the state adjoint and it involves sensitivity kernels which are defined from the propagated and the back-propagated fields. Now it has been shown that full wave form inversions using these sensitivity kernels may be polluted by numerical artefacts. One solution is to use a linear combination of the sensitivity kernels to delete artefacts. In this work, we propose then a new imaging condition which construction is inspired from with some approximations required to keep admissible computational costs. We illustrate the properties of the new imaging condition on

industrial benchmarks like the Marmousi model. In particular, we compare the new imaging condition with other imaging conditions by using as criteria the quality of the image and the computational costs required by the RTM. This work was presented at the the WCCM XI - ECCM V - ECFD VI - Barcelona 2014 Conference and SIAM Conference on IMAGING SCIENCE (SIAM-IS14) Hong Kong Baptist University[67].

6.1.3. Helioseismology

Participants: Juliette Chabassier, Marc Duruflé, Thorsten Hohage.

We have begun to write a software interface that allows to solve an inverse problem using adjoint and regularization methods (iTReg software) while using Montjoie software for the direct problem that must be solved at each iteration of the inversion process.

6.2. Modeling

6.2.1. High-Order Time Schemes

6.2.1.1. Fourth order energy-preserving locally implicit discretization for linear wave equations

Participants: Juliette Chabassier, Sébastien Imperiale.

A family of fourth order coupled implicit-explicit schemes is presented as a special case of fourth order coupled implicit schemes for linear wave equations. The domain of interest is decomposed into several regions where different fourth order time discretization are used, chosen among a family of implicit or explicit fourth order schemes derived in [72]. The coupling is based on a Lagrangian formulation on the boundaries between the several non conforming meshes of the regions. A global discrete energy is shown to be preserved and leads to global fourth order consistency. Numerical results in 1d and 2d illustrate the good behavior of the schemes and their potential for the simulation of realistic highly heterogeneous media or strongly refined geometries, for which using everywhere an explicit scheme can be extremely penalizing. Accuracy up to fourth order reduces the numerical dispersion inherent to implicit methods used with a large time step, and makes this family of schemes attractive compared to second order accurate methods in time. This work has been presented at the Franco-Russian workshop on mathematical geophysics, Sep 2014, Novosibirsk, Russia [58], at the and is the object of a submitted publication to International Journal for Numerical Methods in Engineering.

6.2.1.2. A new modified equation approach for solving the wave equation

Participants: Hélène Barucq, Henri Calandra, Julien Diaz, Florent Ventimiglia.

In order to obtain high-order time-schemes, we are considering an alternative approach to the ADER schemes and to the modified equation technique described in section 3.2 . The two first steps of the construction of the schemes are similar to the previous schemes : we apply a Taylor expansion in time to the solution of the wave equation and we replace the high-order derivatives with respect to the time by high order space operators, using the wave equation. The difference is that we do not use auxiliary variables and we choose to discretize directly the high-order operators in space.

In the framework of the PhD thesis of Florent Ventimiglia, we have extended this new method involving p-harmonic operator to the first order formulation of the acoustic wave equation, which is the formulation discretized in the DIVA platform of TOTAL. In this case, the high order operators in space are not are not powers of the Laplace operator but powers of the gradient. Hence, we also had to adapt the space discretization, and we have extended the DG formulation with centered fluxes proposed in [77] to higher order operators. A numerical analysis of performance in 2D indicates that, for a given accuracy, this method requires less computational costs and less storage than the High-Order ADER Scheme. These results have been presented to the AIMS conference [54]. A paper has been published in ESAIM Proceedings [19].

6.2.2. Finite Element Methods for the time-harmonic wave equation.

6.2.2.1. Goal-Oriented Adaptivity using Unconventional Error Representations

Participants: Vincent Darrigrand, David Pardo, Ignacio Muga, Hélène Barucq.

In the scope of subsurface modelling via the resolution of inverse problems, the so-called goal-oriented adaptivity plays a fundamental role. Indeed, while classical adaptive algorithms were first designed to accurately approximate the energy norm of a problem [69], [70], one requires a good approximation of a specific quantity of interest. An energy norm driven self-adaptive strategy can still be used for that purpose, although it often becomes sub-optimal and unable to provide an accurate solution for the required quantity of interest in a reasonable amount of time.

During the late 90's, to overcome this issue, the so-called goal-oriented strategy appeared, see for instance [82], [81]. The goal-oriented approach consists in expressing the error in the quantity of interest as an integral over the entire computational domain involving the errors of the original and adjoint problems, and then minimise an upper bound of such error representation by performing local refinements.

Most authors, using the adjoint problem, represent the approximation error in the quantity of interest via the global bilinear form that describes the problem in terms of local and computable quantities.

Our methodology, however, is based on the selection of an alternative bilinear form exhibiting better properties than the original bilinear form (e.g. positive definiteness). We represent the residual error functional of the adjoint problem through this alternative form. We can then compute new upper bounds of the error of the quantity of interest in a similar way than with the classical approach. Our main goal is to demonstrate that a proper choice of such alternative form may improve the upper bounds of the error representation.

Moreover, the method proposed here generalises the existing ones, since, in particular, we can select as the alternative bilinear form the one associated to the adjoint problem.

6.2.2.2. Hybridizable Discontinuous Galerkin method for the elastic Helmholtz equations

Participants: Marie Bonnasse-Gahot, Henri Calandra, Julien Diaz, Stéphane Lanteri.

We consider Discontinuous Galerkin (DG) methods formulated on fully unstructured meshes, which are more convenient than finite difference methods on cartesian grids to handle the topography of the subsurface. DG methods and classical Finite Element (FE) methods mainly differ from discrete functions which are only piecewise continuous in the case of DG approximation. DG methods are then more suitable than Continuous Galerkin (CG) methods to deal with hp-adaptivity. This is a great advantage to DG method which is thus fully adapted to calculations in highly heterogeneous media. Nevertheless, the main drawback of classical DG methods is that they are more expensive in terms of number of unknowns than classical CG methods, especially when arbitrarily high order interpolation of the field components is used. In this case DG methods lead to larger sparse linear systems with a higher number of globally coupled degrees of freedom as compared to CG methods with a same given mesh. In that case, we consider a hybridizable Discontinuous Galerkin (HDG) method which principle consists in introducing a Lagrange multiplier representing the trace of the numerical solution on each face of the mesh cells. This new variable exists only on the faces of the mesh and the unknowns of the problem depend on it. This allows us to reduce the number of unknowns of the global linear system. Now the size of the matrix to be inverted only depends on the number of the faces of the mesh and on the number of the degrees of freedom of each face. It is worth noting that for the classical DG method it depends on the number of the cells of the mesh and on the number of the degrees of freedom of each cell. The solution to the initial problem is then recovered thanks to independent elementwise calculation. The principle of the HDG method and 2D results were presented at the WCCM XI - ECCM V - ECFD VI - Barcelona 2014 Conference [41], the EAGE Workshop on High Performance Computing for Upstream [42], the Second Russian-French Workshop "Computational Geophysics" [43] and at the Réunion des Sciences de la Terre 2014 conference [53]. A comparison between HDG method and classical nodal DG method was given on a poster at the Journées Total-Mathias 2014 workshop [66].

6.2.2.3. Helioseismology

Participants: Hélène Barucq, Juliette Chabassier, Marc Duruflé, Damien Fournier, Laurent Gizon.

The finite element code Montjoie 5.2 has been used to solve Helmholtz equation in axisymmetric domain in the configuration of the sun. The efficiency of the code has been compared in three configurations : radial (1-D mesh and spherical harmonics), axisymmetric (2-D mesh), 3-D. The results have convinced our-selves and our partners of Max Planck Institute that the axisymmetric configuration is the most interesting for an inversion procedure, since 3-D computations are too expensive. A more realistic modeling of the sun requires the solution of time-harmonic Galbrun's equations (instead of Helmholtz equation), different formulations have been implemented and studied. It appeared that the different numerical methods are not able to converge to the correct solution for non-uniform flows. The lack of convergence is more obvious for flows with a larger Mach number. Such problems do not appear in Linearized Euler equations, as a result we have proposed simplified Galbrun's equations that converge correctly and provide the same solution as original Galbrun's equations for a null flow. These equations have been implemented in 2-D, axisymmetric and 3-D configuration.

6.2.2.4. Scattering of acoustic waves by a disc - Hypersingular integral equations

Participants: Leandro Farina, Paul Martin, Victor Péron.

Two-dimensional boundary-value problems involving a Neumann-type boundary condition on a thin plate or crack can often be reduced to one-dimensional hypersingular integral equations. Examples are potential flow past a rigid plate, acoustic scattering by a hard strip, water-wave interaction with thin impermeable barriers, and stress fields around cracks. In [29], we generalize some of these results to two-dimensional hypersingular integral equations. Thus, rather than integrating over a finite interval, we now integrate over a circular disc. Two-dimensional hypersingular equations over a disc arise, for example, in the scattering of acoustic waves by a hard disc; this particular application is described in Appendix A. We develop an appropriate spectral (Galerkin) method, using Fourier expansions in the azimuthal direction and Jacobi polynomials in the radial direction. The Hilbert-space arguments used by Golberg are generalized and a convergence theorem is proved by using tensor-product techniques. Our results are proved in weighted L^2 spaces. Then, Tranter's method is discussed. This method was devised in the 1950s to solve certain pairs of dual integral equations. It is shown that this method is also convergent because it leads to the same algebraic system as the spectral method.

6.2.2.5. Finite Element Subproblem Method

Participants: Patrick Dular, Christophe Geuzaine, Laurent Krähenbühl, Victor Péron.

In the paper [26], the modeling of eddy currents in conductors is split into a sequence of progressive finite element subproblems. The source fields generated by the inductors alone are calculated at first via either the Biot-Savart law or finite elements. The associated reaction fields for each added conductive region, and in return for the source regions themselves when massive, are then calculated with finite element models, possibly with initial perfect conductor and/or impedance boundary conditions to be further corrected. The resulting subproblem method allows efficient solving of parameterized analyses thanks to a proper mesh for each subproblem and the reuse of previous solutions to be locally corrected.

6.2.2.6. High Order Methods for Helmholtz Problems in Highly Heterogeneous Media

Participants: Théophile Chaumont-Frelet, Henri Calandra, Hélène Barucq, Christian Gout.

Heterogeneous Helmholtz problems arise in various geophysical application where they modelize the propagation of time harmonic waves through the subsurface. For example, in inversion problems, the aim is to reconstruct a map of the underground based on surface acquisition. This recovery process involves the solution to several Helmholtz problems set in different media, and high frequency solutions are required to obtain a detailed image of the underground. This obervations motivate the design of efficient solver for highly heterogeneous Helmholtz problems at high frequency.

The main issue with the discretization of high frequency problems is the so called "pollution effect" which impose drastic condition on the mesh. In the homogeneous case, it is known that one efficient way to reduce the pollution effect is the use of high order discretization methods. However, high order methods can not be applied as is to highly heterogeneous media. Indeed, they are based on coarser mesh and are not sensitive to fine scale variations of the medium.

We propose to overcome this difficulty by using a multiscale strategy to take into account fine scale heterogeneities on coarse meshes. The method is based on a simple medium approximation method, which can be seen as a special quadrature rule. Numerical experiments in two dimensional geophyscial benchmarks show that high order method coupled with our multiscale approximation medium stragey are cheaper than low order method for a given accuracy. Futhermore, focusing on one dimensional models, we were able to show from a theoretical point of view that our methology reduces the pollution effect even when used on coarse meshes with non-matching interfaces.

This work has been presented at the *WCCM XI - ECCM V - ECFD VI - Barcelona 2014* conference, the *Second Russian-French Workshop "Computational Geophysics"*. A poster has been presented at the *journées Total-Mathias 2014 workshop*. A paper has been submitted for publication to *Math. Of Comp.*.

6.2.3. Boundary conditions.

6.2.3.1. Absorbing Boundary Conditions for Tilted Transverse Isotropic Elastic Media

Participants: Lionel Boillot, Hélène Barucq, Julien Diaz, Henri Calandra.

The seismic imaging simulations are always performed in bounded domains whose external boundary does not have physical meaning. We have thus to couple the wave equations with boundary conditions which aim at reproduce the invisibility of the external boundary. The discretization of these conditions can be an issue. For instance, an efficient condition, once discretized, can induce huge computational costs by filling the matrix which has to be inverted. This is the case of the transparent boundary conditions which are approximated by local Absorbing Boundary Conditions (ABC) that do not increase to much the computational burden. However, the ABC has the drawback to introduce spurious numerical waves which can perturb the RTM results. It is possible to avoid this drawback by applying PML (Perfectly Matched Layers) but it proves to be unstable in anisotropic media. Last year, we proposed a way of construction leading to a stable ABC. The technique is based on slowness curve properties, giving to our approach an original side. We established stability results from long time energy behavior and we have illustrated the performance of the new condition in 2D numerical tests. This year, we extend all these results to 3D case and to arbitrary boundary shapes. The previous paper submission on 2D results has been accepted and released [18]. The recent results in 3D have been presented to the ECCOMAS conference.

6.2.3.2. Derivation of high order absorbing boundary conditions for the Helmholtz equation in 2D.

Participants: Hélène Barucq, Morgane Bergot, Juliette Chabassier, Élodie Estecahandy.

Numerical simulation of wave propagation raises the issue of dealing with outgoing waves. In most of the applications, the physical domain is unbounded and an artificial truncation needs indeed to be carried out for applying numerical methods like finite element approximations. Adapted boundary conditions that avoid the reflection of outgoing waves and provide a well-posed mathematical problem must then be derived. With ideal boundary conditions, the solution on the new mixed boundary valued problem in the truncated domain would actually be equal to the restriction of the mathematical solution in the unbounded domain. However, such ideal boundary conditions, called "transparent boundary conditions", can be shown to be nonlocal, which leads to dramatic computational overcosts. The seek of local boundary conditions, called "absorbing boundary conditions" (ABC), has been the object of numerous works trying to perform efficient conditions based on different techniques of derivation. Among them, the technique of micro-diagonalisation has been employed to the wave equation and more generally to hyperbolic systems in [76], leading to a hierarchy of absorbing local boundary conditions based on the approximation of the Dirichlet-to-Neumann map. A comprehensive review of different used strategies and higher order conditions can be found in [85]. One desirable property of ABCs is that the reflection of the waves on the artificial boundary generates an error of the same order as the one generated by the spatial discretization inside the domain. The computational effort is thus optimized in terms of modeling and numerical inaccuracies. Moreover, the ABC must fit the artificial boundary chosen by the user of the method. In the context of high order spatial discretization (spectral finite elements [74], Interior Penalized Discontinuous Galerkin [68]), there is nowadays a need for high order ABCs that can adapt on non flat geometries since these methods prove very efficient for capturing arbitrary shaped domains.

The aim of the present work is to develop high order ABCs for the Helmholtz equation, that can adapt to regular shaped surfaces. A classical way of designing ABCs is to use Nirenberg theorem [80] on the second order formulation of the Helmholtz equation, which enables us to decompose the operator as a product of two first order operators. Here our approach is to rewrite the Helmholtz equation as a first order system of equations before developing ABCs using M.E. Taylor's micro-diagonalisation method [84]. Then an asymptotic truncation must be performed in order to make the ABC local, and we will see that the high frequency approximation will lead to more usable ABCs than the one stating that the angle of incidence is small. During the process, while increasing the degree of the pseudo differential operator decomposition along with the order of asymptotic truncation, we retrieve classical ABCs that have been found with other techniques by other authors. For now, we have restricted ourselves to two dimensions of space, but despite the fact that 3D generalization should obviously generate more calculation, no further theoretical difficulties are expected.

This work has been the object of a technical report [61] and the obtained conditions have been implemented in Montjoie 5.2 and Houd10ni 5.1 .

6.2.4. Asymptotic modeling.

6.2.4.1. Fast Simulation of Through-casing Resistivity Measurements Using Semi-analytical Asymptotic Models.
Participants: Victor Péron, David Pardo, Aralar Erdozain.

When trying to obtain a better characterization of the Earth's subsurface, it is common to use borehole through-casing resistivity measurements. It is also common for the wells to be surrounded by a metal casing to protect the well and avoid possible collapses. The presence of this metal case highly complicates the numeric simulation of the problem due to the high conductivity of the casing compared to the conductivity of the rock formations. In this study [47] we present an application of some theoretical asymptotic methods in order to deal with complex borehole scenarios like cased wells. The main idea consists in replacing the part of the domain related to the casing by a transmission impedance condition. The small thickness of the casing makes it ideal to apply this kind of mathematical technique. When eliminating the casing from the computational domain, the computational cost of the problems considerably decreases, while the effect of the casing does not disappear due to the impedance transmission conditions. The results show that when applying an order three impedance boundary condition for a simplified domain, it only generates a negligible approximation error, while it considerably reduces the computational cost. For obtaining the numerical results and testing the mathematical models we have developed a Finite Element Code in Matlab. The code works with Lagrange polynomials of any degree as basis functions and triangular shaped elements in two dimensions. The code has been adapted for working with the transmission impedance conditions required by the mathematical models.

6.2.4.2. Modeling the propagation of ultrashort laser pulses in optical fibers.
Participants: Mohamed Andjar, Juliette Chabassier, Marc Duruflé.

In order to model the propagation of an ultrashort laser pulse, the most natural idea is to solve Maxwell's equations in a nonlinear and dispersive medium. Given the considered optical periods (around 10^{-14} seconds), the associated wavelengthes (around 1 millimeter) and the propagation distances (several meters), the direct numerical simulation of these equations by usual numerical techniques (finite elements, explicit time schemes) is impossible because too expensive. The standard procedure is therefore to use approached equations obtained by exploiting legitimate hypotheses in the considered context (slowly varying pulse envelope, narrow spectrum, paraxial approximation ...). These new equations, among them the Nonlinear Schrödinger Equation, are significantly less expensive to solve and we can therefore provide realistic numerical simulations to physicists.

When the pulse propagates in an optical fiber, its spatial profile in the orthogonal plane to the propagation direction in very simple because optical fibers posses a finite (small, often equal to one) number of propagating modes. The equations that originally are stated on a 3D domain can then be written as one spatial dimension equations.

The scientific objective of this internship was to apply the approximation techniques mentioned above in this specific context, in order to obtain one or several equations (depending on the used hypotheses) that model the propagation of ultrashort laser pulses in optical fibers. A matlab code has been developed and integrated in the C++ code Montjoie 5.2 . Numerical simulations have been led in order to observe classical situations of nonlinear fiber optics (Kerr effect, Raman effect, supercontinuum generation, ...).

6.2.4.3. Small heterogeneities in the context of time-domain wave propagation equation : asymptotic analysis and numerical calculation

Participants: Vanessa Mattesi, Sébastien Tordeux.

We have focused our attention on the modeling of heterogeneities which are smaller than the wavelength. The work can be decomposed into two parts : a theoretical one and a numerical one. In the theoretical one, we derive a matched asymptotic expansion composed of a far-field expansion and a near-field expansion. The terms of the far-field expansion are singular solutions of the wave equation whereas the terms of the near-field expansion satisfy quasistatic problems. These expansions are matched in an intermediate region. We justify mathematically this theory by proving error estimates. In the numerical part, we describe the Discontinuous Galerkin method, a local time stepping method and the implementation of the matched asymptotic method. Numerical simulations illustrate these results. Vanessa Mattesi has defended her PhD on this topic[14].

6.2.4.4. Theoretical and numerical investigations of acoustic response of a multiperforated plate for combustion liners

Participants: Vincent Popie, Estelle Piot, Sébastien Tordeux.

Multiperforated plates are used in combustion chambers for film cooling purpose. As the knowledge of the acoustic response of the chamber is essential for preventing combustion instabilities, the acoustic behaviour of the perforated plates has to be modeled. This can be done either by considering the transmission impedance of the plates, or their Rayleigh conductivity.

We have investigated the link between these two quantities thanks to matched asymptotic expansions. Especially the far-field or near-field nature of the physical quantities used in the definition of the impedance and Rayleigh quantity has been enlightened. Direct numerical simulations of the propagation of an acoustic plane wave through a perforated plate are performed and post-treated so that the assumptions underlying the definitions of impedance and Rayleigh conductivity have been checked. The results will be presented at the conference ASME Turbo Expo 2015.

6.3. High Performance methods for solving wave equations

6.3.1. Coupling the DG code with task programming libraries

Participants: Lionel Boillot, Emmanuel Agullo, George Bosilca, Henri Calandra.

The parallelization of the original code is based on a preliminary step of domain decomposition and then on the use of the MPI (Message Passing Interface) library. It is a common choice which works pretty well in most of the classical architectures. However, the parallel efficiency is not optimal and the performance decreases in hybrid architectures. Indeed, we know the number of operations that each sub-domain has to performed but this does not give us the exact time that the computations require. The cluster heterogeneity leads to various automatic optimizations (memory cache, parallel capability, ...) which are difficult to measure. We have decided to tackle this problem by modifying the parallelism with the use of task programming. We have thus rewritten the DIVA algorithm in a graph of tasks without using the MPI library and we have left to the runtime PaRSEC the choice of when and where to execute each task. The numerical experiments we have performed have confirmed the significant improvement of the parallel efficiency on different architectures like ccNUMA machines or Intel Xeon Phi co-processors. Moreover, the proposed solution is portable on these architectures, this means that none or few modifications are required in the code, allowing to focus on algorithmic aspects in order to preserve the performance. These results have been presented to the EAGE HPC workshop and to the HPCC IEEE conference within a paper have been accepted.

MAGNOME Project-Team

6. New Results

6.1. Highlights of the Year

In collaboration with colleagues from the Institut du Vigne et du Vin (ISVV), Bordeaux and the Universidade Nova de Lisboa, Lisbon we used a population genomics approach to investigate the global phylogeography and domestication fingerprints of winemaking yeasts, using a collection of isolates obtained from fermented beverages and from natural environments on five continents. These results appeared in *Nature Communications* [11].

6.2. A Gondwanan imprint of S. uvarum diversity

Domestication of livestock and crops has been amply demonstrated through historical and archeological records, but domestication of microorganisms is much more difficult to establish. In a large-scale study [11] of the wine and cider yeast *Saccharomyces uvarum* conducted with colleagues from the Institut du Vigne et du Vin, Bordeaux and the Universidade Nova de Lisboa, Lisbon, we found the first indications of its domestication in the transition from its habitat in *Nothofagus* (southern beech) trees on the Gondwana mega-continent, to its present-day diversity in the Holarctic. The global phylogeography of these microorganisms was investigated through genome sequencing and comparison of 54 strains isolated on five continents, resulting in the identification of 10^5 high-quality SNPs and a remarkable pattern of introgressions ([11] figure 3 http://dx.doi.org/10.1038/ncomms5044).

The 54 genomes in this study were isolated, selected, and sequenced, and both assembled and aligned against reference genomes. Phylogenies were based on concatenated SNP alignment of selected chromosomes. The structure of the population was investigated using model-based Bayesian clustering.

In addition to the biological result, this study illustrates the ubiquity of an experimental approach based on large-scale sequencing of highly related genomes, in order to isolate tiny differences linked to a trait of interest. This is in contrast to the strategy that was current eight years ago, based on sequencing of a modest number of genomes spanning a much greater evolutionary range.

6.3. Improving inference of metabolic models

Participants: David James Sherman [correspondant], Pascal Durrens, Razanne Issa, Anna Zhukova.

The Pantograph approach uses reference model annotated by *gene associations*, and voting between complementary predictions of homology between reference genes and target genes, to decide whether a reaction that is present in the scaffold ought be be present in the target. A gene association implicitly represents expert knowledge about the role of genes in a compact way. If the gene association can be rewritten into a possibly satisfiable formula, then the corresponding reaction is instantiated in the target model.

Historically, gene associations have been used intuitively by experts during the model design and curation process, and are often inconsistent. We have formalized the construction of gene associations based on the semantics of different interpretations, showing how different boolean formulas should be constructed when the application is *i*) metabolic model inference, *ii*) flux-balance analysis, *iii*) hierarchical modeling, or *iv*) dynamic simulation (Razanne Issa, MS in prep.).

Second, we have refined our strategy for inferring metabolic models using abductive logic. We have shown that given a set of genes as observations in the target organism, and rules for rewriting gene associations while respecting integrity constraints for the model, then the reactions in the target model can be abduced as hypotheses that "explain" the presence of a maximial number of genes in the target genome. The advantage of this approach is that it can invent, through specialization, reactions that are not present *per se* in the reference model. Two classes of reactions can be invented: substrate-specific reactions inferred from expansion in gene families, and transport reactions needed to maintain model integrity for constitutive compartments.

6.4. Knowledge-based generalization of metabolic models

Participants: David James Sherman [correspondant], Pascal Durrens, Razanne Issa, Anna Zhukova.

Large metabolic networks are hard to understand and curate, because the large number of detailed reactions, which are needed for accurate modeling and simulation, obscure the high-level structure of the reaction network. We defined knowledge-based methods that factor similar reactions into “generic” reactions in order to visualize a whole pathway or compartment, while maintaining the underlying model so that the user can later “drill down” to the specific reactions if need be[15], [16] An implementation of this method is available as a Python library (see paragraph 5.3).

Figures 2 and 3 illustrate model generation for *Yarrowia lypolitica* fatty acid oxidation in the peroxisome. Molecular species are represented using SBGN notation: as circular nodes, and the reactions as square ones, connected by edges to their reactants and products. Ubiquitous species are of smaller size and colored gray. Non-ubiquitous species are divided into six equivalence classes, and coloured accordingly. The size of the model does not allow for readability of the species labels, thus we do not show them (figure 2).

The specific model is appropriate for simulation, because it contains all of the precise reactions. The generalized model is suited for a human, because it reveals the main properties of the model and masks distracting details. For example, the generalized model highlights the fact that there is a particularity concerning *C24:0-CoA (stearoyl-CoA)* (yellow): there exists a "shortcut" reaction (orange), producing it directly from another *fatty acyl-CoA* (yellow), avoiding the usual four-reaction beta-oxidation chain, used for other *fatty acyls-CoA*. This shortcut is not obvious in the specific model, because it is hidden among a plethora of similar-looking reactions.

We formally defined the generalization method in [15] and showed how to calculate it using a good approximation to an NP-complete set cover problem. The method was further validated in a collection of 1283 inferred models and revealed, on the one hand, a number of probable errors in the inferred models, and on the other hand, that there exist different families of generalization with a plausible link to different adaptive responses.

6.5. Characterization of STAND protein families

Participants: David James Sherman, Pascal Durrens, Witold Dyrka [correspondant].

In collaboration with Sven Saupe and Mathieu Paoletti from IBGC Bordeaux (ANR Mykimun), we worked on characterization of the STAND protein family in the fungal phylum. We established an *in silico* screen based on state-of-the-art bioinformatic tools, which – starting from experimentally studied sequences from *Podospora anserina* – allowed us to determine the first systematic picture of fungal STAND protein repertoire (ms. in preparation). Most notably, we found evidence of extensive modularity of domain associations, and signs of concerted evolution within the recognition domain [13]. Both results support the hypothesis that fungal STAND proteins, originally described in the context of vegetative incompatibility, are involved in a general fungal immune system. In addition, we investigated improved protein domain representations and elaborated a grammatical modelling method [23], which will be used to elucidate mechanisms of formation and operation of the STAND proteins.

NLR domains identified in this work have been incorporated into the upcoming release of Pfam [0].

To further explor the underlying mechanisms of repeat formation we implemented a stochastic string rewriting system that models the generation process of highly internally conserved repeats. The system is grounded in the biology of the process as it models transformation of repeats through the events of unequal crossing-over and mutation, which are believed to be main mechanisms that produce diversity in repeats. We confirmed that highly variable sites identified on the basis of entropy, are subject to selective pressure towards composition typical for binding sites, which is consistent with the suggested role of recognition epitopes.

[0] http://pfam.xfam.org

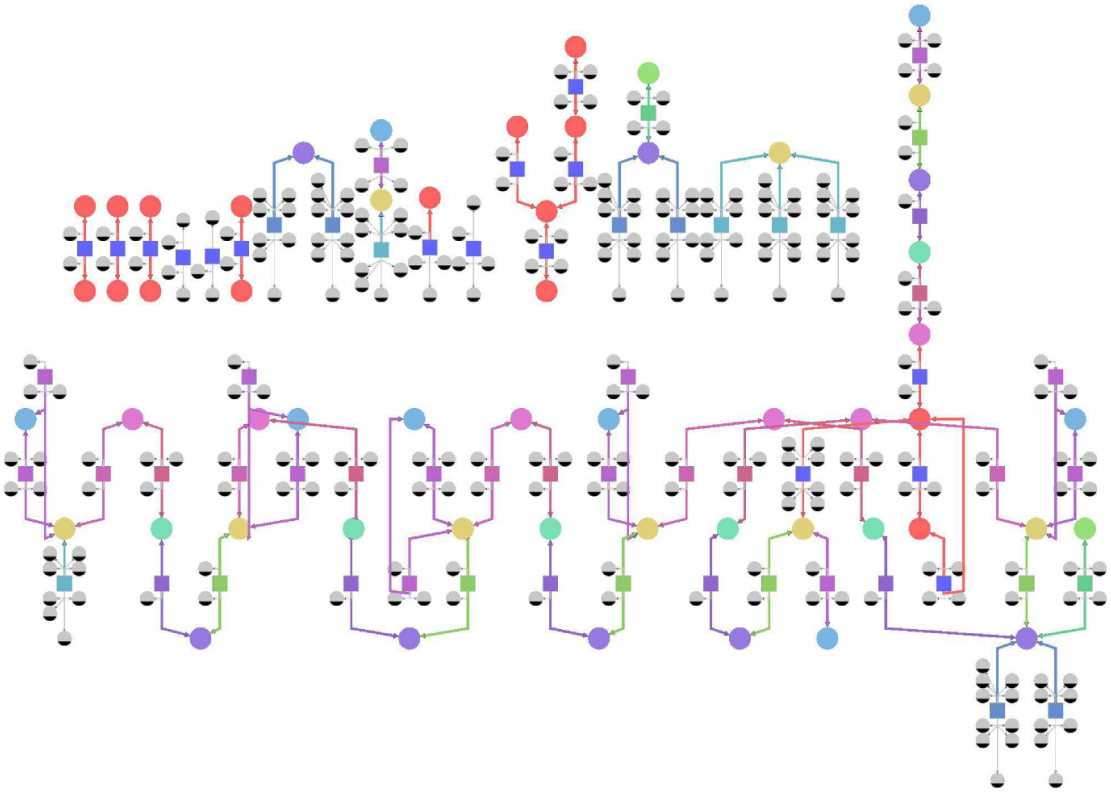

Figure 2. Yarrowia lypolitica fatty acid oxidation model before generalization. Reactions of the specific model are divided into fifteen equivalence classes, represented by different colours. Generally speaking, β-oxidation is a transformation of fatty acyl-CoA (yellow) into dehydroacyl-CoA (violet), then into hydroxyacy fatty acyl-CoA (dark green), 3-ketoacyl-CoA (magenta), and back to fatty acyl-CoA (with a shorter carbon chain); while the specific model describes the same process in more details, specifying those reactions for each of the fatty acyl-CoA species presented in the organisms' cell (e.g. decanoyl-CoA, dodecanoyl-CoA, etc.). This high-level, repetitive structure is obscured by the detail of the individual reactions.

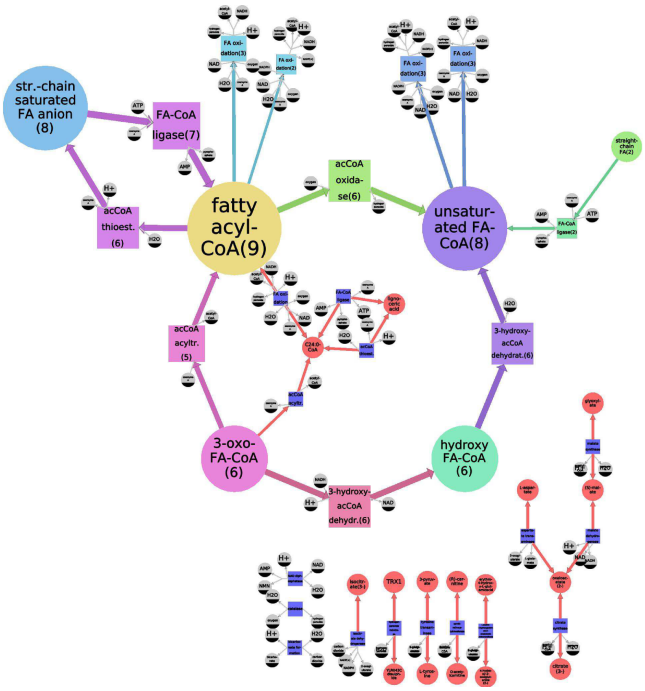

Figure 3. Generalization of the Yarrowia lypolitica fatty acid oxidation model, described as a transformation of fatty acyl-CoA (yellow) into dehydroacyl-CoA (violet), then into hydroxyacy fatty acyl-CoA (dark green), 3-ketoacyl-CoA (magenta), and back to fatty acyl-CoA with a shorter carbon chain. The generalization algorithm identifies equivalent molecular species using an ontology, and groups together reactions that operate on the same abstract species. It finds the greatest generalization the preserves stoichiometry. The generalized model represents quotient species and reactions. For example, the violet dehydroacyl-CoA node is a quotient of hexadec-2-enoyl-CoA, oleoyl-CoA, tetradecenoyl-CoA, trans-dec-2-enoyl-CoA, trans-dodec-2-enoyl-CoA, trans-hexacos-2-enoyl-CoA, trans-octadec-2-enoyl-CoA, and trans-tetradec-2-enoyl-CoA (colored violet in figure 2). In a similar manner, the light-green acyl-CoA oxidase quotient reaction, that converts fatty acyl-CoA (yellow) into dehydroacyl-CoA (violet), generalizes six corresponding light-green reactions of the initial model (figure 2).

MNEMOSYNE Project-Team

6. New Results

6.1. Overview

Though our view is systemic, our daily research activities are concerned with the design, at a given scale of description, of models of neuronal structures, each concerned with a specific learning paradigm. Of course, a major challenge is to keep in mind the systemic view, to put a specific emphasis on the way each neuronal structure communicates with the rest of the system and to highlight how the learning paradigm interplays with other memory systems.

Among the numerous loops involving the brain, the body and the environment, a basic grid of description corresponds to distinguish "Perception Loops", the goal of which is to extract from the inner and outer world sensory invariants helpful to identify and evaluate the current state and to make predictions from previous learning, and "Action Loops", the goal of which is to rely on this sensory, emotional and motivational information to decide, plan and trigger actions for the benefit of the body.

Presently, our team is engaged on the following topics: Concerning perception loops, we are firstly considering the role of the hippocampus and of the posterior cortex in learning high level sensory cues that contribute to pavlovian conditioning in the amygdala. Secondly, we are investigating the role of the thalamus in attentional shifts in the cortex. Concerning Action loops, we are preparing a critical analysis of the current views of the interactions between the prefrontal cortex and the basal ganglia. Finally, we also report here more methodological achievements.

6.2. Pavlovian conditioning

Pavlovian conditioning is an outstanding example of a systemic process involving several cerebral structures and several modes of learning. This year, we have made more precise our model of amygdala [7], with a special emphasis on the variety of its inputs, including by neuromodulation [12]. We have also specifically discussed and contrasted the role of cerebral structures involved in this learning, from the point of view of information processing [11]. In addition to the amygdala, the structures of interest are the hippocampus and the posterior and prefrontal cortex and begin to be investigated in ongoing studies.

6.3. The thalamus is more than a relay

Many recent results in neuroscience indicate that the role of the thalamus in the brain is certainly more important than it used to be considered, particularly concerning its relation with the cortex. Our modeling and bibliographic studies were carried out in the Keops project (*cf.* § 7.2) with our chilean neuroscientist colleagues studying non standard ganglion cells in the retina. Particularly, the PhD work by Carlos Carvajal [1] led us to propose a biologically-founded algorithm for the interplay between the modulatory and driving connections between the thalamus and the cortex, with the strong constraint of proposing a system working on a real visual flow.

6.4. On the computational efficiency of Basal Ganglia models

Many valuable models have been proposed to capture the richness of the fundamental relations between the basal ganglia and prominent brain structures including the prefrontal cortex, the hippocampus and the superior colliculus. To choose among them the mechanisms on which to build the design of the motor pole of our brain-inspired system, a fundamental issue is to evaluate the efficiency of these models in more realistic cases than the ones which are generally considered by the authors [24]. For this reason, we have conducted a systematic study of several basal ganglia computational models to check of their scalability in terms of ation representational space [25]. Unfortunately, we found most of them to not be scalable and some of them to not be reproducible at all.

Another way to explore the computational efficiency of neuronal models is to implement them at lower levels of description. This is currently being done with one model developed in our lab at a level corresponding to a neuronal assembly with a mean activity expressed using a single variable. This mesoscopic approach has been refined to a microscopic scale description level, i.e taking into account individual neurons and synapses. Besides the confirmation of many of the results of initial model with a more detailed formalism, this new model has allowed us to highlight the facilitating role of inhibitory interneurons in the decision-making and action selection processes.

6.5. Distributed Self-Organization

The formation of the sensory homunculus in the primary sensory cortex (SI) is believed to be the result of a dynamic neural self-organization process that starts before birth and lasts for several years, allowing the brain to cope with sensory or brain lesions. The exact neural mechanisms driving this self-organization are not yet known and the role of the somatosensory attention remains unclear in this picture. We thus investigated the influence of somatosensory attention onto the two-dimensional structure of area 3b neuronal receptive fields (RFs) using a computational model [2] based on the dynamic neural field theory. This computational model of SI (area 3b) is able to explain experimental data in the monkey and hypothesizes role for the somatosensory attention in the shaping of SI receptive fields.

SISTM Team

6. New Results

6.1. Highlights of the Year

A work (described below), in collaboration with M. Davis and R. Tibshirani from Standford University, has been published in the "Proceedings of the National Academy of Sciences" : [8].

Females have generally more robust immune responses than males for reasons that are not well-understood. Here we used a systems analysis to investigate these differences by analyzing the neutralizing antibody response to a trivalent inactivated seasonal influenza vaccine (TIV) and a large number of immune system components, including serum cytokines and chemokines, blood cell subset frequencies, genome-wide gene expression, and cellular responses to diverse in vitro stimuli, in 53 females and 34 males of different ages. We found elevated antibody responses to TIV and expression of inflammatory cytokines in the serum of females compared with males regardless of age. This inflammatory profile correlated with the levels of phosphorylated STAT3 proteins in monocytes but not with the serological response to the vaccine. In contrast, using a machine learning approach, we identified a cluster of genes involved in lipid biosynthesis and previously shown to be up-regulated by testosterone that correlated with poor virus-neutralizing activity in men. Moreover, men with elevated serum testosterone levels and associated gene signatures exhibited the lowest antibody responses to TIV. These results demonstrate a strong association between androgens and genes involved in lipid metabolism, suggesting that these could be important drivers of the differences in immune responses between males and females.

6.2. Analysis of purely random forests bias

In collaboration with S. Arlot, we write a research report on some theoretical results about random forests : [30].

Random forests are a very effective and commonly used statistical method, but their full theoretical analysis is still an open problem. As a first step, simplified models such as purely random forests have been introduced, in order to shed light on the good performance of random forests. In this paper, we study the approximation error (the bias) of some purely random forest models in a regression framework, focusing in particular on the influence of the number of trees in the forest. Under some regularity assumptions on the regression function, we show that the bias of an infinite forest decreases at a faster rate (with respect to the size of each tree) than a single tree. As a consequence, infinite forests attain a strictly better risk rate (with respect to the sample size) than single trees. Furthermore, our results allow to derive a minimum number of trees sufficient to reach the same rate as an infinite forest. As a by-product of our analysis, we also show a link between the bias of purely random forests and the bias of some kernel estimators.

HIEPACS Project-Team

6. New Results

6.1. Highlights of the Year

In the context of HPC-PME initiative, we started a collaboration with ALGO'TECH INFORMATIQUE and we have organised one of the first PhD-consultant action implemented by Xavier Lacoste led by Pierre Ramet. ALGO'TECH is one of the most innovative SMEs (small and medium sized enterprises) in the field of cabling embedded systems, and more broadly, automatic devices. The main target of the project is to validate the possibility to use the sparse linear solvers of our team in the area of electromagnetic simulation tools developped by ALGO'TECH. This collaboration will be developed next year in the context of the European project FORSTISSIMO. The principal objective of FORTISSIMO is to enable European manufacturing, particularly SMEs, to benefit from the efficiency and competitive advantage inherent in the use of simulation.

As a conclusion of the OPTIDIS project we organized the first International Workshop on Dislocation Dynamics Simulations that was devoted to the latest developments realized worldwide in the field of Discrete Dislocation Dynamics simulations. This international event held in December 10th to the 12th at "Maison de la Simulation" in Saclay, France and attracted 55 participants from many different countries including England, Germany, France, USA, ... The workshop gathered most of the active researchers working on dislocation dynamics from numerical simulations to experimentatios. Thanks to the success of this workshop, a second one will be scheduled in England during 2016.

6.2. High-performance computing on next generation architectures

6.2.1. Composing multiple StarPU applications over heterogeneous machines: a supervised approach

Enabling HPC applications to perform efficiently when invoking multiple parallel libraries simultaneously is a great challenge. Even if a uniform runtime system is used underneath, scheduling tasks or threads coming from different libraries over the same set of hardware resources introduces many issues, such as resource oversubscription, undesirable cache flushes or memory bus contention.

This work presents an extension of `StarPU`, a runtime system specifically designed for heterogeneous architectures, that allows multiple parallel codes to run concurrently with minimal interference. Such parallel codes run within *scheduling contexts* that provide confined execution environments which can be used to partition computing resources. Scheduling contexts can be dynamically resized to optimize the allocation of computing resources among concurrently running libraries. We introduce a *hypervisor* that automatically expands or shrinks contexts using feedback from the runtime system (e.g. resource utilization). We demonstrate the relevance of our approach using benchmarks invoking multiple high performance linear algebra kernels simultaneously on top of heterogeneous multicore machines. We show that our mechanism can dramatically improve the overall application run time (-34%), most notably by reducing the average cache miss ratio (-50%).
This work is developed in the framework of Andra Hugo's PhD. These contributions have been published in the international journal of High Performance Computing Applications [21].

6.2.2. A task-based $\mathcal{H}$-Matrix solver for acoustic and electromagnetic problems on multicore architectures

$\mathcal{H}$-Matrix is a hierarchical, data-sparse approximate representation of matrices that allows the fast approximate computation of matrix products, LU and LDL^T decompositions, inversion and more. This representation is suitable for the direct solution of large dense linear systems arising from the Boundary Element Method in $O(N \log_2^{\alpha}(N))$ operations. This kind of formulation is widely used in the industry for the numerical simulation of acoustics and electromagnetism scattering by large objetcs. Applications of this approach include

aircraft noise reduction and antenna sitting at Airbus Group. The recursive and irregular nature of these $\mathcal{H}$-Matrix algorithms makes an efficient parallel implementation very challenging, especially when relying on a "Bulk Synchronous Parallel" paradigm. We have considered an alternative parallelization for multicore architectures using a task-based approach on top of a runtime system, namely StarPU. We have showed that our method leads to a highly efficient, fully pipelined computation on large real-world industrial test cases provided by Airbus Group.
This research activity has been conduced in the framework of the EADS-ASTRIUM, Inria, Conseil Régional initiative in collaboration with the RUNTIME Inria project, and is part of Benoit Lize's PhD.

6.2.3. A task-based 3D geophysics application

Reverse Time Migration (RTM) technique produces underground images using wave propagation. A discretization based on the Discontinuous Galerkin (DG) method unleashes a massively parallel elastodynamics simulation, an interesting feature for current and future architectures. We have designed a task-based version of this scheme in order to enable the use of manycore architectures. At this stage, we have demonstrated the efficiency of the approach on homogeneous and cache coherent Non Uniform Memory Access (ccNUMA) multicore platforms (up to 160 cores) and designed a prototype version of a distributed memory version that can exploit multiple instances of such architectures. This work has been conducted in the context of the DIP Inria-Total strategic action in collaboration with the MAGIQUE3D Inria project and thanks to the long-term visit of George Bosilca funded by TOTAL. Geroge's expertise ensured an optimum usage of the PaRSEC runtime system onto which our task-based scheme has been ported.
This work was presented during HPCC conference [27] as well as during a TOTAL scientific event [26].

6.2.4. Resiliency in numerical simulations

For the solution of systems of linear equations, various recovery-restart strategies have been investigated in the framework of Krylov subspace methods to address the situations of core failures. The basic underlying idea is to recover fault entries of the iterate via interpolation from existing values available on neighbor cores. In that resilience framework, we have extended the recovey-restart ideas to the solution of linear eigenvalue problems. Contrary to the linear system case, not only the current iterate can be interpolated but also part of the subspace where candidate eigenpairs are searched.
This work is developed in the framework of Mawussi Zounon's PhD funded by the ANR RESCUE. These contributions have been presented in particuler at the international SIAM workshop on Exascale Applied Mathematics Challenges and Opportunities [40] in Chicago and the Householder symposium [41] in Spa. Notice that theses activities are also part of our contribution to the G8 ESC (Enabling Climate Simulation at extreme scale).

6.2.5. Hierarchical DAG scheduling for hybrid distributed systems

Accelerator-enhanced computing platforms have drawn a lot of attention due to their massive peak computational capacity. Despite significant advances in the programming interfaces to such hybrid architectures, traditional programming paradigms struggle mapping the resulting multi-dimensional heterogeneity and the expression of algorithm parallelism, resulting in sub-optimal effective performance. Task-based programming paradigms have the capability to alleviate some of the programming challenges on distributed hybrid many-core architectures. In this work we take this concept a step further by showing that the potential of task-based programming paradigms can be greatly increased with minimal modification of the underlying runtime combined with the right algorithmic changes. We propose two novel recursive algorithmic variants for one-sided factorizations and describe the changes to the PaRSEC task-scheduling runtime to build a framework where the task granularity is dynamically adjusted to adapt the degree of available parallelism and kernel efficiency according to runtime conditions. Based on an extensive set of results we show that, with one-sided factorizations, i.e. Cholesky and QR, a carefully written algorithm, supported by an adaptive tasks-based runtime, is capable of reaching a degree of performance and scalability never achieved before in distributed hybrid environments.
These contributions will be presented at the international conference IPDPS 2015 [36] in Hyderabad.

6.3. High performance solvers for large linear algebra problems

6.3.1. Parallel sparse direct solver on runtime systems

The ongoing hardware evolution exhibits an escalation in the number, as well as in the heterogeneity, of the computing resources. The pressure to maintain reasonable levels of performance and portability, forces the application developers to leave the traditional programming paradigms and explore alternative solutions. PaStiX is a parallel sparse direct solver, based on a dynamic scheduler for modern hierarchical architectures. In this paper, we study the replacement of the highly specialized internal scheduler in PaStiX by two generic runtime frameworks: PaRSEC and StarPU. The tasks graph of the factorization step is made available to the two runtimes, providing them with the opportunity to optimize it in order to maximize the algorithm efficiency for a predefined execution environment. A comparative study of the performance of the PaStiX solver with the three schedulers - native PaStiX, StarPU and PaRSEC schedulers - on different execution contexts is performed. The analysis highlights the similarities from a performance point of view between the different execution supports. These results demonstrate that these generic DAG-based runtimes provide a uniform and portable programming interface across heterogeneous environments, and are, therefore, a sustainable solution for hybrid environments.

This work has been developed in the framework of Xavier Lacoste's PhD funded by the ANR ANEMOS. These contributions have been presented at the Heterogeneous Computing Workshop held jointly with the international conference IPDPS 2014 [32]. Xavier Lacoste will defend his PhD in February 2015.

6.3.2. Hybrid parallel implementation of hybrid solvers

In the framework of the hybrid direct/iterative MaPHyS solver, we have designed and implemented an hybrid MPI-thread variant. More precisely, the implementation relies on the multi-threaded MKL library for all the dense linear algebra calculations and the multi-threaded version of PaStiX. Among the technical difficulties, one was to make sure that the two multi-threaded libraries do not interfere with each other. The resulting software prototype is currently experimented to study its new capability to get flexibility and trade-off between the parallel and numerical efficiency. Parallel experiments have been conducted on the Plafrim plateform as well as on a large scale machine located at the USA DOE NERSC, which has a large number of CPU cores per socket.
This work is developed in the framework of the PhD thesis of Stojce Nakov funded by TOTAL.

6.3.3. Designing LU-QR hybrid solvers for performance and stability

New hybrid LU-QR algorithms for solving dense linear systems of the form $Ax = b$ have been introduced. Throughout a matrix factorization, these algorithms dynamically alternate LU with local pivoting and QR elimination steps, based upon some robustness criterion. LU elimination steps can be very efficiently parallelized, and are twice as cheap in terms of flops, as QR steps. However, LU steps are not necessarily stable, while QR steps are always stable. The hybrid algorithms execute a QR step when a robustness criterion detects some risk for instability, and they execute an LU step otherwise. Ideally, the choice between LU and QR steps must have a small computational overhead and must provide a satisfactory level of stability with as few QR steps as possible. In this work, we introduce several robustness criteria and we establish upper bounds on the growth factor of the norm of the updated matrix incurred by each of these criteria. In addition, we describe the implementation of the hybrid algorithms through an extension of the PaRSEC software to allow for dynamic choices during execution. Finally, we analyze both stability and performance results compared to state-of-the-art linear solvers on parallel distributed multicore platforms.
These contributions have been presented at the international conference IPDPS 2014 [30] in Phoenix. An extended version has been submitted to JPDC journal.

6.3.4. Divide and conquer symmetric tridiagonal eigensolver for multicore architectures

Computing eigenpairs of a symmetric matrix is a problem arising in many industrial applications, including quantum physics and finite-elements computation for automobiles. A classical approach is to reduce the matrix to tridiagonal form before computing eigenpairs of the tridiagonal matrix. Then, a back-transformation allows one to obtain the final solution. Parallelism issues of the reduction stage have already been tackled in different shared-memory libraries. In this work, we focus on solving the tridiagonal eigenproblem, and we describe a novel implementation of the Divide and Conquer algorithm. The algorithm is expressed as a sequential task-flow, scheduled in an out-of-order fashion by a dynamic runtime which allows the programmer to play with tasks granularity. The resulting implementation is between two and five times faster than the equivalent routine from the INTEL MKL library, and outperforms the best MRRR implementation for many matrices.
These contributions will be presented at the international conference IPDPS 2015 [34] in Hyderabad.

6.4. High performance Fast Multipole Method for N-body problems

Last year we have worked primarily on developing an efficient fast multipole method for heterogeneous architecture. Some of the accomplishments for this year include:

1. implementation of some new features in the FMM library ScalFMM: adaptive variants of the Chebyshev and Lagrange interpolation based FMM kernels, multiple right-hand sides, generic tensorial nearfield...
2. The parallelization and the FMM core parts rely on ScalFMM (OpenMP/MPI) which has been updated all year round. Finally, ScalFMM offers two new shared memory parallelization strategies using OpenMP 4 and StarPU.

6.4.1. Low rank approximations of matrices

New fast algorithms for the computation of low rank approximations of matrices were implemented in a -soon to be- open-source C++ library. These algorithms are based on randomized techniques combined with standard matrix decompositions (such as QR, Cholesky and SVD). The main contribution of this work is that we make use of ScalFMM parallel library in order to power the large amount of matrix to vector products involved in the algorithms. Applications to the fast generation of Gaussian random fields were adressed. Our methods compare good with the existing ones based on Cholesky or FFT and potentially outpass their performances for specific distributions. We are currently in the process of writing a paper on that topic. Extensions to fast Kalman filtering is now considered. This work is done in collaboration with Eric Darve (Stanford, Mechanical Engineering) in the context of the associate team FastLA.

6.4.2. Time-domain boundary element method

The Time-domain Boundary Element Method (TD-BEM) has not been widely studied but represents an interesting alternative to its frequency counterpart. Usually based on inefficient Sparse Matrix Vector-product (SpMV), we investigate other approaches in order to increase the sequential flop-rate. We present a novel approach based on the re-ordering of the interaction matrices in slices. We end up with a custom multi-vectors/vector product operation and compute it using SIMD intrinsic functions. We take advantage of the new order of the computation to parallelize in shared and distributed memory. We demonstrate the performance of our system by studying the sequential Flop-rate and the parallel scalability, and provide results based on an industrial test-case with up to 32 nodes [43], [28]. From the middle of year 2014, we started working on the TD FMM for the BEM problem. A non optimized version is able to solve the TD BEM with the FMM on parallel distributed nodes. All the implementations should be in high quality in the Software Engineering sense since the resulting library is going to be used by industrial applications.
This work is developed in the framework of Bérenger Bramas's PhD and contributes to the EADS-ASTRIUM, Inria, Conseil Régional initiative.

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

6.5.1. Dynamic load balancing for massively parallel coupled codes

In the field of scientific computing, load balancing is a major issue that determines the performance of parallel applications. Nowadays, simulations of real-life problems are becoming more and more complex, involving numerous coupled codes, representing different models. In this context, reaching high performance can be a great challenge. In the PhD of Maria Predari (started in october 2013), we develop new graph partitioning techniques, called co-partitioning, that address the problem of load balancing for two coupled codes: the key idea is to perform a "coupling-aware" partitioning, instead of partitioning these codes independently, as it is usually done. More precisely, we propose to enrich the classic graph model with *interedges*, that represent the coupled code interactions. We describe two new algorithms, called AWARE and PROJREPART, and compare them to the currently used approach (called NAIVE). In recent experimental results, we notice that both AWARE and PROJREPART algorithms succeed to balance the computational load in the coupling phase and in some cases they succeed to reduce the coupling communications costs. Surprisingly we notice that our algorithms do not degrade the global graph edgecut, despite the additional constraints that they impose. In future work, we aim at validating our results on real-life cases in the field of aeronautic propulsion. In order to achieve that, we plan to integrate our algorithms within the Scotch framework. Finally, our algorithms should be implemented in parallel and should be extended in order to manage more complex applications with more than two interacting models.

6.5.2. Graph partitioning for hybrid solvers

Nested Dissection has been introduced by A. George and is a very popular heuristic for sparse matrix ordering before numerical factorization. It allows to maximize the number of parallel tasks, while reducing the fill-in and the operation count. The basic standard idea is to build a "small separator" S of the graph associated with the matrix in order to split the remaining vertices in two parts P_0 and P_1 of "almost equal size". The vertices of the separator S are ordered with the largest indices, and then the same method is applied recursively on the two sub-graphs induced by P_0 and P_1. At the end, if k levels of recursion are done, we get 2^k sets of independent vertices separated from each other by $2^k - 1$ separators. However, if we examine precisely the complexity analysis for the estimation of asymptotic bounds for fill-in or operation count when using Nested Dissection ordering, we can notice that the size of the halo of the separated sub-graphs (set of external vertices belonging to an old separator and previously ordered) plays a crucial role in the asymptotic behavior achieved. In the perfect case, we need halo vertices to be balanced among parts. Considering now hybrid methods mixing both direct and iterative solvers such as HIPS, MaPHyS, obtaining a domain decomposition leading to a good balancing of both the size of domain interiors and the Scalable numerical schemes for scientific applications size of interfaces is a key point for load balancing and efficiency in a parallel context. This leads to the same issue: balancing the halo vertices to get balanced interfaces. For this purpose, we revisit the algorithm introduced by Lipton, Rose and Tarjan which performed the recursion of nested dissection in a different manner: at each level, we apply recursively the method to the sub-graphs But, for each sub-graph, we keep track of halo vertices. We have implemented that in the Scotch framework, and have studied its main algorithm to build a separator, called greedy graph growing.
This work is developed in the framework of Astrid Casadei's PhD. These contributions have been presented at the international conference HIPC 2014 [29] in Goa.

6.6. Application Domains

6.6.1. Dislocation dynamics simulations in material physics

6.6.1.1. Long range interaction

Various optimizations have been performed in the Dislocation Dynamics code OptiDis for the long-ranged isotropic elastic force and energy models using a Fast Fourier based Fast Multipole Method (also known as

Uniform FMM). Furthermore the anisotropic elastic force model was implemented using spherical harmonics expansions of angular functions known as Stroh matrices. Optimizations with respect to the crystallographic symmetries were also considered. Once the corresponding semi-analytic formulae for the force field are derived this method should compare well with existing approaches based on expanding the anisotropic elastic Green's function.

6.6.1.2. Parallel dislocation dynamics simulation

This year we have focused on the improvements of our hybrid MPI-OpenMP parallelism of the OptiDis code. More precisely, we have continued the development of the cache-conscious data structure to manage efficiently large set of data (segments and nodes) during all the steps of the algorithm. Moreover, we have tuned and improved our hybrid MPI-OpenMP parallelism to run simulations with large number of radiation induced defects forming our dislocation network. To obtain a good scalability, we have introduced a better load balancing at thread level as well as process level. By combining efficient data structure and hybrid parallelism we obtained a speedup of 112 on 160 cores for a simulation of half a million of segments.

These contributions have been presented in minisymposia at the 11th World Congress on Computational Mechanics [47], 7th MMM International Conference on Multiscale Materials Modeling [25], [61] and at the International Workshop on DD simulations [62].
This work is developed in the framework of the ANR OPTIDIS.

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

6.6.2.1. MHD instabilities edge localized modes

The last contribution of Xavier Lacoste's thesis deals with the integration of our work in JOREK, a production controlled plasma fusion simulation code from CEA Cadarache. We described a generic finite element oriented distributed matrix assembly and solver management API. The goal of this API is to optimize and simplify the construction of a distributed matrix which, given as an input to PaStiX, can improve the memory scaling of the application. Experiments exhibit that using this API we could reduce the memory consumption by moving to a distributed matrix input and improve the performance of the factorized matrix assembly by reducing the volume of communication. All this study is related to PaStiX integration inside JOREK but the same API could be used to produce a distributed assembly for another solver or/and another finite elements based simulation code.

6.6.2.2. Turbulence of plasma particules inside a tokamak

Concerning the GYSELA global non-linear electrostatic code, the efforts during the period have concentrated on predicting memory requirement and on the gyroaverage operator.

The Gysela program uses a mesh of 5 dimensions of the phase space (3 dimensions in configuration space and 2 dimensions in velocity space). On the large cases, the memory consumption already reaches the limit of the available memory on the supercomputers used in production (Tier-1 and Tier-0 typically). Furthermore, to implement the next features of Gysela (e.g. adding kinetic electrons in addition to ions), the needs of memory will dramatically increase, the main unknown will represents hundreds of TB. In this context, two tools were created to analyze and decrease the memory consumption. The first one is a tool that plots the memory consumption of the code during a run. This tool helps the developer to localize where the memory peak is located. The second tool is a prediction tool to compute the peak memory in offline mode (for production use mainly). A post processing stage combined with some specific traces generated on purpose during runtime allow the analysis of the memory consumption. Low-level primitives are called to generate these traces and to model memory consumption : they are included in the libMTM library (Modeling and Tracing Memory). Thanks to this work on memory consumption modeling, we have decreased the memory peak of the GYSELA code up to 50 % on a large case using 32,768 cores and memory scalability improvement has been shown using these tools up to 65k cores.

The main unknown of the Gysela is a distribution function that represents either the density of the guiding centers, either the density of the particles in a tokamak (depending of the location in the code). The switch between these two representations is done thanks to the gyroaverage operator. In the actual version of Gysela, the computation of this operator is achieved thanks to the so-called Padé approximation. In order to improve the precision of the gyroaveraging, a new implementation based on interpolation methods has been done (mainly by researchers from the Inria Tonus project-team and IPP Garching). We have performed the integration of this new implementation in GYSELA and also some parallel benchmarks. However, the new gyroaverage operator is approximatively 10 times slower than the original one. Investigations and optimizations on this operator are still a work in progress.
This work is carried on in the framework of Fabien Rozar's PhD in collaboration with CEA Cadarache.

6.6.2.3. SN Cartesian solver for nuclear core simulation

High-fidelity nuclear power plant core simulations require solving the Boltzmann transport equation. In discrete ordinate methods, the most computationally demanding operation of this equation is the sweep operation. Considering the evolution of computer architectures, we propose in this work, as a first step toward heterogeneous distributed architectures, a hybrid parallel implementation of the sweep operation on top of the generic task-based runtime system: PaRSEC. Such an implementation targets three nested levels of parallelism: message passing, multi-threading, and vectorization. A theoretical performance model was designed to validate the approach and help the tuning of the multiple parameters involved in such an approach. The proposed parallel implementation of the Sweep achieves a sustained performance of 6.1 Tflop/s, corresponding to 33.9% of the peak performance of the targeted supercomputer. This implementation compares favorably with state-of-art solvers such as PARTISN; and it can therefore serve as a building block for a massively parallel version of the neutron transport solver DOMINO developed at EDF.
Preliminary results have been presented at the international HPCC workshop on HPC-CFD in Energy/Transport Domains [50] in Paris. The main contribution will be presented at the international conference IPDPS 2015 [33] in Hyderabad.

6.6.2.4. 3D aerodynamics for unsteady problems with moving bodies

In the first part of our research work concerning the parallel aerodynamic code FLUSEPA, a first OpenMP-MPI version based on the previous one has been developped. By using an hybrid approach based on a domain decomposition, we achieved a faster version of the code and the temporal adaptive method used without bodies in relative motion has been tested successfully for real complex 3D-cases using up to 400 cores. Moreover, an asynchronous strategy for computing bodies in relative motion and mesh intersections has been developed and has been used for actual 3D-cases. A journal article (for JCP) to sum-up this part of the work is under redaction and a presentation at ISC at the "2nd International Workshop on High Performance Computing Simulation in Energy/Transport Domains" on July 2015 is scheduled.

This intermediate version exhibited synchronization problems for the aerodynamic solver due to the time integration used by the code. To tackle this issue, a task-based version over the runtime system StarPU is currently under development and evaluation. This year was mainly devoted to the realisation of this version. Task generation function have been designed in order to maximize asynchronism in execution. Those functions respect the data pattern access of the code and led to the refactorization of the actual kernels. A task-based version is now available for the aerodynamic solver and is available for both shared and distributed memory. This work will be presented as a poster during the SIAM CSE'15 conference and we are in the process to submit a paper in the Parallel CFD'15 conference.

The next steps will be to validate the correction of this task-based version and to work on the performance of this new version on actual cases. Later, the task description should be extended to the motion and intersection operations.
This work is carried on in the framework of Jean-Marie Couteyen's PhD in collaboration with Airbus Defence and Space Les Mureaux.

PHOENIX Project-Team

6. New Results

6.1. Highlights of the Year

- A best paper award was obtained at ASSETS 2014 (The 16th International ACM SIGACCESS Conference on Computers and Accessibility), by the 5 authors of the paper "Tablet-Based Activity Schedule for Children with Autism in Mainstream Environment" .

BEST PAPERS AWARDS :
[26] **ASSETS 2014 - The 16th International ACM SIGACCESS Conference on Computers and Accessibility**. C. FAGE, L. POMMEREAU, C. CONSEL, E. BALLAND, H. SAUZÉON.

6.2. Technological Support for Self-Regulation of Children with Autism

Children with Autism Spectrum Disorders (ASD) have difficulties to self-regulate emotions, impeding their inclusion in a range of mainstreamed environments. Self-regulating emotions has been shown to require recognizing emotions and invoking specific coping strategies.

In the context of the School+ research project, we have developed an application dedicated to self-regulating emotions in children with ASD. Ten children with ASD have experimentally tested this tablet-based application over a period of three months in a mainstreamed school. A collaborative learning approach, involving parents, teachers and a school aid, was used 1) to train students to operate the tablet and our application autonomously, and 2) to facilitate the adoption of our intervention tool.

This study shows that our application was successful in enabling students with ASD to self-regulate their emotions in a school environment. Our application helped children with autism to recognize and name their emotions, and to regulate them using idiosyncratic, parent-child, coping strategies, supported by multimedia contents.

This work is in the context of the School+ national research project funded by the French Ministry of National Education.

A best paper award was obtained at ASSETS 2014 (The 16th International ACM SIGACCESS Conference on Computers and Accessibility) for this work in October 2014, by the 5 authors of the paper "Tablet-Based Activity Schedule for Children with Autism in Mainstream Environment" [26]: Charles Fage, Léonard Pommereau, Charles Consel, Emilie Balland, and Hélène Sauzéon.

6.3. A Low-Cost approach to the Verification of Daily Activities of Elders

Activities of Daily Living (ADL) are abilities defining the functional status of an individual. Verifying what ADLs are performed by an elder is a decisive factor to determine what kinds and what levels of assistance are needed for an individual and whether aging in place is desirable. The importance of this issue has led a number of researchers to develop a range of Ubicomp approaches that can monitor activities.

In this study, we take these prior results one step further and apply them to the needs of caregiver professionals to monitor elders at their home. Specifically, our approach relies on the following key observation: as people age their daily activities are increasingly organized according to a routine to optimize their daily functioning. As a result, their activities do not need to be recognized but should rather be verified. Deviations are a warning sign of degradation.

We have developed an approach to activity verification. This approach relies on a technological infrastructure that is simple, low-cost and non-intrusive. This infrastructure was deployed in four homes of elders of 83 years of age on average. The same set of sensors was used in the four homes and was placed at strategic locations with respect to their routines to verify the target activities. The analysis of the data collected during five weekdays show that they follow very strict routines that can easily be associated with their main activities.

This work is in the context of the DomAssist project, funded by the following partners: UDCCAS, CG33, CRA, CNSA, Chambre des métiers. A report of the work has been published at the 16th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2014) [25].

6.4. Using virtual reality for studying everyday-like memory and its cognitive correlates

This work consisted in a pilot-study with a comparison approach between aging and traumatic brain injury (TBI) to investigate everyday object memory patterns using a virtual HOMES test.

- Methods: Sixteen young controls, 15 older adults and 15 TBI patients underwent the HOMES test and traditional tests.
- Results: Older adults and TBI patients exhibited similar HOMES performances: poor recall, a greater recognition benefit, high false recognitions, but intact clustering and proactive interference effects. The age-related differences for HOMES measures were mainly mediated by executive functioning, while the HOMES performances in the TBI group were correlated with memory measures.
- Conclusion: The differential cognitive mediating effects for a similar everyday-like memory pattern have been discussed by highlighting the need for more cautious interpretations of cognitive mechanisms behind similar behavioral patterns in different populations especially in clinical and rehabilitation settings.
- Implication for Rehabilitation:
  - Virtual reality might provide ecological scenarios to assess the multiple processes of everyday memory in elderly people as well as in TBI patients.
  - A similar pattern of Everyday-like memory failures might result from different cognitive origins among different neuropsychological patients.
  - The assessment of specific cognitive origins of Everyday-like memory impairments deserves consideration for drawing up relevant rehabilitative programs that match the specific cognitive needs of patients for performing everyday memory tasks.

This work has been published in the journal "Disability and Rehabilitation: Assistive Technology" in November 2014 [14].

RUNTIME Team

6. New Results

6.1. Highlights of the Year

- This year we started very large collaborations with the BULL/Atos company. WE started one European project, one PIA french project and one PhD thesis. The amount of Person Year funded with this project exceed 10. The research we will do with Bull covers resource management, process placement, platform modeling, application modeling, affinity abstraction.
- The StarPU software is used by CEA for automatically distributing linear algebra on their cluster of 144 hybrid nodes.

6.2. Task scheduling over heterogeneous architectures

We continued our work on extending STARPU to master exploitation of Heterogeneous Platforms through dynamic task scheduling, with a now-imminent release of StarPU 1.2.

We have improved the simulation support with SIMGRID, to augment the accuracy of the simulated execution according to the hardware capabilities [30].

We have collaborated with various research projects to leverage the potential of STARPU. We have improved the support for the PASTIX and QR-MUMPS sparse matrix solvers, thus obtaining competitive performance on CPUs and on CPUs+GPUs [25]. We have improved the MPI communication engine of STARPU to get better performance with the EADS hmatrix solver.

We have obtained very good performance and scalability with a Cholesky factorization distributed over a cluster of 144 heterogeneous nodes hosted at CEA.

We have studied the theoretical performance bound that can be achieved for the Cholesky factorization, reproduced the performance of a theorically optimal scheduled, shown that the classical HEFT heuristic is far from it, that more application-specific heuristics allow to get performance closer to the peak, and that the peak is not reachable with simple heuristics, because it requires non-trivial task order inversions.

In relationship with the ADT K'Star effort of building the KLANG-OMP OpenMP compiler and putting together the KASTORS benchmark suite, StarPU has been extended to provide an OpenMP-enabled runtime support for KLANG-OMP. In particular, the StarPU OpenMP Runtime Support implements *preemptible* tasks required for OpenMP, using the concept of continuations, while maintaining interoperability with StarPU regular, non-blocking tasks, and while preserving the heterogeneous, performance model-based scheduling capabilities of StarPU.

The KLANG-OMP C/C++ OpenMP compiler co-developed with Inria Team MOAIS enables plain OpenMP applications to run un-modified on top of the StarPU runtime system, thus significantly increasing the performance portability potential of StarPU.

6.3. Modeling hierarchical platform memory performance with microbenchmarks

Bertrand PUTIGNY developed a new memory performance model based on micro-benchmarks during his PhD. He transforms parallel codes such as OpenMP into memory access skeleton before predicting memory buffer states in caches and using benchmarks outputs to predict the runtime. This model successfully predict the performance behavior of several memory-bound kernels [26].

We also used this model to study the impact of memory caches on the performance on intra-node MPI communication [27].

6.4. Static modeling of clusters of multicore and heterogeneous nodes

We improved the hwloc software to better manage clusters of nodes. This first includes the management of HPC node I/O devices by providing easy ways to retrieve the locality of GPUs and network interfaces. A scalable global view of clusters can be built by factorizing the common topology information that is usually shared by many similar nodes [20]. Finally the topology of the network assembling all these nodes can be exposed in a generic technology-independent manner using the new netloc tool [21] that is now part of hwloc.

6.5. Multithreaded communications

We have proposed a full rewrite of the PIOMAN software, to make it rely on system threads rather than on the now obsolete MARCEL thread scheduler. It makes it more portable, composable with any runtime system used for multithreading, and more scalable. We have shown [19][18] that it features good properties with regard to asynchronous communication progression and multithreaded communications in applications.

6.6. Toplogy-aware load balancing in Charm++

Charm++ implements a fine-grained paradigm based on migratable computing objects. This programming model is designed to run large-scale experiments and provide a dynamic load balancing system to optimize it. Our previous Charm++ load balancer designed for communication-bound applications was improved to scale on large platforms. More precisely, we worked on the network awareness of this algorithm by using LibTopoMap. Our topology-aware load balancing algorithm was also restructured to be parallel and distributed. These enhancements were validated on the Blue Waters supercomputer at Urbana-Champaign, IL. Finally, We have begun to carry out experiments on real application modeling seismic wave propagation.

6.7. Topology-aware ressource allocation

On the one hand SLURM already provides topology aware placement techniques to promote the choice of group of nodes that are placed on the same network level, connected under the same network switch or even placed close to each other so as to avoid long distance communications. On the other hand users can map tasks in a parallel application to the physical processors on the chosen nodes, based on the communication topology.

Our goal is to take in account, in SLURM, placement process, hardware topology, and application communication pattern too. We have implemented a new selection option for the cons_res plugin in SLURM 2.6.5. In this case the usually best fit algorithm used to choose nodes is replaced by Treematch, an algorithm to find the best placement among the free nodes list in light of a given application communication matrix. Tests and evaluation of this feature are in progress.

6.8. Scheduling of dynamic streaming applications on hybrid embedded MPSoCs

The work on the dataflow scheduler has continued so as to improve it: it is now simper and more efficient. Moreover, an H.264 video decoder implementation from STMicroelectronics has been ported onto the developed execution model to conduct more significant experiments. This application exhibits a higher level of complexity and variability, which is the reason why it is well suited for assessing the scheduler's reactivity. Furthermore, an important groundwork has been carried out to enable software support for parts of the application, which enlarges considerably the design space and allow to benefit from better flexibility. In parallel, some earlier work on list scheduling under memory constraints has been extended and published in an international journal [11].

6.9. Performance model for multithreaded applications on multi-core processors

Concerning data locality, researches have shown a tradeof in groupement strategy for process mapping. We have to deal with balanced improvement of several aspects such as threads synchronizations or resource exploitation. Weighting those criterias can only be achieved according to a certain knowledge of both the application and the machine.

Thus, we are working on modeling threads affinity and weights on machines topology to improve a placement method based on the TreeMatch algorithm using new metrics. Several experiences have lead us to the conclusion that it is very hard to identify the key hints and to understand application needs.

Consequently, we are developping a visual tool which displays hardware counters aggregated and mapped on the system topology to identify dynamically those hardware narrows during execution, and understand processes placement effects on them. We hope to achieve a better comprehension of process placement consequences on resources usage by applications.

FLOWERS Project-Team

6. New Results

6.1. Highlights of the Year

PY. Oudeyer and M. Lopes, together with J. Gottlieb (Univ. Columbia, NY) organized the first International Symposium on Neurocuriosity symposium on Information Seeking, Curiosity and Attention, pioneering a gathering of world experts on curiosity from developmental psychology, neuroscience, ethology and computational modelling (see https://openlab-flowers.inria.fr/t/first-interdisciplinary-symposium-on-information-seeking-curiosity-and-attention/21). This was achieved in the context of associated team Neurocuriosity with the cognitive neuroscience lab of J. Gottlieb at Univ. Columbia, NY, US. The first results investigating predictions of theoretical formal models of curiosity on human exploration were also published [25].

O. Mangin obtained the Best thesis poster from Bordeaux doctoral school of mathematics and computer science, for his PhD thesis "The Emergence of Multimodal Concepts: From Perceptual Motion Primitives to Grounded Acoustic Words" [24].

The team, in collaboration with Inaki Iturrate and Luis Montesano, published major results on calibration-free brain-computer interface methods, where incremental machine learning algorithms are used to remove the phase of calibration for an important family of use contexts [44] [45].

In october 2014, the team announced the release of a new version of the Poppy Project platform, dedicated in particular to the use of tools for creating and programming interactive robots in Education and Art. This platform, which is a result of research on the role of morphology in skill acquisition within ERC project Explorers, was selected as finalist for the Global Fab Awards 2014 (https://www.fab10.org/en/awards) which select the best worldwide projects in the Makers ecosystem. It was also presented in major international press and media (https://www.poppy-project.org/in-the-press/), in multiple hackatons and demos, in particular at the major international conference LeWeb (https://www.poppy-project.org/social-life/, and its video on the web was seen 125k times. Poppy Project was presented at Elysée, during a French Tech event, to François Hollande (http://www.inria.fr/centre/bordeaux/actualites/poppy-le-robot-humanoide-a-l-elysee), and in Bordeaux to Axelle Lemaire. Web site: http://www.poppy-project.org

The Flowers team made major achievements in diffusing science and technology towards the general public. The team developped the IniRobot pedagogical kit, for the discovery of computer science and robotics in primary schools. The kit was first developped and evaluated in schools, in collaboration with a group of teachers, and then began to be largely disseminated and used in september 2014 to schools in Talence, Bordeaux, Lormont, and Lille. A dedicated web site has been created, allowing all users and contributors to share their experiences with the kit: https://dm1r.inria.fr/c/kits-pedagogiques/inirobot. PY. Oudeyer was invited to give a TedX talk (https://www.youtube.com/watch?v=AP8i435ztwE, video viewed by more than 9000 people), and was interviewed and invited to talk about our research on major media channels (e.g. Le Monde, Les Echos, France Inter, see http://www.pyoudeyer.com/press/).

6.2. Robotic And Computational Models Of Human Development

6.2.1. Computational Models Of Information-Seeking, Curiosity And Attention

Participants: Manuel Lopes, Pierre-Yves Oudeyer [correspondant], Jacqueline Gottlieb, Adrien Baranes, Pierre Rouanet, Brice Miard, Jonathan Grizou.

6.2.1.1. The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration

Devising efficient strategies for exploration in large open-ended spaces is one of the most difficult computational problems of intelligent organisms. Because the available rewards are ambiguous or unknown during the exploratory phase, subjects must act in intrinsically motivated fashion. However, a vast majority of behavioral and neural studies to date have focused on decision making in reward-based tasks, and the rules guiding intrinsically motivated exploration remain largely unknown. To examine this question we developed a paradigm for systematically testing the choices of human observers in a free play context. Adult subjects played a series of short computer games of variable difficulty, and freely choose which game they wished to sample without external guidance or physical rewards. Subjects performed the task in three distinct conditions where they sampled from a small or a large choice set (7 vs. 64 possible levels of difficulty), and where they did or did not have the possibility to sample new games at a constant level of difficulty. We show that despite the absence of external constraints, the subjects spontaneously adopted a structured exploration strategy whereby they (1) started with easier games and progressed to more difficult games, (2) sampled the entire choice set including extremely difficult games that could not be learnt, (3) repeated moderately and high difficulty games much more frequently than was predicted by chance, and (4) had higher repetition rates and chose higher speeds if they could generate new sequences at a constant level of difficulty. The results suggest that intrinsically motivated exploration is shaped by several factors including task difficulty, novelty and the size of the choice set, and these come into play to serve two internal goals—maximize the subjects' knowledge of the available tasks (exploring the limits of the task set), and maximize their competence (performance and skills) across the task set. This was published in [25].

6.2.1.2. A new experimental setup to study the structure of curiosity-driven exploration in humans

We started evaluating several games that test how humans explore a space of motor tasks of different complexities. Our objective is to observe what exploratory behaviors do people use when learning a new skill. The main hypothesis we are testing is that skills that provide a larger learning progress will be favored and so we will see a progression from the simpler to the more complex skills. Surely there are individual differences and the causes and impact of those differents is a very important research topic. The Abstract Games we created allows us to create several dimensions of complexity for the games. In this task, there are several abstract forms that appear in the screen and the user is able to control them using its own body (tracked using a Kinect sensor), see Fig. 2. The relation between the degrees of freedom and the forms/colors/sizes of the shapes is arbitrary and the user must explore its body to be able to control its behavior. This was published in [58].

6.3. Life-Long Robot Learning And Development Of Motor And Social Skills

6.3.1. Exploration and learning of sensorimotor policies

6.3.1.1. Non-linear regression algorithms for motor skill acquisition: a comparison

Participants: Thibaut Munzer [correspondant], Freek Stulp, Olivier Sigaud.

Endowing robots with the capability to learn is an important goal for the robotics research community. One part of this research is focused on learning skills, where usually two learning paradigms are used sequentially. First, a robot learns a motor primitive by demonstration (or imitation). Then, it improves this motor primitive with respect to some externally defined criterion. We realized a study on how the representation used in the demonstration learning step can influence the performance of the policy improvement step. We provide a conceptual survey of different demonstration learning algorithms and perform an empirical comparison of their performance when combined with a subsequent policy improvement step. These study have been published at the JFPDA conference [61].

During this work, we have discovered that many (batch) regression algorithms (amongst others, locally weighted (projection) regression, Gaussian mixture regression, radial basis function networks, and Gaussian process regression) use only one of two underlying model representations to represent a function: a weighted sum of basis function, or a mixture of linear models. Furthermore, we show that the former is a special case of the latter. This insights provides a deep understanding of the relationship between these algorithms, that, despite being derived from very different principles, use a function representation that can be captured within one unified model. A review article on this topic has been submitted to Neural Networks.

6.3.1.2. Simultaneous On-line Discovery and Improvement of Robotic Skill Options

Participants: Freek Stulp [correspondant], Laura Herlant, Antoine Hoarau, Gennaro Raiola.

The regularity of everyday tasks enables us to reuse existing solutions for task variations. For instance, most door-handles require the same basic skill (reach, grasp, turn, pull), but small adaptations of the basic skill are required to adapt to the variations that exist (e.g. levers vs. knobs). In a joint project with Laura Herlant of Carnegie Mellon University, we developed the algorithm "Simultaneous On-line Discovery and Improvement of Robotic Skills" (SODIRS) that is able to autonomously discover and optimize skill options for such task variations. We formalize the problem in a reinforcement learning context, and use the PI^{BB} algorithm to continually optimize skills with respect to a cost function. SODIRS discovers new subskills, or "skill options", by clustering the costs of trials, and determining whether perceptual features are able to predict which cluster a trial will belong to. This enables SODIRS to build a decision tree, in which the leaves contain skill options for task variations. We demonstrate SODIRS' performance in simulation, as well as on a Meka humanoid robot performing the ball-in-cup task. This work has led to a publication at IROS [64].

6.3.1.3. Simultaneous On-line Discovery and Improvement of Robotic Skill Options

Participants: Freek Stulp [correspondant], Nicolas Alberto Torres, Michael Mistry.

Freek Stulp supervised the Master's thesis project of Nicolas Torres Alberto from the Telecom Physique Strasbourg, which led to a publication at Humanoids'14 [65]. The project focused on improving autonomy in learning inverse dynamics models for computed torque control. In computed torque control, robot dynamics are predicted by dynamic models. This enables more compliant control, as the gains of the feedback term can be lowered, because the task of compensating for robot dynamics is delegated from the feedback to the feedforward term. Previous work has shown that Gaussian process regression is an effective method for learning computed torque control, by setting the feedforward torques to the mean of the Gaussian process. We extend this work by also exploiting the variance predicted by the Gaussian process, by lowering the gains if the variance is low. This enables an automatic adaptation of the gains to the uncertainty in the computed torque model, and leads to more compliant low-gain control as the robot learns more accurate models over time. On a simulated 7-DOF robot manipulator, we demonstrate how accurate tracking is achieved, despite the gains being lowered over time. This is a first step towards life-long learning robots, that continuously and autonomously adapt their control parameters (feedforward *and* feedback) over extended periods of time.

6.3.1.4. Learning manipulation of flexible tools

Participants: Clément Moulin-Frier [correspondant], Marie-Morgane Paumard, Pierre Rouanet.

Clément Moulin-Frier and Pierre Rouanet supervised the internship of Marie-Morgane Paumard from the *Ecole Normale Supérieure de Cachan*, at the Bachelor level. The internship has been realized from May to August 2014. Her report is entitled *Learning the manipulation of flexible tools in developmental robotics: a fishing robot* and is available at this address: https://flowers.inria.fr/clement_mf/files/Paumard_RapportDeStage.pdf.

Learning how to manipulate flexible tools is an harsh issue in robotics, since there is generally no analytical model of the system dynamics available. Learning algorithms are therefore a pivotal tool to control such systems. Marie-Morgane conceived an experiment on the manipulation of a fishing rod by a 2-arm robot equipped with a movement generation and perceptual systems. She studied how an optimization algorithm allows the robot to reach particular position of the hook on the floor. Then, she analyzed the distribution of effects (i.e. final fishhook position) in different contexts as well as optimization performances for particular goals.

6.3.1.5. Learning how to reach various goals by autonomous interaction with the environment: unification and comparison of exploration strategies

Participants: Clément Moulin-Frier [correspondant], Pierre-Yves Oudeyer.

In the field of developmental robotics, we are particularly interested in the exploration strategies which can drive an agent to learn how to reach a wide variety of goals. We unified and compared such strategies, recently shown to be efficient to learn complex non-linear redundant sensorimotor mappings. They combine two main principles. The first one concerns the space in which the learning agent chooses points to explore (motor space vs. goal space). Previous works (Rolf et al., 2010; Baranes and Oudeyer, 2012) have shown that learning redundant inverse models could be achieved more efficiently if exploration was driven by goal babbling, triggering reaching, rather than direct motor babbling. Goal babbling is especially efficient to learn highly redundant mappings (e.g the inverse kinematics of a arm). At each time step, the agent chooses a goal in a goal space (e.g uniformly), uses the current knowledge of an inverse model to infer a motor command to reach that goal, observes the corresponding consequence and updates its inverse model according to this new experience. This exploration strategy allows the agent to cover the goal space more efficiently, avoiding to waste time in redundant parts of the sensorimotor space (e.g executing many motor commands that actually reach the same goal). The second principle comes from the field of active learning, where exploration strategies are conceived as an optimization process. Samples in the input space (i.e motor space) are collected in order to minimize a given property of the learning process, e.g the uncertainty (Cohn et al., 1996) or the prediction error (Thrun, 1995) of the model. This allows the agent to focus on parts of the sensorimotor space in which exploration is supposed to improve the quality of the model. In [59], we have shown how an integrating probabilistic framework allows to model several recent algorithmic architectures for exploration based on these two principles, and compare the efficiency of various exploration strategies to learn how to uniformly cover a goal space. This was published in [59].

6.3.1.6. Reusing Motor Commands to Learn Object Interaction

Participants: Fabien Benureau [correspondant], Pierre-Yves Oudeyer.

We have proposed the Reuse algorithm, that exploit data produced during the exploration of an first environment to efficiently bootstrap the exploration of second, different but related environment. The effect of the Reuse algorithm is to produce a high diversity of effects early during exploration. The algorithm only constrains the environments to share the same motor space, and makes no assumptions about learning algorithms or sensory modalities. We have illustrated our algorithm on a 6-joints robotic arm interacting with a virtual object, and showed that our algorithm is robust to dissimilar environments, and significantly improves the early exploration of similar ones. This was published in [34].

6.3.1.7. Socially Guided Intrinsic Motivation for Robot Learning of Motor Skills

Participants: Mai Nguyen [correspondant], Pierre-Yves Oudeyer.

We have presented a technical approach to robot learning of motor skills which combines active intrinsically motivated learning with imitation learning. Our architecture, called SGIM-D, allows efficient learning of high-dimensional continuous sensorimotor inverse models in robots, and in particular learns distributions of parameterised motor policies that solve a corresponding distribution of parameterised goals/tasks. This is made possible by the technical integration of imitation learning techniques within an algorithm for learning inverse models that relies on active goal babbling. In an experiment where a robot arm has to learn to use a flexible fishing line , we have illustrated that SGIM-D efficiently combines the advantages of social learning and intrinsic motivation and benefits from human demonstration properties to learn how to produce varied outcomes in the environment, while developing more precise control policies in large spaces. This was published in [28].

6.3.1.8. A social learning formalism for learners trying to figure out what a teacher wants them to do

Participants: Thomas Cederborg [correspondant], Pierre-Yves Oudeyer.

We have elaborated a theoretical foundation for approaching the problem of how a learner can infer what a teacher wants it to do through strongly ambiguous interaction or observation. This groups the interpretation of a broad range of information sources under the same theoretical framework. A teacher's motion demonstration, eye gaze during a reproduction attempt, pushes of good/bad buttons and speech comment are all treated as specific instances of the same general class of information sources. These sources all provide (partially and ambiguously) information about what the teacher wants the learner to do, and all need to be interpreted concurrently. We introduce a formalism to address this challenge, which allows us to consider various strands of previous research as different related facets of a single generalized problem. In turn, this allows us to identify important new avenues for research. To sketch these new directions, several learning setups were introduced, and algorithmic structures are introduced to illustrate some of the practical problems that must be overcome. This was published in [26].

6.3.2. Task learning from social guidance

6.3.2.1. Inverse Reinforcement Learning in Relational Domains

Participants: Thibaut Munzer [correspondant], Bilal Piot, Mathieu Geist, Olivier Pietquin, Manuel Lopes.

We introduced a first approach to the Inverse Reinforcement Learning (IRL) problem in relational domains. IRL has been used to recover a more compact representation of the expert policy leading to better generalize among different contexts. Relational learning allows one to represent problems with a varying number of objects (potentially infinite), thus providing more generalizable representations of problems and skills. We show how these different formalisms can be combined by modifying an IRL algorithm (Cascaded Supervised IRL) such that it handles relational domains. Our results indicate that we can recover rewards from expert data using only partial knowledge about the dynamics of the environment. We evaluate our algorithm in several tasks and study the impact of several experimental conditions such as: the number of demonstrations, knowledge about the dynamics, transfer among varying dimensions of a problem, and changing dynamics.

6.4. Autonomous And Social Perceptual Learning

6.4.1. Unsupervised and online non-stationary obstacle discovery and modelling using a laser range finder

Participants: Guillaume Duceux, David Filliat [correspondant].

Recognizing objects is an important capability for assistance robots, but most methods rely on vision and a heavy training procedures to be able to recognize some objects. Using laser range finders has shown its efficiency to perform mapping and navigation for mobile robots. However, most of existing methods assume a mostly static world and filter away dynamic aspects while those dynamic aspects are often caused by non-stationary objects which may be important for the robot task. We propose an approach that makes it possible to detect, learn and recognize these objects through a multi-view model, using only a planar laser range finder. We show using a supervised approach that despite the limited information provided by the sensor, it is possible to recognize efficiently up to 22 different object, with a low computing cost while taking advantage of the large field of view of the sensor. We also propose an online, incremental and unsupervised approach that make it possible to continuously discover and learn all kind of dynamic elements encountered by the robot including people and objects. These results have been published at the IROS conference [40].

6.4.2. Task oriented representations by discriminative modulation of a generative learning method

Participants: Mathieu Lefort, Alexander Gepperth [correspondant].

PROPRE (which stands for PROjection - PREdiction) is a generic and modular unsupervised neural learning paradigm that extracts meaningful concepts of multiple data flows based on predictability across stimuli. It consists on the combination of three modules. First, a topological projection of each data flow on a self-organizing map. Second, a decentralized prediction of each projection activity from each other map activities. Third, a predictability measure that quantifies the prediction error. This measure is used to modulate the projection learning so that to favor the mapping of predictable stimuli across data flows. This model was applied to the visual supervised classification of the pedestrian orientation. The modulation of the visual representation learning by the predictability measure (quantifying the ability to detect the orientation of the pedestrian) improves significantly classification performances of the system independently of the predictability measure used [55]. Moreover, PROPRE provides a combination of interesting functional properties, such as online and incremental learning [56].

6.4.3. *Learning of multimodal representations based on the self-evaluation of their predictability power*

Participants: Mathieu Lefort, Thomas Kopinski, Alexander Gepperth [correspondant].

PROPRE paradigm (see section 6.4.2) was also applied to the classification of gestures caught from two time-of-flight (ToF) cameras. In this context, the predictability measure acts as a self-evaluation module that biases the learned representations towards stimuli correlated across modalities, i.e. related to the ability of one camera to predict the other one. We show in [57] that this unsupervised multimodal representations learning improves the gesture recognition performance, compared to isolated camera representations learning, even not as much as supervised one.

6.4.4. *Resource-efficient online learning of classification and regression tasks*

Participants: Mathieu Lefort, Thomas Kopinski, Thomas Hecht, Alexander Gepperth [correspondant].

This activity investigates the coupling of generative and discriminative learning (SOM and regression) to achieve incremental learning that stays resource-efficient when the number of input and output dimensions is high. On the one hand, we apply this technique to sensory classification problems where input dimensionalities can exceed 10000 in the presence of multiple categories. On the other hand, we target the learning of forward and inverse regression models for robotics, possibly combining proproceptive with sensory information which again leads to high data dimensionality. A special kind of regression task we consider in this context is optimal integration of sensory information, where the most likely underlying value must be inferred from several noisy sensor readings. In contrast to popular approaches like XCF or LWPR, our approach achieves efficiency by avoiding a precise partitioning of the input space, relying on a dimensionality-reduced topological projection of the input space instead. While this achives slightly inferior results on standard benchmarks, we can treat high-dimensional incremental learning problems that are inaccessible to other algorithms, and especially to LWPR. This activity has resulted in two submissions to ESANN 2015 and one to IEEE Transactions on Autonomous Mental Development.

6.4.5. *Indoor semantic mapping on a mobile robot*

Participants: Louis-Charles Caron [correspondant], Alexander Gepperth, David Filliat.

Semantic mapping is the act of storing high-level information in a persistent map of the environment. The semantic information considered here is the identity of objects encountered by a mobile robot in an indoor environment [35]. The robot runs a SLAM algorithm and builds a map using a laser range finder. The semantic information is collected by analysing the point cloud provided by an RGB-D camera mounted on the robot. The choice of features used to describe the objects, the type of fusion and the recognition algorithm influence the overall capacity of the algorithm. Shape features perform very well, but are blind to changes in color. The fusion of different types of features can reduce the recognition rates on some objects but increases the overall figure. This increase is more significant as the number of objects to recognitze gets larger [36]. After running the object recognition algorithm, the identity of the objects is stored alongside the map. The stored information influences future recognition attempts on objects that were already seen by the robot to improve the recognition process. A 3-d map along with a snapshot and the identity of each object seen is displayed to a user.

6.5. Robot Design And Morphological Computation

6.5.1. Rapid morphological exploration with the Poppy humanoid platform.

Participants: Matthieu Lapeyre [correspondant], Steve N'Guyen, Alexandre Le Falher, Pierre-Yves Oudeyer.

In the paper [53], we discuss the motivation and challenges raised by the desire to consider the morphology as an experimental variable on real robotic platforms as well as allowing reproducibility and diffusion of research results in the scientific community. In this context, we present an alternative design and production methodology that we have applied to the conception of Poppy humanoid, the first complete 3D printed open-source and open-hardware humanoid robot. Robust and accessible, it allows exploring quickly and easily the fabrication, the programming and the experimentation of various robotic morphologies. Both hardware and software are open-source, and a web platform allows interdisciplinary contributions, sharing and collaborations. Finally we conduct an experiment to explore the impact of four different foot morphologies on the robot's dynamic when it makes a footstep. We show that such experimentation can easily be achieved and shared in couple of days at almost no cost.

6.6. Educational Technologies

6.6.1. KidLearn

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

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.

Intelligent Tutoring System (ITS) are computer environments designed to guide students in their learning. Through the proposal of different activities, it provides teaching experience, guidance and feedback to improve learning. The FLOWERS team has developed several computational models of artificial curiosity and intrinsic motivation based on research on psychology that might have a great impact for ITS. Results showed that activities with intermediate levels of complexity, neither too easy nor too difficult but just a little more difficult that the current level, provide better teaching experiences. The system is based on the combination of three approaches. First, it leverages Flowers team's recent models of computational models of artificial curiosity and intrinsic motivation based on research in psychology and neuroscience. 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. In 2014, we have run a second pilot experiment in elementary schools of Région Aquitaine, where 7-8 year old kids could learn elements of mathematics thanks to an educational software that presented the right exercises at the right time to maximize learning progress. [69], [37], [38], [39]

6.6.2. Education and the Poppy project

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

The Poppy platform was initially designed for research purposes and even more specifically for studying biped locomotion and human-robot interaction. However, it has been designed with open science goals in mind, both to share our research and create tools for researchers. As we are convinced of the need for multidisciplinary contributions in order to improve the state of the art in the robotics field, we decided right from the beginning to use and create modern and easy-to-use tools. This choice has strongly affected the way we designed our platform. Indeed, being simple to use, easily reproducible and hackable, modular, 3D printable and as plug 'n play as possible lead to the development of hardware (Poppy) and software (pypot) tools that can be also used by non-expert people.

Thus Poppy meets a growing societal need: education and training in technologies combining computer science, electronics and mechanics, as well as a training tool for the emergent revolutionary 3D printing process. Since October 2013 (open source release), we have been contacted by several Fablabs, universities, engineering schools and even high schools. We have had the opportunity to meet with educational teams and it appears they are looking for new motivational tools for group projects.

In this context, the Poppy platform appears well suited. Indeed, it integrates advanced and yet easily accessible techniques (3D printing, Arduino, Python) in an embodiment that motivates students and the wider public. With its openness, design and rather low-cost, Poppy is highly hackable and provides a unique context for learning and experimenting with these technologies in a Do-It-Yourself (DIY) way.

The paper [54] describes the use of the Poppy platform as a tool for scientific researches as well as educationnal and artistic applications.

Several experiments with Poppy in middle and high schools, science museums and Fablabs in France and abroad are already underway and will be discussed in the upcoming **Partnerships and Cooperations** sections.

6.6.3. Expression of emotions with Poppy Humanoid

Participants: Fabien Benureau [correspondant], Matthieu Lapeyre.

Two students in 3rd year of the Cognitive Science major at the University of Bordeaux led a TER project this year using Poppy under the supervision of Fabien Benureau, exploring how the attitude towards robots influences how humans recognise the emotion they try to express. Poppy having no facial expression — or face — yet, the students expressed the five expressions they selected (anger, surprise, joy, sadness, disgust) with body movements alone. They videotaped the sequences of movements (videos are available here http://python.sm.u-bordeaux2.fr/ter/2014/sc/desprez-zerdoumi/?page_id=289) and created an experiment asking volunteers to guess which emotion was displayed. The form also included the Negative Attitude towards Robots Scale (NARS), to investigate the possible correlation between fear of robot and the ability to identify their emotional attitude. The results showed no correlation between the two, although it was admitted that the experiment would have to be improved and ran again before any conclusion could be made.

6.7. Interactive Learning and user adaptation

6.7.1. Interactive learning from unlabeled instructions

Participants: Grizou Jonathan [correspondant], Itturate Inaki, Montesano Luis, Pierre-Yves Oudeyer, Manuel Lopes.

Interactive learning deals with the problem of learning and solving tasks using human instructions. It is common in human-robot interaction, tutoring systems, and in human-computer interfaces such as brain-computer ones. In most cases, learning these tasks is possible because the signals are predefined or an ad-hoc calibration procedure allows to map signals to specific meanings. In this work, we addressed the problem of simultaneously solving a task under human feedback and learning the associated meanings of the feedback signals. This has important practical application since the user can start controlling a device from scratch, without the need of an expert to define the meaning of signals or carrying out a calibration phase. We proposed an algorithm that simultaneously assign meanings to signals while solving a sequential task under the assumption that both, human and machine, share the same a priori on the possible instruction meanings and the possible tasks. This work was published in a conference paper [45] and a journal paper will be submitted in January 2015.

We communicated about this work to the human-robot interaction (HRI) community. A robot equiped with our algortihm would be able to interact with a human without knowing in advance the specific communicative signals used by the human. This work was published in the HRI Pionneer workshop [46].

This work was presented during the thesis defense of Jonathan Grizou entitled: Learning from Unlabeled Interaction Frames, on October 24, 2014. The video, slides, and thesis manuscript can be found at: http://jgrizou.com/projects/thesis-defense/

6.7.2. Calibration-Free BCI Based Control

Participants: Grizou Jonathan [correspondant], Itturate Inaki, Montesano Luis, Pierre-Yves Oudeyer, Manuel Lopes.

We applied previous work on interactive learning from unlabeled instructions [45] to Brain-Machine Interaction problem, leading to a Calibration-Free brain computer interfaces. So far in such brain-computer interfaces (BCI), an explicit calibration phase was required to build a decoder that translates raw electroencephalography signals from the brain of each user into meaningful instructions. Our method removes the calibration phase, and allows a user to control a device to solve a sequential task. We performed experiments where four users use BCI to control an agent on a virtual world to reach a target without any previous calibration process. Our approach is promising for the deployments of BCI applications out of the labs. This work was published in a conference paper [44] and a journal paper will be submitted in January 2015.

This work was presented during the thesis defense of Jonathan Grizou entitled: Learning from Unlabeled Interaction Frames, on October 24, 2014. The video, slides, and thesis manuscript can be found at: http://jgrizou.com/projects/thesis-defense/

6.8. Studying the Co-Construction of Interaction Protocols in Collaborative Tasks with Humans

6.8.1. Experimental Setups for User Study of Alignment in Asymmetric Interactions

Participants: Anna-Lisa Vollmer [correspondant], Jonathan Grizou, Manuel Lopes, Katharina Rohlfing, Pierre-Yves Oudeyer.

In interaction, humans align and effortlessly create common ground in communication, allowing efficient collaboration in widely diverse contexts. Robots are still far away from being able to adapt in such a flexible manner with non-expert humans to complete collaborative tasks. Challenges include the capability to understand unknown feedback or guidance signals, to make sense of what they refer to depending on their timing and context, and to agree on how to organize the interaction into roles and turns.

As a first step in approaching this issue, we investigated the processes used by humans to negotiate a protocol of interaction when they do not already share one. We developed a new experimental setup, where two humans have to collaborate to solve a task. The channels of communication they can use are constrained and force them to invent and agree on a shared interaction protocol in order to solve the task. These constraints allow us to analyze how a communication protocol is progressively established through the interplay and history of individual actions.

We consider a remote construction task, where one user (user A) knows what to build but do not have access to the construction site while its partner (user B) is at the site but do not know what to do. By constraining the communicative channel between the two partners, we study how, and if, they will agree on a similar set of signals to convey information and what type of information they tend to produce. The experimental setup consist of box with button, a video recording system and two screens. User A can send signals to user B by pressing buttons (fig. 14). Signals are displayed on a screen (fig. 14) at user B side. User A is not aware of what is displayed on user B screen, neither user B is aware of the relation between button presses and screen events. The video of user B construction scene is streamed to a screen at user B side. The task consist of bulding arbitrary construction (fig. 14) using colored toy bricks (fig. 14).

The various data recolted during these interaction sequences (fig. 15) allow us to study the Co-Construction of Interaction Protocols. This work was published in a conference paper [66].

6.9. Other

6.9.1. A Framework for Proactive Assistance

Participants: Alexandre Armand, David Filliat [correspondant].

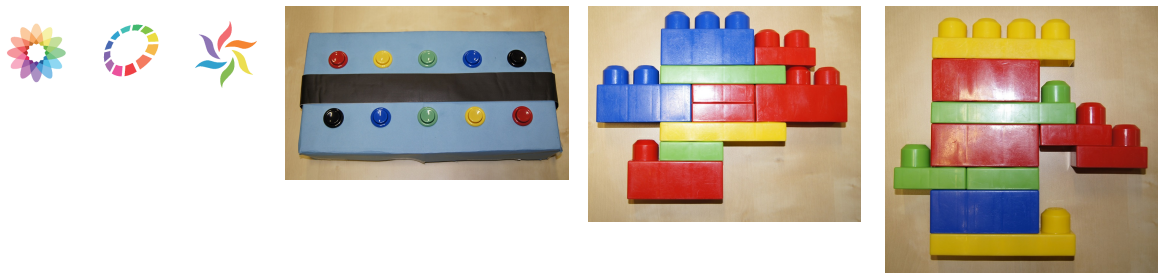

Figure 14. Three examples of sign displayed on the learner screen; The box and the button use as an interface for the teacher to communicate with the learner; Examples of construction presented to the teacher.

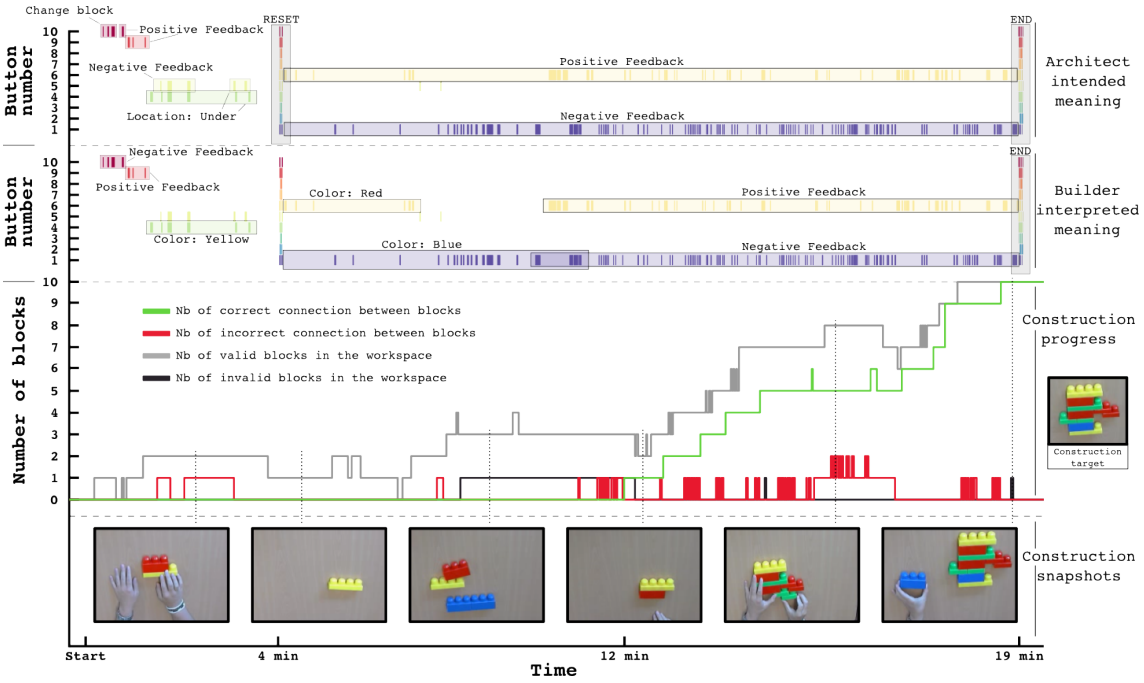

Figure 15. Timeline for one experiment of an architect and a builder collaborating towards building the construction target (right hand side). The top and middle part show the timeline of button presses associated with the intended meaning from the architect (top) and the understood meaning from the builder (middle).

We worked in collaboration with Renault on the problems of adapting driving assitance systems by learning individual drivers behaviours and of integrating more advanced perception in these systems. Advanced Driving Assistance Systems usually provide assistance to drivers only once a high risk situation has been detected. Indeed, it is difficult for an embedded system to understand driving situations, and to predict early enough that it is to become uncomfortable or dangerous. Most of ADAS work assume that interactions between road entities do not exist (or are limited), and that all drivers react in the same manner in similar conditions. We propose a framework that enables to fill these gaps. On one hand, an ontology which is a conceptual description of entities present in driving spaces is used to understand how all the perceived entities interact together with the subject vehicle, and govern its behavior. On the other hand, a dynamic Bayesian Network enables to estimate the driver situation awareness with regard to the perceived objects, based on the ontology inferences, map information, driver actuation and learned driving style. This work was published in a workshop [33] and a conference paper [32].

MANAO Project-Team

5. New Results

5.1. Highlights of the Year

We are still developping our expertise in fitting techniques. As an illustration, we have solved of a long-standing problem in fluid capture: the non-invasive three-dimensional digitization of dynamic gas flows *including their three-dimensional velocity fields* [17] (cf. Figure 8). We solve the three-dimensional flow tracking problem by fitting a full 3D Navier-Stokes simulation to the acquired data. To our knowledge, this is a world-first in this area that considerably improves the results by incorporating high-level prior knowledge into the estimate. The resulting mathematical framework can be generalized easily and lends itself to editing operations. The technique has applications, e.g., in aerospace engineering. We are exploring the possibilities with ONERA, the French space agency. In fact, parts of the developed techniques have been validated by them and are now being installed in a wind tunnel facility for real-world tests.

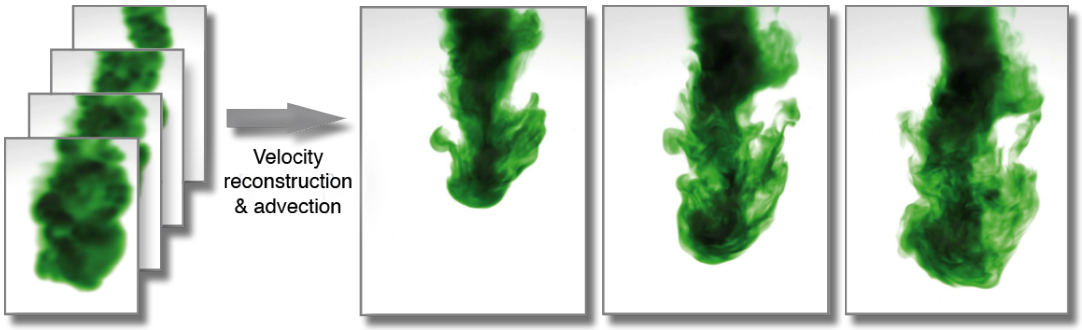

Figure 8. Low-resolution captures obtained by tomographic scanning (left) are used as inputs to our method which estimates physically plausible dense velocity fields. Such velocity fields fully determine the fluid state and can be applied in a variety of applications including fluid super-resolution (right) allowing capture to be integrated into pipelines for visual effects simulation.

This year, the collaboration between Optics and Computer Graphics has grown to a now long-term project, under the initiative of the MANAO team. First, from an institutional point of view, a framework agreement has been signed the 10th of July 2014 between the IOGS and Inria. This is an important and institutional recognition of the potential trans-disciplinary impacts of our work. Second, we have begun to set-up the COEL experimentation facility inside the LP2N laboratory. It has been made possible thanks to the support of the "Région Aquitaine" and upcoming supports from l'"Initiative d'excellence de l'université de Bordeaux". With this trans-disciplinary experimentation facility – rather unique in Europe – we can now put into practice a long-term vision of the researches that we want to achieve.

In term of visibility, we managed to published our first paper in the Optics scientific community [15], highlighting our trans-disciplinary research. We have also been part of the final and transnational exhibition of the V-Must.net network of excellence: Keys2Rome - http://keys2rome.eu. It was launched simultaneously in Rome, Sarajevo, Amsterdam and Alexandria on September 23, 2014. The exhibition uses immersive technology to present and connect these regional cultures within the Roman Empire, highlighting their diversity and commonality over centuries of Roman rule. Our spatial augmented reality solution [21] was included in this event.

5.2. Analysis and Simulation

5.2.1. Importance Sampling of Realistic Light Sources

Realistic images can be rendered by simulating light transport with Monte Carlo methods. The possibility to use realistic light sources for synthesizing images greatly contributes to their physical realism. Among existing models, the ones based on environment maps and light fields are attractive due to their ability to capture faithfully the far-field and near-field effects as well as their possibility of being acquired directly. Since acquired light sources have arbitrary frequencies and possibly high dimensions (4D), using such light sources for realistic rendering leads to performance problems. We have investigated [12] how to balance the accuracy of the representation and the efficiency of the simulation (cf. Figure 9). The work relies on generating high quality samples from the input light sources for unbiased Monte Carlo estimation [74]. This is a foundation work that has leaded to new sampling techniques for physically-based rendering with time-varying environment lighting [73] and light field light sources. The results show that physically accurate rendering with realistic light sources can be achieved in real time.

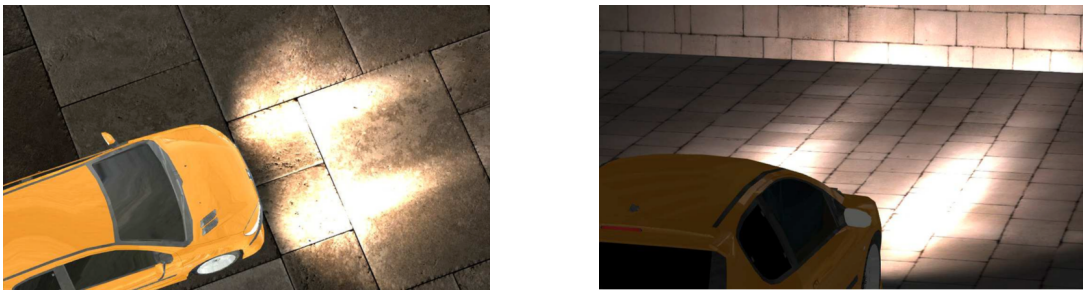

Figure 9. Our new light importance sampling technique estimates direct lighting interactively (7-9 fps) with only 200 samples per pixel that are distributed among the different images of the light field luminaire. The car headlights are represented by the same light field composed of 11 × 9 images (256 × 256 pixels).

5.2.2. Frequency Analysis of Light Scattering and Absorption

We have proposed [14] an innovative analysis of absorption and scattering of local light fields in the Fourier domain, and derived the corresponding set of operators on the covariance matrix of the power spectrum of the light field. This analysis brings an efficient prediction tool for the behavior of light along a light path in participating media. We leverage this analysis to derive proper frequency prediction metrics in 3D by combining per-light path information in the volume. Our key contribution is to show that analyzing local light fields in the Fourier domain reveals the consistency of illumination in such media, and provides a set of simple and useful rules to be used to accelerate existing global illumination methods.

5.3. Acquisition and Display

5.3.1. Three-Dimensional, Dynamic, Full State Fluid Capture and Manipulation

Participant: I. Ihrke

We have explored [17] the connection between fluid capture, simulation and proximal methods, a class of algorithms commonly used for inverse problems in image processing and computer vision. Our key finding is that the proximal operator constraining fluid velocities to be divergence-free is directly equivalent to the pressure-projection methods commonly used in incompressible flow solvers. This observation lets us treat the inverse problem of fluid tracking as a constrained flow problem all while working in an efficient, modular framework. In addition it lets us tightly couple fluid simulation into flow tracking, providing a global prior that significantly increases tracking accuracy and temporal coherence as compared to previous techniques. We demonstrate how we can use these improved results for a variety of applications, such as re-simulation, detail enhancement, and domain modification. We furthermore give an outlook of the applications beyond fluid tracking that our proximal operator framework could enable by exploring the connection of deblurring and fluid guiding.

5.3.2. Measurements and Analysis of Retro-reflective Materials

Participants: L. Belcour, R. Pacanowski

We have compared [15] performance of various analytical retro-reflecting BRDF models to assess how they reproduce accurately measured data of retro-reflecting materials. We have also introduced a new parametrization, the back vector parametrization, to analyze retro-reflecting data and we have shown that this parametrization better preserves the isotropy of data. Furthermore, we have updated existing BRDF models to improve the representation of retro-reflective data. This work was supported by the development of the ALTA library [23].

5.3.3. Kaleidoscopic Imaging

Participants: I. Reshetouski, I. Ihrke

Kaleidoscopes have a great potential in computational photography as a tool for redistributing light rays. In time-of-flight imaging the concept of the kaleidoscope is also useful when dealing with the reconstruction of the geometry that causes multiple reflections. Our work [13] is a step towards opening new possibilities for the use of mirror systems as well as towards making their use more practical. The focus of this work is the analysis of planar kaleidoscope systems to enable their practical applicability in 3D imaging tasks. We have analyzed important practical properties of mirror systems and developed a theoretical toolbox for dealing with planar kaleidoscopes. Based on this theoretical toolbox, we have explored the use of planar kaleidoscopes for multi-view imaging and for the acquisition of 3D objects [90]. The knowledge of the mirrors positions is crucial for these multi-view applications. On the other hand, the reconstruction of the geometry of a mirror room from time-of-flight measurements is also an important problem. We therefore employ the developed tools for solving this problem using multiple observations of a single scene point.

5.3.4. Interactive Spatial Augmented Reality

Participants: B. Ridel, P. Reuter, X. Granier

We have proposed the *Revealing Flashlight* [21], a new 6-degree-of-freedom interaction and visualization technique in spatial augmented reality that helps to reveal the details of cultural heritage artifacts. We locally and interactively highlight them by projecting an expressive visualization. The Revealing Flashlight can be used by archaeologists, for example, to help decipher inscriptions in eroded stones, or by museums (cf. Figure 10) to let visitors interactively discover the features and meta-information of cultural artifacts. A permanent exhibition is now running at the Allard Pierson Museum, and others museums are asking us to set-up similar installations. It was part of the final trans-European showcase of the V-MusT.net project.

5.4. Rendering, Visualization & Illustration

5.4.1. Computing Smooth Surface Contours with Accurate Topology

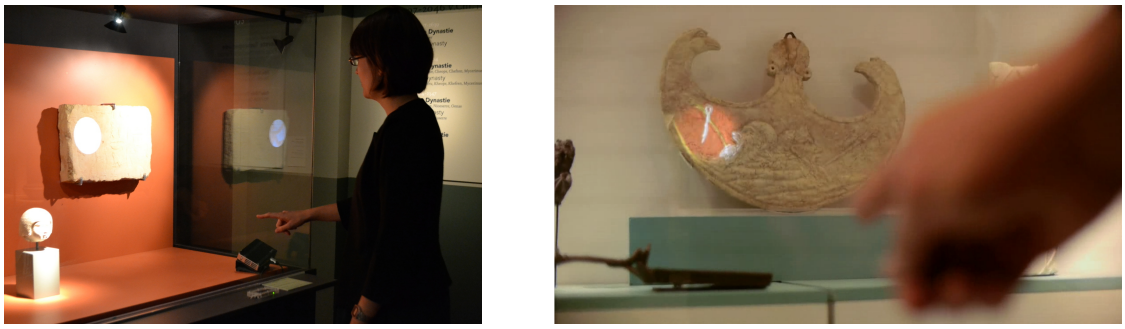

Figure 10. "The Revealing Flashlight" lets visitors explore ancient artifacts interactively. (Left) Allard Pierson Museum - Amsterdam. (Right) Keys2Rome exhibition in Museo dei Fori Imperiali - Roma.

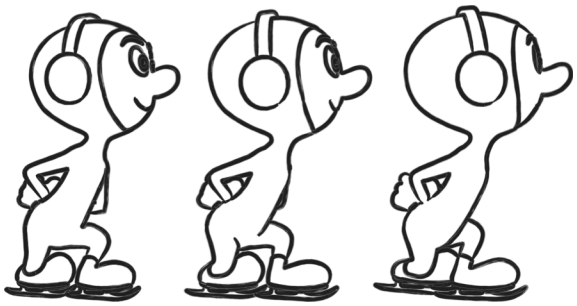

Figure 11. Contours stylized with tapered strokes [16]. Our method avoids classical breaks and gaps, producing more coherent animated strokes. Red © Pixar

We have introduced [16] a method for accurately computing the visible contours of a smooth 3D surface for stylization. This is a surprisingly difficult problem, and previous methods are prone to topological errors, such as gaps in the outline. Our approach is to generate, for each viewpoint, a new triangle mesh with contours that are topologically-equivalent and geometrically close to those of the original smooth surface. The contours of the mesh can then be rendered with exact visibility. The core of the approach is Contour-Consistency, a way to prove topological equivalence between the contours of two surfaces. Producing a surface tessellation that satisfies this property is itself challenging; to this end, we introduce a type of triangle that ensures consistency at the contour. We then introduce an iterative mesh generation procedure, based on these ideas. This procedure does not fully guarantee consistency, but errors are not noticeable in our experiments. Our algorithm can operate on any smooth input surface representation; we use Catmull-Clark subdivision surfaces in our implementation.

5.5. Editing and Modeling

5.5.1. Tomography-Based Volume Painting

Participant: I. Ihrke

Although volumetric phenomena are important for realistic rendering and can even be a crucial component in the image, the artistic control of the volume's appearance is challenging. Appropriate tools to edit volume properties are missing, which can make it necessary to use simulation results directly. Alternatively, high-level modifications that are rarely intuitive, e.g., the tweaking of noise function parameters, can be utilized. We have introduced [18] a solution to stylize single-scattering volumetric effects in static volumes. Hereby, an artistic and intuitive control of emission, scattering and extinction becomes possible, while ensuring a smooth and coherent appearance when changing the viewpoint. Our method is based on tomographic reconstruction, which we link to the volumetric rendering equation. It analyzes a number of target views provided by the artist and adapts the volume properties to match the appearance for the given perspectives. Additionally, we describe how we can optimize for the environmental lighting to match a desired scene appearance, while keeping volume properties constant. Finally, both techniques can be combined. We demonstrate several use cases of our approach and illustrate its effectiveness.

5.5.2. Implicit Skinning

Participant: G. Guennebaud

In collaboration with IRIT (Toulouse), we extended our *implicit skinning* method to a new approach for interactive character skinning called *elastic implicit skinning*. The method simulates skin contacts between limbs as well as the effect of skin elasticity (Figure 12). In addition, we go a step further towards the automation of the rigging process: our method doesn't require the definition of skinning weights. Elastic implicit skinning takes the best features of the recent implicit skinning method, and makes it robust to extreme character movements. While keeping the idea of implicit skinning, namely approximate the character by 3D scalar fields in which mesh-vertices are appropriately re-projected, we depart from the processing pipeline used so far. Implicit skinning is history independent and uses an initial skinning solution (e.g., linear blending or dual quaternions) to correct vertex positions at each frame. Our new approach is history dependent; the mesh directly tracks the iso-surfaces of the scalar field over time. Technically our solutions include: new implicit surface composition operators and a tangential relaxation scheme derived from the as-rigid-as possible energy. This work [101] has been presented at SIGGRAPH Asia this year.

5.5.3. Multi-scale Editing

Participant: G. Guennebaud

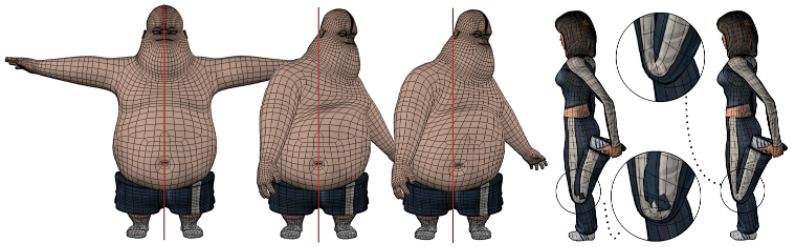

Figure 12. Illustration of the benefits of our novel elastic implicit skinning technique.

In the continuation of our Growing Least Square approach [5] for the multi-scale analysis of shape, we developed a novel tool that enables the direct editing of surface features in large point-clouds or meshes [19]. This is made possible by a novel multi-scale analysis of unstructured point-clouds that automatically extracts the number of relevant features together with their respective scale all over the surface. Then, combining this ingredient with an adequate multi-scale decomposition allows us to directly enhance or reduce each feature in an independent manner. Our feature extraction is based on the analysis of the scale-variations of locally fitted surface primitives combined with unsupervised learning techniques. Our tool may be applied either globally or locally, and millions of points are handled in real-time. The resulting system enables users to accurately edit complex geometries with minimal interaction.

5.5.4. Manipulation of Anisotropic Highlights

Participants: B. Raymond, P. Barla, G. Guennebaud, X. Granier

We have developed [20] a system for the direct editing of highlights produced by anisotropic BRDFs, which we call anisotropic highlights. We first provide a comprehensive analysis of the link between the direction of anisotropy and the shape of highlight curves for arbitrary object surfaces. The gained insights provide the required ingredients to infer BRDF orientations from a prescribed highlight tangent field. This amounts to a non-linear optimization problem, which is solved at interactive framerates during manipulation. Taking inspiration from sculpting software, we provide tools that give the impression of manipulating highlight curves while actually modifying their tangents. Our solver produces desired highlight shapes for a host of lighting environments and anisotropic BRDFs.

POTIOC Project-Team

6. New Results

6.1. Highlights of the Year

- Acceptance of the ANR project "ISAR" (Interacting with Spatial Augmented Reality) lead by Martin Hachet (Potioc)
- Publication of "Teegi" (Tangible EEG Interface) at UIST14 [15] and more than 13000 views on vimeo until December 2014 (http://vimeo.com/potioc/teegi)

6.2. Teegi -Tangible EEG Interface- and MindMirror for interactive visualization of brain activities

Participants: Jérémy Frey, Renaud Gervais, Fabien Lotte, Martin Hachet.

Typical brain activity visualization tools are usually hard to understand and interpret for novice users. With advances in neurotechnologies (notably BCI) and HCI/AR, we explored the design of new ways to visualize our own brain activity in real-time, for which we proposed two new systems.

We designed Teegi, a Tangible EEG Interface that enables novice users to get to know more about something as complex as brain signals, in an easy, engaging and informative way [15]. To this end, we have designed a new system based on a unique combination of spatial augmented reality, tangible interaction and real-time neurotechnologies (see Figure 5). With Teegi, a user can visualize and analyze his or her own brain activity in real-time, on a tangible character that can be easily manipulated, and with which it is possible to interact. Users can also reveal some specific EEG phenomenons (e.g., sensorimotor rhythms) still using a tangible approach by placing dedicated "mini-teegi" (small pupets) in a designated area on the interaction zone. The whole system has been designed with educational psychology tools in mind to ensure an efficient learning. An explorative study has shown that interacting with Teegi seems to be easy, motivating, reliable and informative. Overall, this suggests that Teegi is a promising and relevant training and mediation tool for the general public.

In addition, together with colleagues from Inria Rennes (team Hybrid), we introduced a novel augmented reality paradigm called ”the Mind-Mirror” which enables such an experience of seeing ”through your own head”, visualizing your brain ”in action and in-situ” [23]. Our approach relies on the use of a semi-transparent mirror positioned in front of a computer screen. A virtual brain is displayed on screen and automatically follows the head movements thanks to an optical face-tracking system. The brain activity is extracted and processed in real-time thanks to an EEG cap wore by the user. A rear view is also proposed thanks to an additional web-cam recording the rear of user’s head (see Figure 6).

6.3. Interaction in mobile augmented reality

Participants: Asier Marzo, Benoît Bossavit, Martin Hachet.

Nowadays, handheld devices are capable of displaying augmented environments in which virtual content overlaps reality. To interact with these environments it is necessary to use a manipulation technique. The objective of a manipulation technique is to define how the input data modify the properties of the virtual objects. Current devices have multi-touch screens that can serve as input. Additionally, the position and rotation of the device can also be used as input creating both an opportunity and a design challenge. In this project we compared three manipulation techniques which namely employ multi-touch, device position and a combination of both. A user evaluation on a docking task revealed that combining multi- touch and device movement yields the best task completion time and efficiency. Nevertheless, using only the device movement and orientation is more intuitive and performs worse only in large rotations. This work has been presented at the ACM Symposium on Spatial User Interaction 2014 [21].

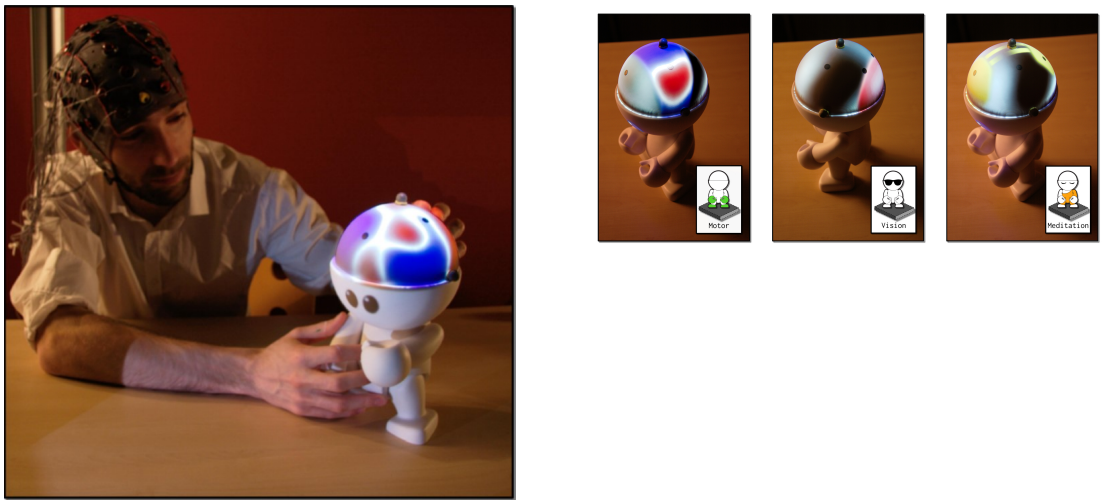

Figure 5. Teegi : a Tangible EEG Interface based on augmented reality.

Figure 6. The Mind Mirror, a new real-time visualization tool of the user's own brain activity based on augmented reality.

Figure 7. Touch (left), tilte (middle) and AR interaction techniques.

In this project we have furthermore evaluated controls based on Augmented Reality (AR), Tilt and Touch for a Point and Shoot Mobile Game (see Figure 7). A user study (n=12) was conducted to compare the three controls in terms of player experience and accuracy. Tilt and AR controls provided more enjoyment, immersion and accuracy to the players than Touch. Nonetheless, touch caused fewer nuisances and was playable under more varied situations. Despite the current technical limitations, we suggest to incorporate AR controls into the mobile games that supported them. Nowadays, AR controls can be implemented on handheld devices as easily as the more established Tilt and Touch controls. However, this study is the first comparison of them and thus its findings could be of interest for game developers. This work has been presented at ISMAR - MASH'D [22].

6.4. CurSAR: Interacting with Spatial Augmented Reality with 2D Input Devices

Participants: Renaud Gervais, Jérémy Frey, Martin Hachet.

Spatial Augmented Reality (SAR) opens interesting perspectives for new generations of mixed reality applications. Compared to traditional HCI contexts, there is little work that studies user performance in SAR. We did an experiment that compared pointing in SAR versus pointing in front of a screen using standard pointing devices (mouse and tablet). The results showed that the users tend to interact in SAR in a way that is similar to the screen condition, without a big loss of performance.

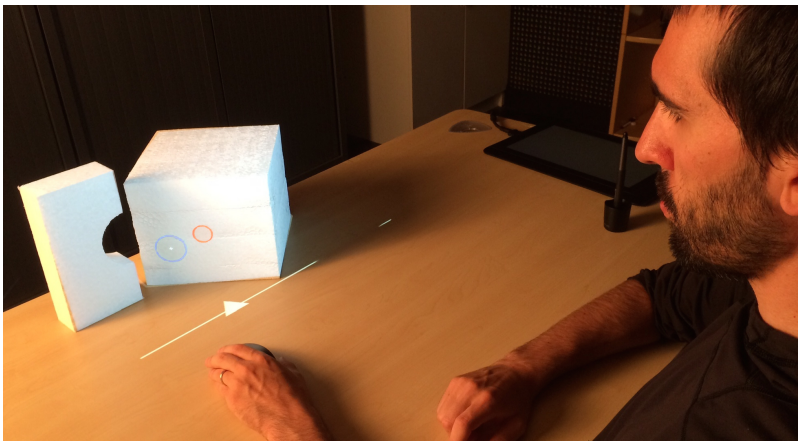

Figure 8. Pointing in spatial augmented reality

6.5. Creative Coding on Objects

Participants: Renaud Gervais, Jérémy Laviole, Asier Marzo, Martin Hachet.

In a near future scenario, we will replace some of our everyday objects with counterparts in form of Computational Objects (COs). COs look similar to the original object; however, inside them there are input sensors, output devices such as displays and a CPU. Furthermore, COs still convey the context and meaning that the original object had. For instance, a clock is associated with time and thus users could expect its CO version to display time-related data. We suggest that any user should be able to easily code new appearances and behaviors for his or her own objects. Using creative coding as a base, we propose to add the notions of affordances and conventions to this programming context. Moreover, we suggest that COs could be used as a creativity support tool although modifying their behavior beyond conventions could confuse the user. Finally, we reckon that with the proper tools, users could also make physical modifications to COs. For example, a retractile cord can be attached to the clock and be used to pull data out and display them in a linear layout.

This work has been presented as a poster at TEI 2014 [30].

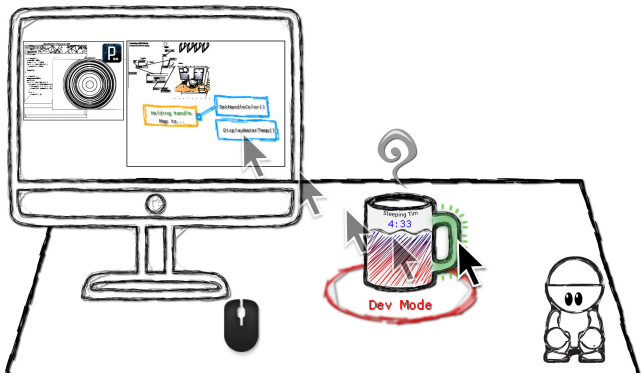

Figure 9. Sketch for a creative coding on objects scenario

6.6. Physiological sensors: bridging human-computer interaction

Participants: Jérémy Frey, Dennis Wobrock, Aurélien Appriou, Christian Mühl, Fabien Lotte, Martin Hachet.

Physiological sensors are not limited to research and medical facilities anymore. They are getting more and more affordable and they become widely accessible to users, as denoted by the popularity of smartphone apps and wearables that track heart rate during fitness activities. Before long, we may see a wide range of sensors embedded into consumer electronic devices. This trend has already started with the arrival of "smartwatches" that could – among other things – detect users' heart beats covertly.

We anticipated this opportunity in order to increase engagement in human-computer interaction, more specifically in human-agent interaction. In [14] we demonstrated that we could increase the social presence of embodied agents – that is, of virtual beings – by simply mirroring the heart rate of users. The "similarity-attraction" effect induces positive emotions toward persons or things that look like us or react as we do. An agent that is associated to a heart beat at the same pace as the user is found more sympathetic. The "similarity-attraction" effect, applied to physiological computing, could help with little effort to improve the acceptance of embodied agents and robots by the general public. (See Figure 10 for setup).

Furthermore, we have taken advantage of physiological sensors in order to evaluate different sorts of human-computer interfaces prior to their release. First, we showed that we could reliably estimate the user's mental workload levels from his/her EEG signals, across different contexts involving different levels of social stress [10]. Then, based on those results, we used a combination of electrocardiography (measure of heart beats), galvanic skin response (measure of sweat on the skin) and electroencephalography (EEG, measure of brain activity) to assess the workload of users during 3D manipulation tasks. The first preliminary results seem to indicate that we might be able to discriminate the parts of the interaction that provokes a high cognitive stress, hence that needs to be improved. This work is in line with the evaluation of visual comfort. We presented earlier this year a pilot study documenting how different virtual depths could cause different levels of discomfort [17], and how this discomfort translates to EEG activity.

Pervasive technologies and physiological computing may be a key component to bridge the gap that too often keeps dividing machines and general public. We believe that it'll help make computers more enjoyable and more usable.

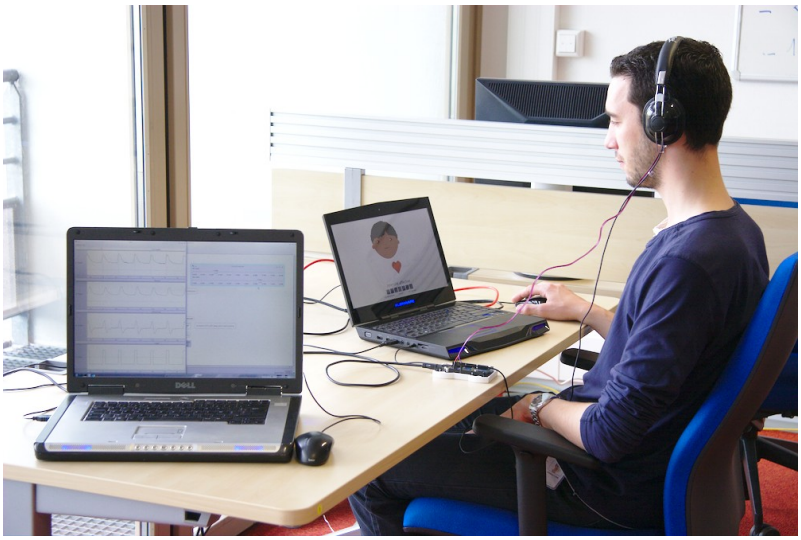

Figure 10. Experimental setup used to study the physiological "similarity-attraction" effect in human-agent interaction.

6.7. Training Approaches for Brain-Computer Interfaces

Participants: Alison Cellard, Martin Hachet, Camille Jeunet, Fabien Lotte, Christian Mühl, Julia Schumacher.

While recent research on Brain-Computer Interfaces (BCI) has highlighted their potential for many applications, they remain barely used outside laboratories due to a lack of robustness. Spontaneous BCI (i.e., mental imagery-based BCI) often rely on mutual learning efforts by the user and the machine, with BCI users learning to produce stable EEG patterns (spontaneous BCI control being widely acknowledged as a skill) while the computer learns to automatically recognize these EEG patterns, using signal processing. Most research so far was focused on signal processing, mostly neglecting the humans in the loop.

Indeed, even if it has been advocated in one of our previous publications (see activity report 2013) that current human training approaches for spontaneous BCI are most likely inappropriate, based on theoretical models, we still needed practical confirmations that users' modest performances at controlling a BCI could be partly due to these inappropriate training protocols. Thus, in our work, we proposed to study standard BCI training protocols without EEG signals, i.e., without a BCI [31]. In particular, we studied how people could learn to do two simple motor tasks using the same training tasks and feedback as the one given to motor imagery BCI users. More precisely, we asked subjects to learn to draw on a graphic tablet a triangle and a circle (the correct size, angles and speed of drawing of these two shapes being unknown to the subject) that can be recognized by the system, using a synchronous training protocol and an extending bar as feedback, like for motor imagery based BCI training. Our results show that most subjects (out of N=20 subjects) improved with this feedback and practice (i.e., the shapes they drew were increasingly more accurately recognized by the system), but that 15% of them completely failed to learn how to draw the correct shapes, despite the simplicity of the motor tasks. This suggests that part of BCI illiteracy/deficiency is likely due to the training protocols currently used.

From the huge variability in users' performances at BCI mastery emerged the following question: Why do some people manage to learn using these protocols and others do not? Our hypothesis here was that these protocols are not adapted to some users' profiles. Thus, we designed an experiment in which we looked for correlations between the personality and cognitive profile of the users and their ability to learn to control a MI-

BCI. Our current results (N=18) show that 1) performances are strongly correlated with users' spatial abilities and 2) we can reliably predict these performances using a model including different psychological factors (like abstractedness, self-reliance or tension). These results are very encouraging as they could lead to reflections about 1) exercices to improve users' spatial abilities and 2) solutions to take into account users' cognitive and personality profiles in BCI training approaches.

Furthermore, it is more and more claimed that visual feedback is not ideal for BCIs as they are conceived for interaction situations in which the visual channel is often overtaxed. Thus, tactile feedback might appear to be more relevant. In order to test this hypothesis, we proposed a study aiming at comparing a standard visual feedback with an equivalent tactile feedback in an appealing training environment containing visual distractors (to mimic an interaction context in which the visual channel is overtaxed). Users had to learn to perform motor-imagery tasks as well as a counting task, and received either a visual or vibrotactile feedback (see Figure 11). Our main result (N=18) is the fact that people receiving tactile feedback perform significantly better (at Motor-Imagery and counting task). This kind of result should encourage the BCI community to replace standard BCI protocols by more motivating training environments and multimodal feedback.

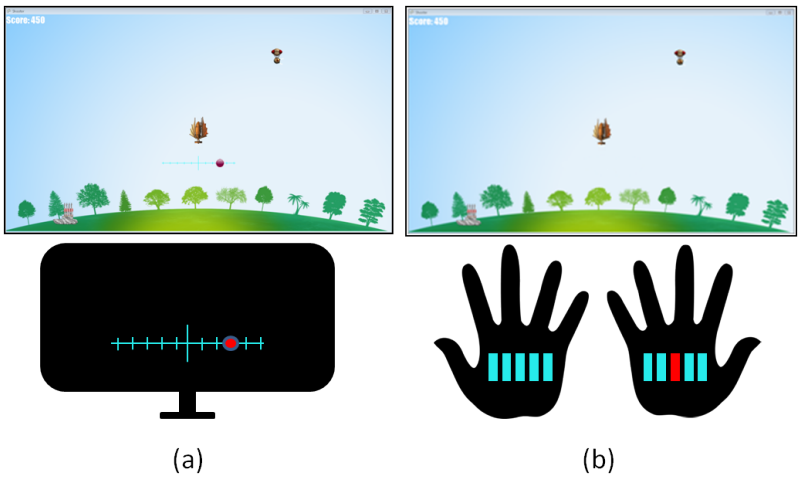

Figure 11. Illustration of the combination of an appealing training environment and a vibrotactile feedback (right, compared to a standard visual feedback, left)

Still regarding the feedback, we explored what kind of information could help the user to perform better mental imagery tasks. As such, we look for physiological features that could predict whether a mental task will be correctly recognized by the BCI, and that could be understood by the user. Among the different features we explored, it appears that the user's relaxation (from a muscular point of view), as measured in EMG activity collected by EEG channels, is one of such features. We are currently building and exploring new BCI training protocols that provide additional information about the user's muscular relaxation as complementary feedback.

6.8. EEG signal processing

Participants: Alison Cellard, Nicoletta Caramia, Fabien Lotte.

Spatial filters are powerful tools for EEG classification for BCI design, able to reduce spatial blurring effects. In particular, optimal spatial filters have been designed to classify EEG signals based on band power features. Unfortunately, there are other relevant EEG features for which no optimal spatial filter exists. This is the case for Phase Locking Value (PLV) features, which measure the synchronization between 2 EEG channels.

Therefore, we proposed to create such a pair of optimal spatial filters for PLV-features [13]. To do so, we optimized a functional measuring the discriminability of PLV-features based on a genetic algorithm. An evaluation of our algorithm on a motor imagery EEG data set showed that using optimized spatial filters led to higher classification performances, and that combining the resulting PLV features with traditional methods boosts the overall BCI performances.

We also wrote a chapter that is an introductory overview and a tutorial of signal processing techniques that can be used to recognize mental states from EEG signals in BCI [26]. More particularly, this chapter presented how to extract relevant and robust spectral, spatial and temporal information from noisy EEG signals (e.g., Band Power features, spatial filters such as Common Spatial Patterns or xDAWN, etc.), as well as a few classification algorithms (e.g., Linear Discriminant Analysis) used to classify this information into a class of mental state. It also briefly touched on alternative, but currently less used approaches.

6.9. Navigation techniques in 3D digital cities on mobile touch devices

Participants: Jacek Jankowski, Thomas Hulin, Martin Hachet.

This project was part of "Villes transparentes" research project in collaboration with Mappy (Solocal group) and Vectuel - VirtuelCity initiated in 2013. It aimed at characterizing today's most common interaction techniques for street-level navigation in 3D digital cities, for mobile touch devices, in terms of their efficiency and usability. To do so, we conducted a user study, where we compared target selection (Go-To), rate control (Joystick), position control, and stroke-based control navigation metaphors (see Figure 12). The results suggest that users performed best with the Go-To interaction technique. The subjective comments showed a preference of novices towards Go-To and expert users towards the Joystick technique. This work has been published at the 3DUI 2014 conference [18].

(a) Target selection (b) Rate control (c) Position control (d) Stroke-based control

Figure 12. Four techniques for navigating in a 3D city on a mobile touch device.)